

# Spatial Inequalities and Satellite Monitored Data: Analysis of Well-Being and Inclusive Wealth

クリバリ, チエリー, イエレマ

<https://hdl.handle.net/2324/4784589>

---

出版情報 : Kyushu University, 2021, 博士 (工学), 課程博士  
バージョン :  
権利関係 :



**Spatial Inequalities and Satellite Monitored Data: Analysis of Well-Being and Inclusive Wealth**

**Coulibaly Thierry Yerema**  
**January, 2022**

# **Spatial Inequalities and Satellite Monitored Data: Analysis of Well-Being and Inclusive Wealth**

A Thesis Submitted

In Partial Fulfillment of the Requirements

For the Degree of

**Doctor of Engineering**

By

**Coulibaly Thierry Yerema**



to the

DEPARTMENT OF CIVIL ENGINEERING

GRADUATE SCHOOL OF ENGINEERING

**KYUSHU UNIVERSITY**

Fukuoka, Japan

January, 2022

DEPARTMENT OF CIVIL ENGINEERING

GRADUATE SCHOOL OF ENGINEERING

**KYUSHU UNIVERSITY**

Fukuoka, Japan

**CERTIFICATE**

The undersigned hereby certify that they have read and recommended to the Graduate School of Engineering for the acceptance of this thesis entitled, “*Spatial economic inequalities and satellite monitored data*” by **COULIBALY Thierry Yerema** in partial fulfillment of the requirements for the degree of **Doctor of Engineering**.

Dated: January 2021

Thesis Supervisor:

---

Prof. Shunsuke MANAGI, Ph.D.

Examining Committee:

---

Prof. Kenichi TSUKAHARA, Ph.D.

---

Assoc. Prof. Yoshinao OEDA, Dr Eng.

## **ABSTRACT**

Lack of economic data at the subnational level prevents the creation and the assessment of policy on spatial inequality. Therefore, I use satellite-generated images of the surface of the earth to assess the efficacy of and the need for sub-national policy whose performance and necessity remain unknown owing to data scarcity. The robustness of the satellite images in representing human activities has been established by previous economics studies. Features such as crops, roads, buildings, or human led air pollutants and light emissions can be pictured by satellites at a high resolution. Consequently, I leverage the ability of these data to explain human activities at a subnational level to carry out five empirical economic studies assessing either the need for or the efficacy of policies addressing spatial inequality.

Thus, while Chapter 1 focus on defining the key terms of the topic, the three subsequent chapter focus on three issues that may require the creation of policies on spatial inequality. Chapter 2 explores the absence of consensus on factors influencing the impact of air pollution on well-being across and within countries. Since this lack of consensus is linked to the lack of consistent measure of pollution across space, I use monthly nitrogen dioxide changes monitored by satellite data to provide a new perspective and an eventual need for policy on this issue. Then, Chapter 3 looks into evidence of inequalities of political power, as well as disparities in investment rates and resource exploitation between areas around capital cities and the other areas in African countries. For this purpose, I assess the potential need for policy to correct these potential subnational sustainability inequities by valuing the variation of subnational capital stocks by means of satellite data. Afterwards, Chapter 4 investigates whether the provision of public services can explain the higher well-being in urban areas relative to rural areas in

developing countries. For this purpose, I compare the well-being in rural areas with that of urban areas where public service provision is low, that is informal settlements, while controlling for confounding effects such as economic activities and population density measured by satellite images.

After focusing of issues relative to potential need for policy to address spatial inequality, I use South Africa (the most unequal country worldwide) as a case study to assess the performance of policy that target spatial inequalities in two case studies. Chapter 5 assesses the relevance of policy enforced during the hosting of the World Cup 2010 in South Africa. This event was supposed to increase the income and the life conditions in host cities. For this assessment, I predict total income of South African cities during the period of the event based on satellite monitored night light data since the South African government does not monitor periodic city-level economic data. Then, Chapter 6 performs analyses relative to the South Africa government efforts to eradicate the legacy of the inequitable water policy under apartheid. This legacy may continue to impact water services in economically less developed and rural areas in South Africa. I perform an analysis using satellite night-time light data as a proxy of economic development and surface water quality at the smallest area of water management to scrutinize any income or racial inequality regarding water pollution because no economic index is disaggregated enough to enable this type of analysis. Finally, Chapter 7 relates the key findings of our work and serves as a conclusion.

## ACKNOWLEDGEMENTS

This work is dedicated to my father Yezouma Coulibaly and my mother Madeleine Coulibaly née Tamini.

I need to acknowledge that this thesis would have not existed without the support and the kindness of many people. First, I need to express my utmost gratitude to my advisor Professor Shunsuke Managi for his guidance, encouragement, and support during my study in Japan. He gave me life and professional lessons in the academic world that I would never have imagined perceiving before I arrived in Japan. I am grateful to Assistant Professor Mihoko Tegawa Wakamatsu, and Assistant Professor Moinul Islam for their active mentorship during my research activities and to Assistant Professor Chiaki Matsunaga for her kindness and help directed towards the completion of my academic life.

I would also like thank and acknowledge the support to past and present members of Managi's Laboratory that I have had the pleasure to work with and be friend with. Professor Akinori Kitsuki, Professor Fukai Hiroki, Professor Wataru Nozawa, Professor Sunbin Yoo, Professor Alexander Ryota Keeley, Professor Shuichi Tsugawa, Professor Darshana Rajapaksa, Junya Kumagai, Clarence Tolliver, Jun Xie, Chi Zhang, Binqi Zhang, Moegi Igawa, Boh Shi, Mike Li, Dyah Ika Rinawati, Zulfhazli Abdullah, Kento Komatsubara, Yasumoto Daisuke, Qiuyi Chen, Ebrahim Aly, Rafid Mahful, , Jayanti Mandasari, Kong Joo Shin, Yoshitaka Tanaka, Hiroki Onuma, Toshihiko Kitamura, Guodong Du, Shunta Komuro Liang Yuan, Kei Takahashi, Naoto Tada, and Ryota Nonaka. Thank you, as well, to the research lab support staff, Yuriko Kikuchi, Naoko Endo, Mayo Miyaishi, Naomi Kurokawa, Yumiko Ono, Mika Sugi, and Seiko Haraguchi.

Next, my gratitude goes out to the Ministry of Education, Culture, Sports, Science and Technology of Japan (MEXT) for providing the scholarship and financial support during my work in Japan.

Finally, I am indebted to my family who, through their suggestions and with their discussions shaped this dissertation in many ways. In particular, my sisters Clarisse Ouedraogo, Josette Coulibaly and Edith Coulibaly whose support enabled me to write this thesis. My love Alessandra Ferrer whose priceless contribution to this work goes beyond any description. My elder brothers in Japan, Ronald Nguele, Guillaume Faye, Yemba Baruti and Brian Omondi for their support and advice in my first steps in the world of academia.



## **TABLE OF CONTENTS**

|                                                                                                        |            |
|--------------------------------------------------------------------------------------------------------|------------|
| <b>ABSTRACT .....</b>                                                                                  | <b>ii</b>  |
| <b>TABLE OF CONTENTS .....</b>                                                                         | <b>vi</b>  |
| <b>LIST OF FIGURES .....</b>                                                                           | <b>xi</b>  |
| <b>LIST OF TABLES.....</b>                                                                             | <b>xiv</b> |
| <b>Chapter 1 Introduction.....</b>                                                                     | <b>1</b>   |
| 1.1. Spatial economic inequalities.....                                                                | 1          |
| 1.1.1 Definitions.....                                                                                 | 1          |
| 1.1.2 Lack of data to tackle spatial economic inequality .....                                         | 4          |
| 1.2. The use of satellite data in economics .....                                                      | 6          |
| 1.2.1 Description of satellite data .....                                                              | 6          |
| 1.2.2 Novelty of satellite data in economics research and issues on spatial economic inequalities..... | 7          |
| 1.3. Contributions to literature .....                                                                 | 10         |
| 1.4. The outline of the thesis.....                                                                    | 11         |
| References.....                                                                                        | 15         |
| <b>Chapter 2 The multinational and heterogeneous burden of air pollution on well-being .....</b>       | <b>16</b>  |
| 2.1. Introduction .....                                                                                | 16         |
| 2.2. Heterogeneous impact of air pollution on well-being with respect to income.....                   | 20         |
| 2.2.1 The impact of air pollution on well-being .....                                                  | 20         |
| 2.2.2 The heterogeneous impact of pollution at the individual level.....                               | 21         |
| 2.2.3 The heterogeneous impact of pollution of on well-being internationally .....                     | 23         |
| 2.3. Data.....                                                                                         | 26         |
| 2.3.1 Sample .....                                                                                     | 26         |

|                                                                                                                 |           |
|-----------------------------------------------------------------------------------------------------------------|-----------|
| 2.3.2 Measure of well-being .....                                                                               | 26        |
| 2.3.3 Factors determining well-being.....                                                                       | 28        |
| 2.4. Empirical approach.....                                                                                    | 30        |
| 2.4.1 The impact of air pollution on well-being .....                                                           | 30        |
| 2.4.2 Hypotheses testing of the heterogeneous impact of air pollution with regard to income .....               | 36        |
| 2.5. Results and discussion .....                                                                               | 37        |
| 2.5.1 The impact of air pollution on well-being .....                                                           | 37        |
| 2.5.2 Heterogeneous impacts of changes in pollution with regard to income.....                                  | 42        |
| 2.5.3 Heterogeneous impacts of changes in pollution with regard to national income characteristics.....         | 43        |
| 2.5.4 Policy implications .....                                                                                 | 47        |
| 2.6. Conclusions .....                                                                                          | 49        |
| References.....                                                                                                 | 51        |
| Appendix 2 .....                                                                                                | 56        |
| Appendix 2 - Section A: Supplementary tables and figures.....                                                   | 56        |
| Appendix 2 - Section B: The impact of air pollution on affective well-being: .....                              | 77        |
| Appendix 2 - Section C: Measuring well-being as a single item.....                                              | 82        |
| <b>Chapter 3 An evaluation of the sub-national difference in inclusive wealth sustainability in Africa.....</b> | <b>84</b> |
| 3.1. Introduction .....                                                                                         | 84        |
| 3.2. Methods.....                                                                                               | 88        |
| 3.2.1 Estimation of Inclusive wealth.....                                                                       | 88        |
| 3.2.1.1 Natural capital.....                                                                                    | 91        |
| ● Non-renewable natural capital .....                                                                           | 91        |
| ● Renewable natural capital.....                                                                                | 93        |
| 3.2.1.2. Human capital .....                                                                                    | 95        |

|                                                                                                                     |            |
|---------------------------------------------------------------------------------------------------------------------|------------|
| 3.2.1.3 The manufactured capital .....                                                                              | 96         |
| 3.2.2 Difference in IW growth between key and secondary provinces and factors<br>underlying it.....                 | 99         |
| 3.3. Results.....                                                                                                   | 100        |
| 3.3.1 The spatial distribution of African wealth .....                                                              | 100        |
| 3.3.2 Sustainability of wealth in African provinces and the difference between key<br>and secondary provinces ..... | 107        |
| 3.3.2.1. Descriptive analysis.....                                                                                  | 107        |
| 3.3.2.2. Differences in IW growth between key and secondary provinces .....                                         | 111        |
| 3.4. Discussions and conclusion.....                                                                                | 117        |
| References.....                                                                                                     | 121        |
| Appendix 3 .....                                                                                                    | 126        |
| <b>Chapter 4 Populations in slums are happier than rural populations: The case of<br/>Mumbai .....</b>              | <b>141</b> |
| 4.1. Introduction .....                                                                                             | 141        |
| 4.2. Literature review.....                                                                                         | 146        |
| 4.3. Data.....                                                                                                      | 149        |
| 4.3.1 Sample .....                                                                                                  | 149        |
| 4.3.2 Definitions of the variables.....                                                                             | 150        |
| 4.4. Empirical approach.....                                                                                        | 156        |
| 4.5. Results.....                                                                                                   | 158        |
| 4.6. Discussion .....                                                                                               | 162        |
| 4.7. Conclusions .....                                                                                              | 165        |
| References.....                                                                                                     | 167        |
| Appendix 4.....                                                                                                     | 171        |
| <b>Chapter 5 Using satellite images to assess the economic impact of the World Cup at<br/>the city level.....</b>   | <b>173</b> |

|                                                                                                         |            |
|---------------------------------------------------------------------------------------------------------|------------|
| 5.1. Introduction .....                                                                                 | 173        |
| 5.2. The World Cup and the economy of South African host cities .....                                   | 178        |
| 5.3. Data and summary statistics .....                                                                  | 181        |
| 5.4. Methods.....                                                                                       | 185        |
| 5.4.1 The spatially varying relationship between total income and night light data .....                | 185        |
| 5.4.2 Estimation of differential income growth in host cities with a panel data ...                     | 188        |
| 5.5. Results and discussion .....                                                                       | 194        |
| 5.5.1 The relationship between economic development and night light.....                                | 194        |
| 5.5.2 Differential income growth in host cities .....                                                   | 199        |
| 5.5.2.1 Yearly difference in income growth between host cities and other cities .....                   | 199        |
| 5.5.2.2 Persistent income growth between host cities and other cities.....                              | 203        |
| 5.5.2.3 Heterogeneous income growth among host cities between host cities and other cities .....        | 206        |
| 5.6. Conclusions .....                                                                                  | 209        |
| References.....                                                                                         | 211        |
| Appendix 5 .....                                                                                        | 216        |
| Appendix 5 - Section A: Supplementary tables and figures.....                                           | 216        |
| Appendix 5 - Section B: SCM results .....                                                               | 227        |
| <b>Chapter 6 Differences in water policy efficacy across South African water management areas .....</b> | <b>232</b> |
| 6.1. Introduction .....                                                                                 | 232        |
| 6.2. South African water policies and local economic development .....                                  | 236        |
| 6.3. Methodology .....                                                                                  | 240        |
| 6.3.1 Geographically weighted regression .....                                                          | 240        |
| 6.3.2 The spatial Durbin model .....                                                                    | 242        |

|                                                                                                                                   |            |
|-----------------------------------------------------------------------------------------------------------------------------------|------------|
| 6.4. Data.....                                                                                                                    | 246        |
| 6.5. Results and discussion .....                                                                                                 | 250        |
| 6.5.1 The geographical distribution of the relationship between economic<br>development and water quality .....                   | 250        |
| 6.5.2 The impact of economic development and its statistical significance .....                                                   | 258        |
| 6.5.3 Factors underlying the regional differences in the impact of economic<br>development on water pollution across regions..... | 264        |
| 6.6. Conclusions and policy implications .....                                                                                    | 270        |
| References.....                                                                                                                   | 274        |
| Appendix 6.....                                                                                                                   | 280        |
| Appendix 6 - Section A: Supplementary tables and figures.....                                                                     | 280        |
| Appendix 6 - Section B: Pre-estimation tests.....                                                                                 | 284        |
| Appendix 6 - Section C: GWR results performance.....                                                                              | 287        |
| Appendix 6 - Section D: Additional results .....                                                                                  | 292        |
| <b>Chapter 7 Conclusion .....</b>                                                                                                 | <b>298</b> |
| References.....                                                                                                                   | 302        |

## LIST OF FIGURES

|                                                                                                     |     |
|-----------------------------------------------------------------------------------------------------|-----|
| Figure 2-1: Global distribution of NO <sub>2</sub> vertical column density .....                    | 32  |
| Figure 2-2 Schematic diagram of the instrumental variable strategy .....                            | 34  |
| Figure 2-3 Monthly change in NO <sub>2</sub> concentration and subjective well-being.....           | 35  |
| Figure 2-4 Locations of the respondents (1) .....                                                   | 60  |
| Figure 2-5 Locations of the respondents (2) .....                                                   | 61  |
| Figure 2-6 Locations of the respondents (3) .....                                                   | 62  |
| Figure 2-7 Locations of the respondents (4) .....                                                   | 63  |
| Figure 2-8 Locations of the respondents (5) .....                                                   | 64  |
| Figure 2-9 Locations of the respondents (6) .....                                                   | 65  |
| Figure 2-10 Histogram of subjective well-being.....                                                 | 66  |
| Figure 2-11 Well-being and pollution.....                                                           | 67  |
| Figure 2-12 Wellbeing and pollution by country .....                                                | 68  |
| Figure 3-1 Levels of capital per capita.....                                                        | 102 |
| Figure 3-2 Nature of capital .....                                                                  | 103 |
| Figure 3-3 Variations of capital per capita.....                                                    | 110 |
| Figure 3-4 Explanations of the difference in IW growth between key and secondary<br>provinces ..... | 116 |
| Figure 3-5 Histogram of variation of IWI.....                                                       | 132 |
| Figure 3-6 Histogram of the natural, human, and produced capital per capita in 2012                 | 133 |
| Figure 3-7 Total IWI.....                                                                           | 134 |

|                                                                                                                   |     |
|-------------------------------------------------------------------------------------------------------------------|-----|
| Figure 3-8 Variations of all forms of capital.....                                                                | 135 |
| Figure 3-9 Variation in the corporants of human capital.....                                                      | 136 |
| Figure 3-10 Variation in the corporants of produced capitals.....                                                 | 137 |
| Figure 3-11 Variation in the corporants of renewable natural capital.....                                         | 138 |
| Figure 3-12 Variation in the corporants of non-renewable natural capital.....                                     | 139 |
| Figure 5-1 Night light and South African total income in 2007.....                                                | 182 |
| Figure 5-2 Trends in light growth between host cities and other cities.....                                       | 193 |
| Figure 5-3 GWR estimates of the relationship between total income and night light in 2007.....                    | 195 |
| Figure 5-4 The difference in predicted income growth between host cities and other cities from 2002 to 2013 ..... | 202 |
| Figure 5-5 National South African GDP from 1990 to 2018.....                                                      | 216 |
| Figure 5-6 Provincial South African GDP from 1990 to 2018 .....                                                   | 217 |
| Figure 5-7 The distribution of the land cover in 2010.....                                                        | 218 |
| Figure 5-8 The distribution of the population in 2010.....                                                        | 219 |
| Figure 5-9 Trends of the growth of night light between host cities and other cities ....                          | 220 |
| Figure 5-10 Prediction accuracy of the GWR model .....                                                            | 221 |
| Figure 5-11 The difference in income growth between host cities and other cities from 1992 to 2013.....           | 225 |
| Figure 5-12 SCM results from (Pfeifer et al., 2018).....                                                          | 229 |
| Figure 5-13 Replication of SCM results from (Pfeifer et al., 2018) with our data .....                            | 230 |
| Figure 5-14: Account of population, urban and cropland areas in SCM analysis.....                                 | 231 |

|                                                                                 |     |
|---------------------------------------------------------------------------------|-----|
| Figure 6-1 The impact of the level of development on water quality indices..... | 256 |
| Figure 6-2 The impact of economic growth on water quality indices .....         | 257 |
| Figure 6-3 Sample repartition.....                                              | 281 |
| Figure 6-4Night light map in 2010 .....                                         | 282 |
| Figure 6-5 Land use repartition in South Africa.....                            | 283 |
| Figure 6-6Clusters among dependent variables .....                              | 285 |
| Figure 6-7 Cluster in the logarithm of night-light digital number.....          | 286 |
| Figure 6-8 Residual distribution of the GWR .....                               | 288 |
| Figure 6-9 Average economic development impact on water quality .....           | 289 |
| Figure 6-10 Economic growth impact on water quality .....                       | 290 |



## LIST OF TABLES

|                                                                                                                   |     |
|-------------------------------------------------------------------------------------------------------------------|-----|
| Table 2-1 Estimated impact of monthly changes in air pollution on well-being .....                                | 41  |
| Table 2-2 Heterogeneous impact of air pollution with regard to national-level characteristics .....               | 46  |
| Table 2-3 Countries summary statistics .....                                                                      | 56  |
| Table 2-4 Summary statistics for categorical variables .....                                                      | 58  |
| Table 2-5 Summary statistics .....                                                                                | 59  |
| Table 2-6 Full estimation results of the estimated impact of monthly changes in air pollution on well-being ..... | 69  |
| Table 2-7 Full estimation results of the heterogeneous impact of air pollution.....                               | 73  |
| Table 2-8 Impact of pollution on affective well-being .....                                                       | 78  |
| Table 3-1 Growth in IWI and its components from 2012 to 2018 .....                                                | 112 |
| Table 3-2 Difference in capital formation within countries between key and secondary provinces .....              | 113 |
| Table 3-3 Correspondence between land cover and ecosystem values .....                                            | 126 |
| Table 3-4 Oil reserves per country in 2018.....                                                                   | 128 |
| Table 3-5 Gas reserves per country in 2018.....                                                                   | 129 |
| Table 3-6 Coal reserves per country in 2018 .....                                                                 | 130 |
| Table 3-7 Commodities' price.....                                                                                 | 131 |
| Table 3-8 Difference in income growth between key provinces and secondary provinces .....                         | 140 |
| Table 4-1 Sample characteristics .....                                                                            | 155 |

|                                                                                                                                             |     |
|---------------------------------------------------------------------------------------------------------------------------------------------|-----|
| Table 4-2 Differences in well-being across areas in Mumbai.....                                                                             | 160 |
| Table 5-1 Summary statistics .....                                                                                                          | 184 |
| Table 5-2 The OLS estimates of the relationship between total income and night light in 2007.....                                           | 197 |
| Table 5-3 The comparative ability of GWR and OLS estimates to predict provincial income .....                                               | 198 |
| Table 5-4 The differential income growth in host cities .....                                                                               | 205 |
| Table 5-5 The differential income growth across host cities .....                                                                           | 208 |
| Table 5-6 Estimation of the difference of economic activities between host and non-host cities for every taken individually .....           | 222 |
| Table 5-7 Stadiums and venues .....                                                                                                         | 226 |
| Table 6-1 The average impact of economic development on water quality .....                                                                 | 262 |
| Table 6-2 The average impact of economic development on water quality after controlling for the impact of land use .....                    | 263 |
| Table 6-3 The average impact of economic development on water quality in eastern regions and western regions without control variables..... | 268 |
| Table 6-4 The average impact of economic development on water quality in former black-dominated lands with the control variables .....      | 269 |
| Table 6-5 Summary statistics .....                                                                                                          | 280 |
| Table 6-6 GWR performance with AICc estimation of bandwidth of the weight matrix .....                                                      | 287 |
| Table 6-7 GWR performance with cross-validation estimation of bandwidth of the weight matrix.....                                           | 291 |
| Table 6-8 SDM regression with log night-time light as the only human-related dependent variable .....                                       | 292 |
| Table 6-9 The estimation results with the spatial Durbin model.....                                                                         | 293 |

|                                                                                                                                                                       |     |
|-----------------------------------------------------------------------------------------------------------------------------------------------------------------------|-----|
| Table 6-10 SDM with a dummy variable for the East region without the control variables (results associated to Table 6-3) .....                                        | 294 |
| Table 6-11 SDM with a dummy variable for the eastern region and control variables (Results associated to Table 6-11) .....                                            | 295 |
| Table 6-12 The impacts of economic development on water quality in the eastern regions with the control variables .....                                               | 296 |
| Table 6-13 The average impact of economic development on water quality in former black-dominated areas with control variables (results associated to Table 6-4) ..... | 297 |

## **Chapter 1 Introduction**

### **1.1. Spatial economic inequalities**

#### *1.1.1 Definitions*

Economic inequality can be defined as the difference in welfare and income across individuals. In economics, these inequalities are caused by the diversity in individuals' income expectations and their engagements in various types of activity to earn their income. In the absence of coercion, individuals who provide the largest contributions to society, via the provision of goods and services, earn more than others, everything else being equal.

Inequality in this context of free trade is legitimate because while trading goods and services everyone pays the price, s/he judges fair during his/her trading interactions. Thus, the wealthiest individuals are those with the highest efficiency and innovative ideas owing to their ability to produce the largest quantities and most valuable goods to society respectively. In the context of free trade, the concentration of income and wealth in the hands of few people, the sources of economic inequality, shows that populations value goods and services produced by certain individuals more than that of others. These others, that is the least wealthy, experience a reduction of their relative income and welfare proportionally to their low added value to society. This archetype of economic inequality is the fuel that constraints the least efficient workers to innovate, to improve their efficiencies or to change activities to provide higher contributions to society in order to receive a higher income.

This same economic process determines the concept of regional economic inequality that is the difference in wealth between the wealthiest and the least wealthy areas. Regional economic inequality is the result of the agglomeration of high means of production, and highly productive individuals in certain areas. Populations in these areas are relatively wealthier than others because they provide goods and services highly valued by society. This is the reason underlying the economic development of some cities over other cities and over rural areas or even the development of some countries over other neighboring countries. A major consequence of the agglomeration of wealthy individuals in certain region is the access to valuable amenities, public goods, and services in these regions but not in others because only wealthy individuals are able to pay for these services.

Theoretically, regional economic inequality is not an issue. If an individual want to enjoy higher incomes and better amenities of wealthy regions, this individual should be free to migrate in a wealthy region. In classic economics, this setting is supposed to provide the largest freedom to each individual as one chooses what s/he wants for her/his life and the collection of choices by the mass results in an aggregate allocation of resources that reflect the will of societies.

However, most researchers, not only economists, suggest that we should reduce regional economic inequality or at least monitor them. Two main reasons determine such a stance. First, political, economic, and societal sciences suggest that regional economic inequalities are often linked to the coercion of some individuals by others. The extensive works in political economics of (Acemoglu & Robinson, 2013) or that of (LEWIS, 1972)

for instance, show that accumulation of wealth correlates with accumulation of political power across individuals and places.

Although the laws that govern our societies show that each citizen can elect the political representatives of her/his choice through “one” vote, these laws fail to regulate the money political representatives can put in their election campaigns, and or can give to citizens to get their votes. It results that political representatives of countries and regions are generally wealthy individuals whose laws may not align with that of individuals with lower socio-economic status and not living in wealthy areas. An example of this instance is exhibited in South Africa where highly developed cities coexist with impoverished slums. In addition to the presence natural economic processes that split the wealthy and the least wealthy populations between cities and slums, (Acemoglu & Robinson, 2013) show that wealthy South African populations have ensured for years that the least wealthy populations remain impoverished in slums by the presence of oppressive laws and policy. Thus, economic inequalities are often linked to political imbalance. In such kind of setting, regional economic inequality may lead to corruptions (Jong-sung & Khagram, 2005), civil wars (Østby, Nordås, & Rød, 2009) and lower national economic development (Alesina, Michalopoulos, & Papaioannou, 2016).

Second, economic forces that should push the least wealthy individuals to freely migrate toward wealthy areas, as dictates in the classic economic models, apply in the long-term but not in the short-term. Cost of living, the slow shift in demands and offers in the job or the housing markets in wealthy regions or even cultural differences between least wealthy and wealthy regions are elements that prevents least wealthy individuals to freely move in a wealthy region in the short-term. Therefore, most economist think that

there is a need to assist economic forces so that individuals unsatisfied with their income and wealth can migrate more easily.

Owing to this failure of economic system, several political entities at the national and international levels aim at reducing spatial economic inequalities as they consider it undesirable. Indeed, most countries in the world have ministries of urban development and that of rural development, where the goals of these entities is to manage economic inequality. Internationally, the SDG 1 (End poverty in all its forms everywhere), SDG 10 (Reduce inequality within and among countries) and SDG 11 (Make cities and human settlements inclusive, safe, resilient, and sustainable) for instance put some focus on the reduction of spatial economic inequality within and across nations worldwide.

#### *1.1.2 Lack of data to tackle spatial economic inequality*

Despite efforts made to address spatial economic inequalities, two major challenges limit the creation and the efficacy of policy for this issue. They are the collection of comparable and reliable data across countries and the availability of data within countries. Indeed, national efforts to provide statistical data widely diver in terms of quality of the output. Thus, the creation of policies and their assessments to tackle spatial inequality will be affected by measurement error because statistical data differ across countries. As exposed by (Henderson, Storeygard, & Weil, 2012)'s work on the measurement of national income, the quality of data widely varies from one country to another. To point attention at this issue, institutions such as the World Bank, the IMF, and Penn World Tables (PWT) rank the reliability of data by countries. In the Penn World

Tables, countries are given grades corresponding to subjective data quality, with a grade of A indicating a margin of error of 10%, B indicating 20%, C indicating 30%, and D indicating 40%. The grading is based in part on the ability to construct good purchasing power parity measures, but also reflects a country's capacity to produce reliable national income accounts and domestic price indices. These grading systems reveal that the quality of data in developing countries is substantially worse than that of developed countries.

In addition to the reliability of data at the national level, a second issue is that in most countries statistical data are not available on any consistent basis at the sub-national level. Most effort to monitor and share statistical data at the subnational level only exist in few developed countries. In other countries, the infrastructure and the capital of statistical agencies do not allow a reliable estimation of subnational data. Furthermore, most efforts to provide subnational level data focus only on income and population related dataset. Despite the growing issue of ecosystem loss, pollution, or climate change, subnational information on these issues is not consistent in developing as well as in developed countries. The reason being the lack of resources of statistical agencies to continuously monitor data on these issues. It is often prohibitively expensive to monitor information on all issues in all areas of countries. For instance, (Zou, 2021) shows for instance that only developed regions within countries have a reliable and frequent measurement of air pollution levels in USA. Thus, the state of subnational data collection across countries and within countries heavily impact the ability to develop policy tackling spatial economic inequality. This two issues led several economists to use remotely sensed information to address to further explore economic analyses.



## **1.2. The use of satellite data in economics**

As described by Dave Donaldson and Adam Storeygard in their seminal review on the use of satellite data in economics, photographic images of the Earth from above, recorded initially by cameras attached to balloons, kites, or pigeons have been used in several scientific disciplines since the 1900s. In the past decades however, a tremendous progress took place in remote sensing and allied fields such as computer science, engineering, and geography, remote sensing data. This progress made the use of data from satellite monitoring common in economics.

### *1.2.1 Description of satellite data*

Satellite images are pictures of the earth surface taken by satellites orbiting above it. Several satellites and techniques are used record pictures at different resolutions, different locations and during different timeframe. Thus, most of these images are recorded by highly funded aero-spaces institutions such as the NASA in the USA, or JAXA in Japan. Extensive processing is typically required to make raw satellite data comparable and interpretable. Thus, there is an active remote sensing literature developing and refining these methods and applying them to new and existing datasets. For instance, the area covered in each overpass of a satellite or other sensor over the earth is known as its “swath.” Within a given swath of satellite image of the earth, a satellite views the locations it is passing directly over (at nadir) from a right angle, but it is simultaneously viewing other locations at different angles. Hence, there is a need of perform an “orthorectification” to transforms the resulting images to account for these differences in angles, as well as an analogous distortion due to elevation differences. Nevertheless, economists almost never use raw satellite data since most interest on the

transformation of raw satellite data to more informative come from areo-space engineers. Consequently, our interest in satellite data to perform economic analysis is restricted to processed dataset.

Economists have used data derived from many different satellites and sensors. Table 1 lists a few of the most prominent. A comprehensive database of essentially all declassified and unclassified remote sensing satellites ever launched except for the very smallest, complete with information on onboard sensors and application areas, can be found in the World Meteorological Organization's OSCAR database: <http://www.wmo-sat.info/oscar/>. A series of eight satellites in the Landsat program have been collecting data since 1972. The MODIS sensor has flown on two satellites since 1999. It provides coarser resolution data than Landsat but has been useful in producing high-frequency imagery, and data on important phenomena including particulate pollution and fires. Night lights data from the Defense Meteorological Satellite Program–Operational Linescan System (DMSP-OLS), and Visible Infrared Imaging Radiometer Suite (VIIRS) sensors are even coarser, spatially and spectrally, but their simplicity has made them easy for economists to use. The 2000 Shuttle Radar Topography Mission (SRTM) provides the most comprehensive elevation data to date. Finally, a rapidly increasing trove of high-resolution images is being captured by private sector firms, including Spot Image and Digital Globe and, more recently, Planet and Terra Bella.

### *1.2.2 Novelty of satellite data in economics research and issues on spatial economic inequalities*

In their seminal work Dave Donaldson and Adam Storeygard group the main advantages of remote sensing data to economists into three categories: 1) access to information difficult to obtain by other means; 2) unusually high spatial resolution; and 3) wide geographic coverage.

The first advantage of satellite data they describe is simply that remote sensing technologies can collect panel data at low marginal cost, repeatedly, and at large scale on proxies for a wide range of hard-to-measure characteristics. This data relates to plethora of information such as night lights, precipitation, wind speed, flooding, topography, forest cover, crop choice, agricultural productivity, urban development, building type, roads, pollution, beach quality, and fish abundance. It is usually prohibitively expensive to measure accurately this information without remote sensing, and there are settings in which the official government counterparts of some remotely sensed statistics (such as pollution or forestry) may be subject to manipulation.

The second advantage of remote sensing data sources in Dave Donaldson and Adam Storeygard is that they are typically available at a substantially higher degree of spatial resolution than are traditional data. Much of the publicly available satellite imagery used by economists provides readings for each of the hundreds of billions of 30-meter-by-30-meter grid cells of land surface on Earth. Many economic decisions (particularly land use decisions such as zoning, building types, or crop choice) are made at approximately this same level of spatial resolution. Building on the fact that some innovative identification strategies used by economists exploit stark policy changes that occur at geographic boundaries; these high-spatial-resolution data enable more accurate and reliable study design.

The third key advantage of remotely sensed data lies in their wide geographic coverage. Only rarely do social scientists enjoy the opportunities, afforded by satellites, to study data that have been collected in a consistent manner—without regard for local events like political strife or natural disasters—across borders and with uniform spatial sampling on every inhabited continent. Equally important, many research satellites (or integrated series of satellites) offer substantial temporal coverage, capturing data from the same location at weekly or even daily frequency for several decades and counting.

These three advantages of satellite data are conducive to progress in the field of spatial economic inequalities. For instance, as there is no periodic economic index within African countries, (Hodler & Raschky, 2014) show that emissions of light at night in president's birth province increase by 13% more than in other provinces of countries starting when the president take office. By assuming that growth in night light satellite data can be a proxy of economic growth, (Hodler & Raschky, 2014) deduce that this difference in light growth is evidence of significant inequality of economic growth between regions where the president is born and other regions in Africa. This assumption that night light data can be a proxy of economic development was leveraged in several studies tackling spatial economic inequality in issues relative to ethnic inequalities, differences of economic development between urban and rural areas or differences of economic development between cities on the coast and those inner land for instance (Michalopoulos & Papaioannou, 2018).

Similarly, leaps have been made in the domain spatial inequalities relative environmental quality and pollutions. For instance, with satellite data measuring the amount of air pollutants in the stratosphere, (Zou, 2021) showed that current monitoring of air pollution within countries are likely biased and that heavily polluted areas benefited

from this situation. This bias came from the fact that economic agents responsible for pollution were aware of monitoring patterns and used this knowledge to decrease their level of pollution according the monitoring varying strictness in time.

Despite these advances provided by satellite data images, many efforts remain to be made for spatial policies. According to (Todes & Turok, 2018), the three main spatial policies are the spatial rebalancing, space-neutral and the place-based policies. While the spatial policies aim at narrowing the prosperity gap between regions and reduce unemployment in poorer areas, the space-neutral policies seek to generate the maximum aggregate economic growth via agglomeration economies and economic integration across regions. Finally, the place-based policies assist each region to its development potential in proportion to its local economy. None of these policies prevail on the other and the choice between either of them depends on the priorities of national and local governments as well as local economies. However, the creation and the assessment of these policies require data that are often not available through governmental agencies. Owing to the ability of satellite data to provide access to information difficult to obtain by other means of their high unusually high spatial resolution, their wide geographic coverage, they can provide a pedestal for the creation and the assessment of spatial policies. As few studies focus on the investigation of this claim, this thesis seeks to provide more evidence in that direction.

### **1.3. Contributions to literature**

The contribution of this thesis to the economic literature is double. It consists in:

- assessing spatial policies whose performance was unknown because of data unavailability

- analysing contexts where spatial policies are needed but inexistent because of data unavailability in the making of background analysis.

#### **1.4. The outline of the thesis**

To emphasize the pertinence of satellite images for spatial policies, I carry out three studies divided in three chapters.

**Chapter 2** is an exploration of the absence of consensus on the factors influencing the impact of air pollution on well-being across countries that are linked to lack of international air pollution data. With monthly nitrogen dioxide changes monitored by satellite used as an equivalent index of pollution worldwide, I estimate how individual and national income influence the impact of air pollution on the well-being of respondents to an international survey. To identify the burden of pollution on well-being, I focus on the impact of monthly changes in pollution induced by monthly changes in planetary boundary layers' height. Within countries, I find that air pollution reduces the well-being of low-income individuals and people living in highly polluted areas to a greater extent. Thus, it appears that populations in developing countries and countries with high income inequality suffer the highest burden of air pollution on their well-being than those in developed countries and countries with low-income inequality. However, populations in highly polluted countries and highly productive countries are the least affected by pollution. These results show that using an economic index to assess distress associated with pollution is inappropriate for addressing its heterogeneous burden. Furthermore, habituation to pollution by populations in highly polluted countries may threaten international collaboration for pollution reduction.

**Chapter 3** is an investigation into the subnational sustainability in Africa. Rapid population and economic growth in African countries raise concern over the sustainability of their development since their wealth greatly originate from natural endowments. To answer this matter, several pieces of research assessed the sustainability of the development of African countries at the national. However, national level research conceals subnational sustainability inequities embodied by evidence of inequalities of political power and disparities in investment rates and resource exploitation between areas around capital cities and the other areas of countries. Here, I estimate the sustainability of wealth across subnational provinces between 2012 and 2018, with the inclusive wealth accounting framework valuing the natural, manufactured, and human capitals. The estimates, based on subnational spatial data, reveal important differences in composition and amount of wealth within countries. I show that inclusive wealth per capita of African provinces grew by 3% on average and is embodied by investment in human and manufactured capitals offsetting natural capital depletion. However, there are unequal performances of subnational sustainability as, owing to their capital compositions, wealth in provinces with national capital cities grew by 8% faster than that of others. This may act as an important driver of population growth in these provinces by enhancing migration. Overall, the results reveal the potential for sustainability improvement within nations by replicating subnational successful sustainable development path. Furthermore, these estimates can assist decision-makers to plan need-based interventions within countries in creating sustainable development whose benefit extend nationally.

**Chapter 4** is examination into the difference in well-being between slum and rural areas in Greater Mumbai. The concept of urban-rural gradient of happiness suggests that

levels of happiness in urban areas are lower than those in rural areas. The urban-rural gradient of happiness applies in developed but not in developing countries. Previous studies suggest that the low supply of public services and economic development in developing countries' rural areas influence this absence of gradient of happiness. Thus, this study investigates whether the gradient of happiness exists in developing countries when rural areas are compared to urban areas' informal settlements because these settlements are known to receive few public services and few benefits from urban economic growth. I use a spatial error model to compare the happiness and life satisfaction of samples from Greater Mumbai's slums and rural areas, controlling for socio-economic profiles of the respondents to a survey and the average population and light emissions in their locations. Our results show that populations in slums have, on average, a subjective happiness, and a life satisfaction higher than that of rural populations. This result suggests that lifestyle in urban areas in developing countries is linked to higher happiness independently of the difference in levels of development between urban and rural areas.

**Chapter 5** is an assessment of the economic impact of the World Cup 2010 in South Africa (the most unequal country worldwide). South African government attempted to use this event to increase income and life conditions in host cities. However, the effect of the event on economic performances in South African cities remain unknown since the economic benefits of mega-events are noticeable at the city-level and the South African government does not monitor periodic city-level economic data. Thus, I generate an economic proxy to quantify the income growth in the economies of the South African host cities linked to the event. Based on the economic literature, I generate a city-level income variation index by (1) performing a geographically-weighted regression between



night-time lighting data and total city-level income to account for spatial inequality and by (2) using this relationship to predict total income from 1992 to 2013 with night light data. With this index, my comparison of income growth in host and non-host cities with panel data analysis shows similar income growth between the two types of cities. However, the cities that invest the most in building or renovating their stadiums experienced 9% lower income growth than other host cities between 2010 and 2013. These findings suggest a limited ability of the World Cup to stimulate economic growth and the presence of inefficient investment.

**Chapter 6** describes the investigation the legacy of the inequitable water policy under apartheid as it may continue to impact water services in economically less developed and rural areas in South Africa. I perform an analysis using satellite night-time light data as a proxy of economic development and surface water quality at the smallest area of water management to scrutinize any income or racial inequality regarding water pollution. I find a spatial discrepancy in the relationship between water quality and economic development; economic development improves water quality in western provinces and in former white-dominated areas, while it generally degrades water quality in other regions of the country. These results suggest the inability of the government to equitably provide the same standards of water policies nationwide and the presence of inequitable policies as legacies of apartheid at the smallest level of water policy implementation.

## References

- Acemoglu, Daron, and James Robinson. 2013. *Why Nations Fail: The Origins of Power, Prosperity, and Poverty*. 1st Editio. New York: Crown Business.
- Alesina, Alberto, Stelios Michalopoulos, and Elias Papaioannou. 2016. "Ethnic Inequality." *Journal of Political Economy* 124 (2): 428–88. <https://doi.org/10.1086/685300>.
- Henderson, J Vernon, Adam Storeygard, and David N Weil. 2012. "Measuring Economic Growth from Outer Space." *American Economic Review* 102 (2): 994–1028. <https://doi.org/10.1257/aer.102.2.994>.
- Hodler, Roland, and Paul A Raschky. 2014. "Regional Favoritism." *Quarterly Journal of Economics* 129 (2): 995–1033. <https://doi.org/10.1093/qje/qju004>.Advance.
- Jong-sung, You, and Sanjeev Khagram. 2005. "A Comparative Study of Inequality and Corruption." *American Sociological Review* 70 (1): 136–57. <https://doi.org/10.1177/000312240507000107>.
- LEWIS, W. ARTHUR. 1972. "Reflections on Unlimited Labor." In *International Economics and Development*, 75–96. Elsevier. <https://doi.org/10.1016/B978-0-12-216450-7.50017-7>.
- Michalopoulos, Stelios, and Elias Papaioannou. 2018. "Spatial Patterns of Development: A Meso Approach." *Annual Review of Economics* 10 (1): 383–410. <https://doi.org/10.1146/annurev-economics-080217-053355>.
- Østby, Gudrun, Ragnhild Nordås, and Jan Ketil Rød. 2009. "Regional Inequalities and Civil Conflict in Sub-Saharan Africa." *International Studies Quarterly* 53 (2): 301–24. <https://doi.org/10.1111/j.1468-2478.2009.00535.x>.
- Todes, Alison, and Ivan Turok. 2018. "Spatial Inequalities and Policies in South Africa: Place-Based or People-Centred?" *Progress in Planning* 123 (July): 1–31. <https://doi.org/10.1016/j.progress.2017.03.001>.
- Zou, Eric Yongchen. 2021. "Unwatched Pollution: The Effect of Intermittent Monitoring on Air Quality." *American Economic Review* 111 (7): 2101–26. <https://doi.org/10.1257/aer.20181346>.

## **Chapter 2 The multinational and heterogeneous burden of air pollution on well-being**

### **2.1. Introduction**

Air pollution is a global issue responsible for more than 4 million premature deaths every year (X. Li, Jin, & Kan, 2019). However, its burden is widely heterogeneous across countries and across communities within countries. Nearly 89% of air pollution-related deaths occur in low-income and middle-income countries (Tessum et al., 2019). Within nations, death is more prevalent among minorities and poorer populations (Ash & Boyce, 2018).

Air pollution also harms life choices through its impact on people's well-being (see (Levinson, 2020b; Y. Li, Guan, Tao, Wang, & He, 2018) for a review of the literature). However, researchers do not agree on the heterogeneity of the impact of air pollution on well-being. On the one hand, empirical evidence, such as Levinson (2012) in the US and Ambrey et al. (2014) in Australia, show that the harm of pollution on an individual's self-assessed well-being (SWB) increases with income. Across countries, although populations in developing countries are unable to bear the cost of effective environmental policies, some studies show that they are less concerned by environmental issues than those in developed countries (Xudong Chen, Huang, & Lin, 2019; Franzen & Vogl, 2013) and prioritize jobs, income, and consumption over clean air. This perspective suggests that air pollution has a small weight in the construction of their well-being and air pollution-driven morbidity rates may be a consequence of the tendency of low-income populations to live in areas with high levels of pollution (Banzhaf, Ma, & Timmins, 2019).

However, other evidence suggests that the burden of air pollution may be the highest for low-income individuals' well-being because of their inability to take necessary pecuniary preventive measures against it (B. O. Deschênes, Greenstone, & Shapiro, 2017; J. G. Zivin & Neidell, 2013). Internationally, populations in developing countries may be exposed to undesirable pollution because their countries are more likely to be economically unequal (Acemoglu & Robinson, 2013), and therefore, linked to few economic and political leverage opportunities for the majority of the population to act against pollution (Boyce, 2007). In this context, countries that do not implement policies against air pollution owing to the lack of demand from its population to reduce it may help perpetuate unwanted air pollution-driven morbidity rates (X. Zhang, Zhang, & Chen, 2017).

Nevertheless, previous studies are limited by the availability of international data that link individual-level information and high-resolution pollution exposure (Li et al., 2018). Previous works investigating the impact of exposure to air pollution on well-being focus either on (1) single countries with data at the high resolution or (2) several countries with data averaged at the country level (see survey on the topics by Li et al., 2018 and Levinson, 2020). In the first instance, the estimation results do not provide a sufficient international perspective and cannot be used to explore the heterogeneities of the impact of air pollution with regard to national characteristics. In the latter instance, the estimations may be biased, because they fail to account for important differences in pollution exposure within countries (see Figure 2-1) and they disregard endogenous factors affecting pollution and well-being simultaneously (Levinson, 2020b).

Furthermore, both types of research lack focus on developing countries and may provide inferences that do not apply globally.

Thus, I aim to provide multinational evidence of the heterogeneous harm of air pollution on well-being influenced by individual and national income-related factors. I carry out a survey on well-being in 30 developed and developing countries responsible for 76% of CO<sub>2</sub> emissions (World Bank estimations) and 66% of nitrogen dioxide (NO<sub>2</sub>) emissions (estimations based on my dataset) worldwide in 2016 (see Table 2-3). I pair the address of the respondents of this survey with a proxy of outdoor air pollution, the gridded satellite measurement of vertical column density (VCD) of NO<sub>2</sub> particles suspended between the earth's surface and the lowest troposphere limit (in molecules/cm<sup>2</sup>). Then, I model well-being as a function of the index monthly changes in NO<sub>2</sub> particles at the grid level, income, and other individual and environmental characteristics. In this model, I isolate the impact of pollution on well-being from other factors that also affect well-being, by focusing on changes in NO<sub>2</sub> uniquely resulting from changes in the planetary boundary layers' height.

This study uses the NO<sub>2</sub> VCD monitored by the GOME-2 satellite as an index of pollution, because of its global coverage at an 80×40 km resolution; the availability of the data during the sample collection; the link of NO<sub>2</sub> to lower well-being (Levinson, 2020b) and health (O. Deschênes, Greenstone, & Shapiro, 2017; Schwartz, Bind, & Koutrakis, 2017); and the ability of NO<sub>2</sub> to be a proxy of other air pollutants (Levy, Mihele, Lu, Narayan, & Brook, 2014).

Hence, this study fills the literature gap related to the narrow international perspectives of studies investigating the impact of air pollution on SWB raised in a survey

by Li et al. (2018). As described by Levinson (2020) and Li et al. (2018), previous international studies of the impact of air pollution on well-being use data averaged at the national level and, thereby, disregard important subnational disparities displayed in Figure 2-1. One notable exception is the work of Ferreira et al. (2013), who investigate the impact of air pollution on well-being with subnational data but restricted to 23 European countries. Our study is unique in estimating the multinational impact of air pollution with (1) data of the exposure at the grid level (80×40 km) and (2) data stemming both from developed and developing countries spread across five continents (Table 2-3). Therefore, I adopt a global view to explore the global issue of air pollution.

Further, this is the first study to show the pertinence of global satellite monitored air pollution as a tool for investigating the impact of pollution on well-being. Most previous studies use indexes of pollution from monitoring stations. However, this data may lack a subnational consistent index in highly polluted and poor areas because of subnational discrimination in ambient air monitoring (Grainger & Schreiber, 2019) while developing countries may lack a pollution monitoring index (Barwick, Li, Lin, & Zou, 2019). With this regard, the NO<sub>2</sub> VCD, used as proxy of outdoor pollution, provides an equivalent assessment of the impact of air pollution on well-being in developed and developing countries as well as in subnational areas where monitoring is not properly conducted.

Finally, this work helps justify the assumption that clean air should be perceived as a normal good for which populations in developing countries are unable to pay despite higher exposure to pollution and disutility (Greenstone & Jack, 2015). The empirical basis of this assumption has not been established, and an analysis with samples from the UK

and China suggest that habituation to pollution may disprove this assumption (Y. Li et al., 2019). I show that, although air pollution is most detrimental for populations in developing and unequal countries, populations in countries with high levels of pollution develop some habituation to the effect of pollution.

## **2.2. Heterogeneous impact of air pollution on well-being with respect to income**

### *2.2.1 The impact of air pollution on well-being*

Exposure to air pollution can obstruct people's emotions (Y. Li et al., 2019; X. Zhang et al., 2017), work productivity (He, Liu, & Salvo, 2019a; J. G. Zivin & Neidell, 2012), length of time spent outdoors (G. J. Zivin & Neidell, 2009), or school performance (Balakrishnan & Tsaneva, 2021). Because the effects of air pollution are varied, numerous researchers have estimated its impact on the concept of well-being. Well-being is an aggregate of the perception of factors, such as mood, income, productivity, social relationships, and the environment, which lead people to think and experience their lives in positive versus negative ways (Diener, Oishi, & Tay, 2018). It is an appropriate index to assess the effects of air pollution on people's perceptions of their lives as it conveniently encompasses their likes and dislikes.

To evaluate the causal impact of pollution on well-being, some quasi-experimental studies have shown that changes in either an exogenous variable that affects pollution but not well-being or the short-time variations of high-resolution observations of pollution level may provide robust estimates. This effect is often quantified as willingness to pay (WTP), a monetary valuation of the effective demand for clean air

generally used to inform environmental policy. WTP can be expressed as the part of income an individual is willing to give up for cleaner air while maintaining the same level of well-being (Levinson, 2020b). For instance, using a quasi-experimental study design in Germany, Luechinger (2009) estimate the harm of SO<sub>2</sub> as WTP for a reduction of SO<sub>2</sub> by 1  $\mu\text{g}/\text{m}^3$  equal to USD 232. While Du et al. (2018) estimate the harm of pollution as the WTP for a reduction of 1  $\mu\text{g}/\text{m}^3$  of SO<sub>2</sub> of USD 1899 in Beijing and USD 2234 in Shanghai. Building on the leaps yielded by these quasi-experimental studies in understanding the impact of pollution in single countries, there is reason to believe that air pollution affects well-being internationally. Thus, I propose the following hypothesis.

**H0.** Air pollution has a statistically significant impact on well-being.

Nevertheless, uncertainty remains about the reasons for the large differences in its impact across individuals and countries. In different ways, several studies suggest that the prevalent socio-economic cause of divergent impact of air pollution on well-being across countries and individuals is their income.

### *2.2.2 The heterogeneous impact of pollution at the individual level*

Despite general agreement in the literature on the higher harm of air pollution on low-income individuals' health, no clear conclusion exist for well-being. One perspective suggests that the harm of pollution on well-being is largest among high-income people at the individual level. According to this line of thought, high-income individuals have less economic problems to worry about and therefore, are more affected by other concerns, such as air pollution. For instance, Franzen and Vogl (2013) show that within countries environmental concerns increase with income for a sample of 31 countries. This argument



is similar to that of Liu et al. (2016), who show that awareness of and dissatisfaction with pollution is highest for high-income individuals in a sample of three Chinese cities. Specific to well-being, Levinson (2012) shows that as individual income increases, the impact of daily air pollution on well-being increases in the US. In Queensland, Australia, Ambrey et al. (2014) find that, when exposed to excess air pollution, the likelihood of individuals reporting the highest level of life satisfaction is lowest for highest income populations.

However, other studies oppose these findings by stressing that low-income individuals are more likely to be exposed to high levels of air pollution and less likely to act effectively against its harm on their well-being (B. O. Deschênes et al., 2017; J. G. Zivin & Neidell, 2013). They argue that low-income populations consistently ingest more pollution than high-income populations owing to their inability to take pecuniary measures against pollution (e.g., residency selection or the purchase of air filters). Subsequently, air pollution affects well-being through channels beyond awareness, such as work productivity (Hausman & Stolper, 2020). Accordingly, under the assumption that disutility in purchasing a pecuniary preventive measure by high-income individuals is lower than the disutility resulting from the impact of air pollution on oneself, air pollution has a worse impact on low-wage earners' well-being.

Considering these conflicting perspectives specific mostly to individual countries, I aim to provide a multinational view on this topic. I postulate the following hypothesis.

**H1.** The impact of air pollution on well-being is significantly lower for low-income populations.

Related to the concept of residential sorting, I further explore the effect of income on the impact of pollution on well-being. Residential sorting pushes relatively low-income populations to live in highly polluted areas owing to their inability to pay for amenities (Ash & Boyce, 2018; Banzhaf et al., 2019). Building on Levinson (2020) and Li et al. (2019), who argue that the well-being of populations in highly polluted areas may be less affected by pollution because they become habituated to pollution, I propose the following hypothesis.

**H2.** The impact of air pollution on well-being is significantly lower for individuals living in highly polluted areas.

### *2.2.3 The heterogeneous impact of pollution of on well-being internationally*

Discrepancies of findings of the impact of pollution on well-being at the micro-economic level contributes to two conflicting views at the macro-economic level (Greenstone & Jack, 2015). The first perspective suggests that there is a relatively low valuation of clean air by populations in developing countries owing to the two-phase relationship between income per capita and pollutant emissions, or the environmental Kuznets curve (EKC) (Beja, 2011; Xudong Chen et al., 2019). The EKC stipulates that the beginning of economic growth is related to increases in pollution emissions because, despite their inability to afford environmental regulations, low-income countries are keener for income than for clean air. After income per capita increases past a certain level, communities value clean air highly, and regulatory policies become more effective. Hence, pursuit of the benefits of economic development at the expense of air pollution by the

low-income populations in these countries contributes to the increasing productivity, linked to pollution (Greenstone & Jack, 2015). In addition to the relatively limited monitoring of pollution in developed countries (Barwick et al., 2019), this perspective suggests lower awareness of air pollution (Franzen & Vogl, 2013), which could result in lower effects of pollution on overall well-being in less developed countries.

Another perspective suggests that higher pollution levels in developing countries may be the result of income inequality rather than relative indifference to pollution (Banzhaf et al., 2019; Boyce, 2007). An unequal distribution of income may lead to a low WTP for pollution reduction. In an unequal country, if the majority of the population is too poor to express its dissatisfactions with dirty air, political elites and wealthy private parties interested in goods production and consumption may intensify their activities even if polluting (Banzhaf et al., 2019; Boyce, 2007; Oswald et al., 2020). This case of unequal economic capacity is recurrent in several developing countries, where the roots of national poverty are explained by the concentration of income and political power by elite populations (Acemoglu & Robinson, 2013). Therefore, this inequity in bargaining power may explain large levels of national emissions in developing countries and countries with high income inequality, and large impact of air pollution on well-being nation-wide.

Thus, I scrutinize the various national income-related characteristics that, according to these opposing perspectives, may influence the impact of air pollution on well-being. I investigate the effect of countries' level of development (income per capita) on the impact of air pollution on well-being. Assuming that level of development raise the awareness to pollution, I propose the hypothesis.

**H3.** The impact of air pollution on well-being is significantly lowest for populations in developing countries.

Then, I investigate the influence of national income inequality on the impact of pollution on well-being. Assuming that levels of development raise the awareness to pollution, and that countries with large inequality are less developed, I propose the following hypothesis.

**H4.** The impact of air pollution on well-being is significantly lowest for populations in countries with high income inequality.

Next, I explore the divergence of the impact of pollution on well-being linked to nationally emitted NO<sub>2</sub>. Since NO<sub>2</sub> may reflect disinterest in reducing national pollution, I propose the following hypothesis.

**H5.** The impact of air pollution on well-being is significantly lowest for populations in highly polluted countries.

Finally, I investigate the influence of total national income on the impact of pollution. As national income is the main yardstick of economic development, which can, however, lead populations to overlook pollution. I propose the following hypothesis.

**H6.** The impact of air pollution on well-being is significantly lowest for populations in highly productive countries.

## **2.3. Data**

### *2.3.1 Sample*

This study used the results of an international survey on well-being of 89,266 respondents from 30 countries collected by Nikkei Research Inc. (Chapman, Fujii, & Managi, 2019). These countries are Australia, Brazil, Canada, Chile, China, Colombia, Czech, France, Germany, Greece, Hungary, India, Indonesia, Italy, Japan, Malaysia, Mexico, Myanmar, Netherlands, Philippines, Poland, Romania, Russia, South Africa, Spain, Sweden, Thailand, Turkey, US, Vietnam. They were chosen based on the survey company's ability to collect data there and budgetary constraints. The sample sizes by countries are described in Table 2-3 in the Appendix. After setting goals to collect data from each country's population age and gender categories, the respondents were randomly selected nationwide to fill these demographics via the internet. As a result, my sample is geographically dispersed within countries (see Figures 2-4 – 2-9). We ensured the respondents understood the questionnaire and transnational equivalence of the answers by translating and checking the questions and answers several times. In each country, the survey was conducted for a month between June 2015 and March 2017 (see Table 2-3). This survey investigated respondents' self-assessment of their well-being, and socio-economic status, the statistics for which are displayed in Table 2-4 in the Appendix.

### *2.3.2 Measure of well-being*

I use so-called cognitive well-being (or life satisfaction) as the index of SWB and ensure the robustness of my inferences with an index of affective well-being, also known

as hedonic well-being (see Appendix Section B). Our analysis focus on a single-item measure of well-being, since aggregating multiple items to form an SWB index is statistically inconsistent (Cronbach's  $\alpha < 0.5$ ), in my view, owing to the international nature of the survey (see Appendix Section C).

Cognitive well-being is assessed based on the answer to the question: "All things considered, how satisfied are you with your life these days?" This SWB, recorded on a 10-point Likert-type scale (0=lowest level of satisfaction; 10=highest) is a judgmental assessment of well-being by respondents widely established by researchers (Diener et al., 2018). This index can capture the proportion of dissatisfaction in lives that are linked to exposure to pollution. In particular, it includes dissatisfaction resulting from awareness of pollution, largely attributed to high income, and dissatisfaction with pollution-related lower productivity or health, which may be prevalent among low-income populations.

This SWB index is a common measure of total utility in economics and psychology (Kahneman, Wakker, & Sarin, 1997). Although the locution "these days" in the question refers to a vague timeframe, I assume that the SWB index differs significantly according to monthly changes in pollution based on the model of total utility developed by Kahneman et al. (1997). In this model, the total utility that can be assimilated as SWB is the sum of (1) instant utility, understood as utility at the exact moment (day or hour) of the respondent's answer, and (2) remembered utility, understood as a monotonic sum of retrospective evaluations of past episodes that decrease when an individual experiences a negative event (and increase with a positive event). To construct remembered utility, events that occurred in the distant past were assumed to have lower importance, as individuals may forget (S. Chen, Oliva, & Zhang, 2018). Therefore, I

limited SWB to be the combination of utility during the survey and the utility of a period effectively remembered by an individual.

By apprehending the SWB index through the lens of the model of total utility, we could deduce that it is relevant for evaluating the impact of a 1-month change in air pollution on well-being, for two reasons. First, the importance of instant utility over the average remembered utility in an individual's assessment of total utility matters significantly only when he or she is affected by an extraordinary experience during (or shortly before) the evaluation of total utility (Kahneman et al., 1997). However, since extraordinary events (e.g., an extraordinary increase in pollution) are unlikely to occur on average, we could assume that respondents' SWB is a significant representation of their average remembered utility (Figure 2-10 in the Appendix shows that SWB approximately follows a normal distribution). Second, since previous studies show that significant impacts of air pollution on aspects involved in the SWB construction, such as awareness of pollution (Oltra & Sala, 2018), productivity (He, Liu, & Salvo, 2019b), and mental illness (S. Chen et al., 2018), are noticeable within a 1-month period, we could assume that the negative impact of pollution within 1-month is effectively remembered.

### *2.3.3 Factors determining well-being*

As for the index of outdoor air pollution, I collected satellite-derived monthly NO<sub>2</sub> concentration data from the Global Ozone Monitoring Experiment (GOME-2) satellite (ESA, n.d.). This satellite measures the global daily tropospheric NO<sub>2</sub> VCD through an optical process and models a rigorous monthly average of NO<sub>2</sub> at an 80×40 km resolution,

usable in a broad range of scientific studies (Hickman et al., 2021; Richter, Begoin, Hilboll, & Burrows, 2011).

I then used age, gender, number of children, education level, occupation, self-assessed health, and marital status of respondents as control variables in my regressions. With regard to respondents' household income collected as income ranges, I transformed the categorical ranges of income into real values by taking the midpoint of the corresponding range and multiplying this midpoint by the purchasing power parity relative to USD, to account for transnational differences in currency and life cost (Jebb, Tay, Diener, & Oishi, 2018; Levinson, 2012).

Furthermore, I included wind speed, precipitation data, and minimum and maximum monthly temperatures as control variables, because they are linked to both pollution and well-being (S. Ferreira et al., 2013; Levinson, 2012). These global maps were collected from the TerraClimate dataset (Abatzoglou, Dobrowski, Parks, & Hegewisch, 2018) at a 4-km resolution for the corresponding months of the survey in every country. The monthly PBLH, used as an instrumental variable, was collected from (NASA, n.d.-a) at a resolution of  $0.5^\circ \times 0.625^\circ$ . I matched all these raster data with respondents' locations at the grid level. Finally, proximity to sea and roads is known to impact levels of air pollution and SWB (Levinson, 2020b). Thus, distances to the closest coast and the closest roads of each respondent were accounted for, using the NASA database (NASA, n.d.-b) and the Natural Earth database (NACIS (North American Cartographic Information Society), n.d.), respectively.

When investigating the heterogeneous air pollution burden depending on national factors, I relied on the 2016 World Bank dataset indexes. The summary statistics of



variables are displayed in Appendix Tables 2-4 and 2-5, at the individual-level and in Table 2-3 at the country level.

## 2.4. Empirical approach

### 2.4.1 *The impact of air pollution on well-being*

This analysis uses the following function:

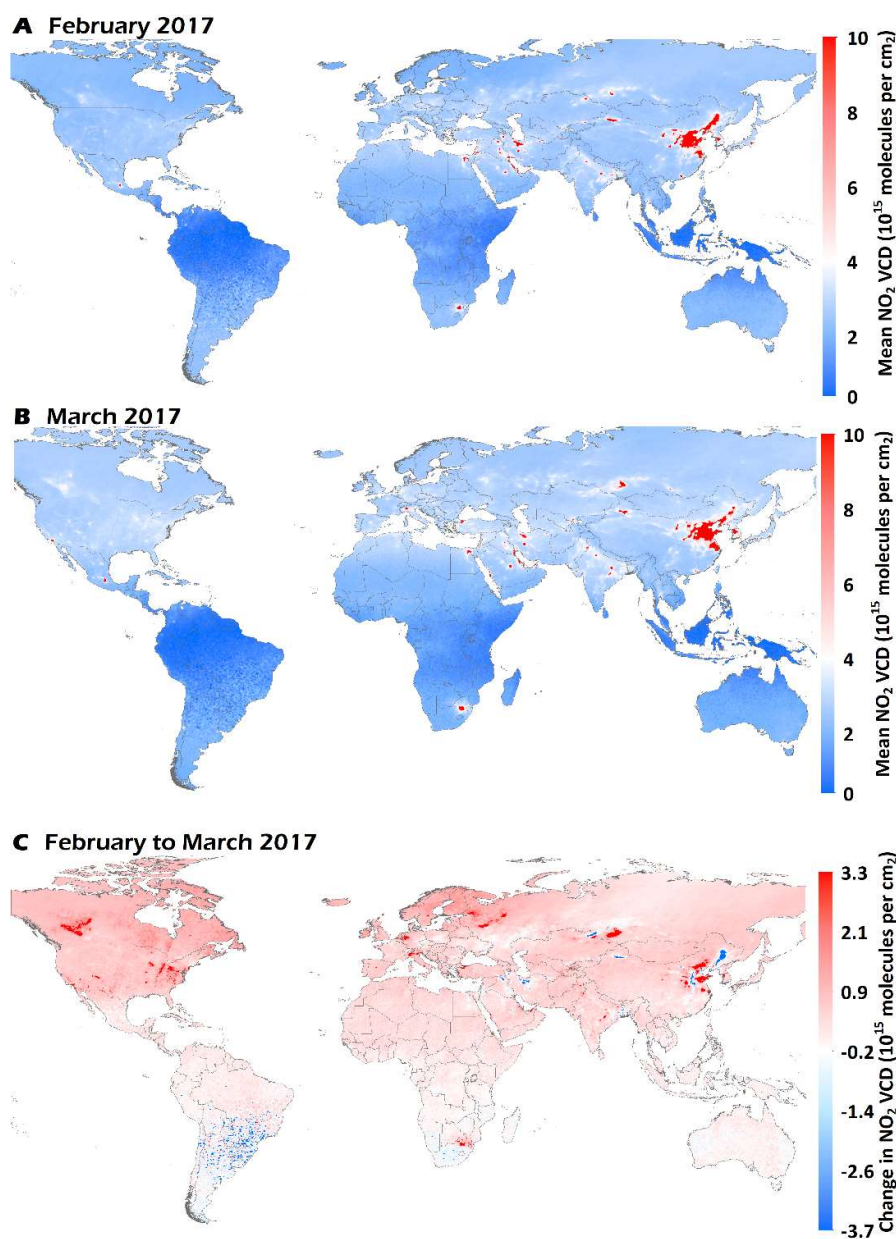
$$SWB_i = \theta_0 \Delta P_i + \sum_{k=1}^n \gamma_k X_{ki} + \sum_{c=1}^l \delta_c Z_{ci} + \eta_c + \epsilon_i \quad (2.1)$$

to explore the impact of air pollution on well-being. I use ordinary least squares (OLS) to calibrate this model (calibration with ordered Probit model is performed as a robustness check). In this function,  $SWB_i$  refers to the SWB of respondent  $i$ . The term  $\Delta P_i$  is a 1-month change in air pollution to which respondent  $i$  is exposed. The term  $X_{ki}$  is a set of individual variables highly linked to SWB, which comprises the logarithm of income, gender, number of children, age (in quadratic form), marital status, occupation status, education level, subjective health assessment, distance to the closest road, and inverse distance to the coast. Variables represented by  $Z_{ci}$  are precipitation, minimum and maximum temperatures, and wind speed. These climatic variables are included in quadratic form because I expect extreme climatic events to harm well-being, while mild climatic features may augment well-being. The term  $\eta_c$  represents country dummy variables, set to capture national differences in perception or language of the respondents, and the national economic characteristics impacting respondents' answers (Oishi, 2010).

Finally,  $\theta_0$ ,  $\gamma_k$ , and  $\delta_c$  are the coefficients of the variables that I seek to estimate, while  $\epsilon_i$  is the stochastic error term of respondent  $i$ .

Our function uses a 1-month change in pollution instead of the average monthly pollution for three reasons. The relationship between the average pollution during the month of the survey and well-being may be biased because (1) people become habituated to average pollution over time; (2) those who prioritize clean air the most are likely to live in areas with cleaner air; and (3) several other factors affecting well-being simultaneously affect levels of pollution (see Figure 2-2 and Levinson, 2020). Using the 1-month change in monthly pollution (instead of the average) allows us to estimate the impact of pollution that is comparable within nations and across developed cities, because the exposure to these changes in pollution are evenly spread across areas (see Figure 2-1). Consequently, the relationship between the monthly change in pollution and SWB is negative, while the potentially biased relationship between the monthly average level of pollution and SWB is positive (see Figure 2-3, as well as Figures 2-11 and 2-12 in the Appendix).

Figure 2-1: Global distribution of NO<sub>2</sub> vertical column density



**Note:** These maps illustrate the identification strategy of the study in March 2017. The mean NO<sub>2</sub> VCD in (A) February 2017 and (B) March 2017 show that the most polluted regions in the world are developed cities. To avoid amalgamating the effect of NO<sub>2</sub> with lifestyles in these cities, I use the changes in NO<sub>2</sub> between these months (C), since these changes are roughly evenly distributed worldwide.

Nevertheless, the sole use of change in air pollution may not be robust to an omitted variable that causes pollution and simultaneously affects well-being. Increases in monthly traffic jams, which may impact stress and air pollution simultaneously, or increases in home heating, which may impact happiness and pollution, are examples of this bias.

Therefore, I employ an instrumental variable strategy to provide estimations uncorrelated to sources of pollution with the changes in monthly planetary boundary layer height ( $\Delta PBLH_i$ ) as an instrumental variable (see identification strategy in Figure 2-2). The PBLH is the distance between the Earth's surface and the troposphere, which is the lowest layer of the Earth's atmosphere and the area where most air pollution is impacted by vertical transport. The PBLH is inversely related to air pollution concentrations, suggesting that in two places with the same levels of air pollution, the place with the lowest PBLH would have the largest air pollution concentration. Since PBLH is mainly influenced by climatic processes (Schwartz et al., 2017; Xiang et al., 2019), its variations are not directly linked to well-being but can affect it through air pollution. To model the relationship between pollution and PBHL, I perform a linear regression with PBLH in quadratic form. I justify this choice by the necessity to use linear regressions in the two-stage least squares (2SLS) model (Angrist & Pischke, 2009), and by the argument of Schwartz et al. (2017) suggesting a non-linear relationship between PBHL and air pollution.

Hence, I can reformulate the null hypothesis stipulating that air pollution affects well-being with the parameter of interest in (2.1) as follows:

$$\mathbf{H0} : \theta_0 \neq 0 \quad (2.2)$$

Figure 2-2 Schematic diagram of the instrumental variable strategy

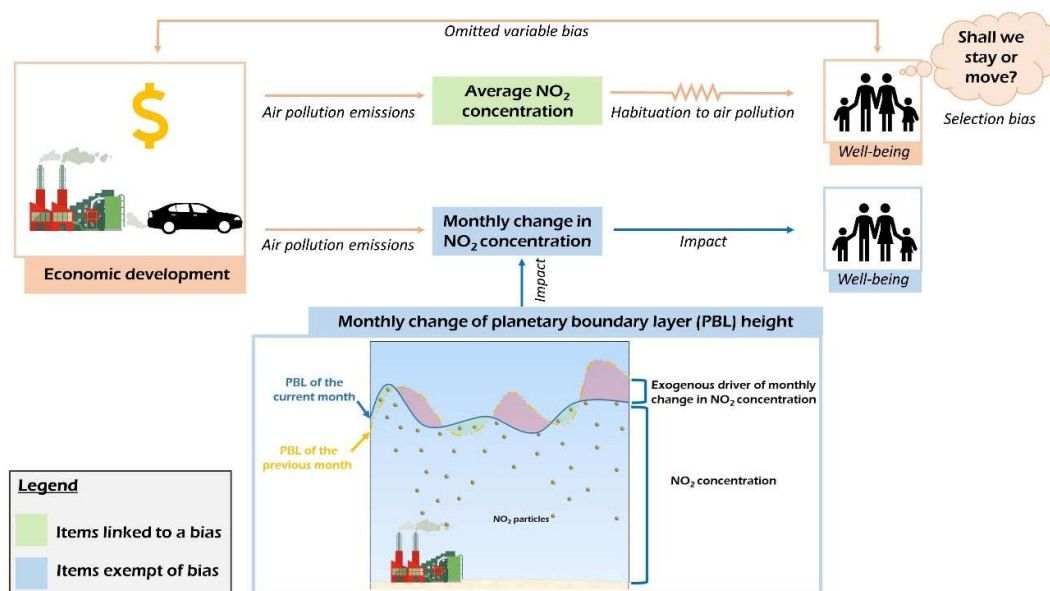
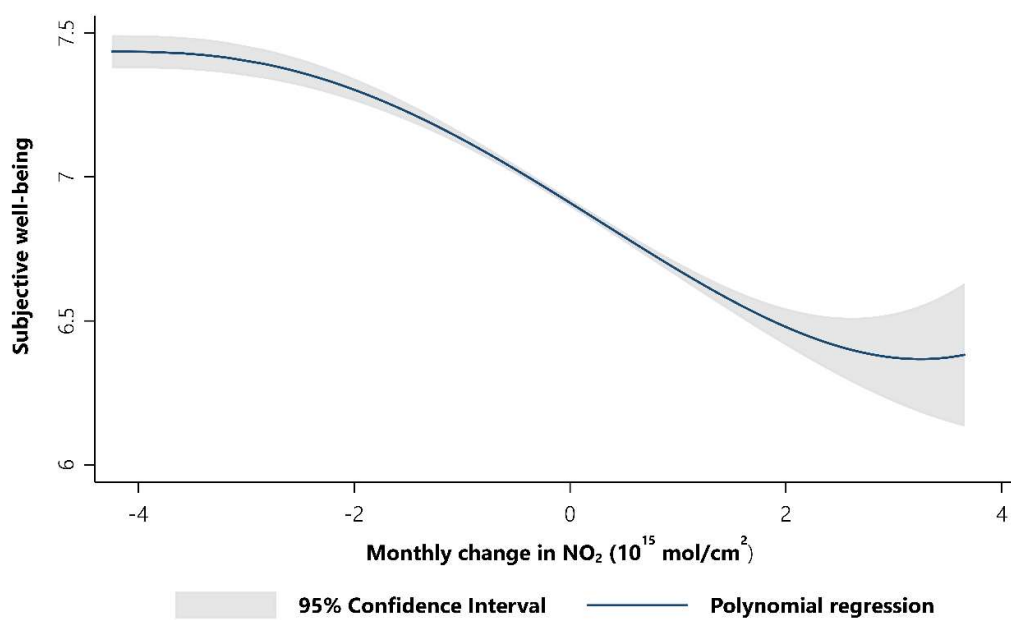


Figure 2-3 Monthly change in NO<sub>2</sub> concentration and subjective well-being



#### 2.4.2 Hypotheses testing of the heterogeneous impact of air pollution with regard to income

To investigate the heterogeneous impact of air pollution on well-being, I regress the following model:

$$SWB_i = \theta_1 \Delta P_i + \theta_f \Delta P_i \times F_i + \sum_{k=1}^n \gamma_k X_{ki} + \sum_{c=1}^l \delta_c Z_{ci} + \eta_c + \epsilon_i \quad (2.3)$$

The difference between models (2.1) and (2.3) is the term  $F_i$  which is introduced as an interaction term with pollution. This term is the individual or national factor that causes a heterogeneous impact of air pollution on well-being.

To determine the heterogeneous impact of air pollution, I refer to the sign of the coefficient  $\theta_f$  in different model specifications. Different model specifications are created by changing the nature of the variable  $F_i$  to investigate the hypotheses raised in Section 2.2. At the individual level,  $F_i$  refers to personal income relative to the average income in my sample to test H1. In the second iteration,  $F_i$  represents the average monthly pollution at the grid level during the month of the survey to investigate H2. I can formulate the null hypotheses of H1 and H2 with regard to the coefficients  $\theta_f$  in (2.3) as follows:

$$\mathbf{H1} : \theta_{income} < 0 \quad (2.4)$$

$$\mathbf{H2} : \theta_{average\ pollution} > 0 \quad (2.5)$$

In the investigation of national-level hypotheses, the value of  $F_i$  is informed by national monthly averages of NO<sub>2</sub> emissions, GDP, and the Gini index to represent the national amount of air pollution faced by populations, levels of productivity, and the level

of inequality, respectively. However, I do not directly interact these macro-economic indexes with the impact of air pollution, because they are already accounted for by the country's indicator variables ( $\eta_c$ ) in equation (2.3). Following Dell et al. (2012), I create dummy variables from groups of countries generated according to these macroeconomic indexes. These groups are labeled productive, unequal, and polluted countries, and are defined as having above-median GDP, Gini index, and total NO<sub>2</sub>, respectively, among the countries in my dataset. In the case of H4, to categorize the level of development, I simply use the developing country taxonomy of the World Bank to create a developing country dummy (see categories of countries in Table A1). With these dummy variables, I investigate H3–H6. I formulate national-level null hypotheses with regard to the coefficients  $\theta_f$  in (2.3) as follows:

$$\mathbf{H3} : \theta_{developing\ country} > 0 \quad (2.6)$$

$$\mathbf{H4} : \theta_{unequal\ country} > 0 \quad (2.7)$$

$$\mathbf{H5} : \theta_{polluted\ country} > 0 \quad (2.8)$$

$$\mathbf{H6} : \theta_{productive\ country} > 0 \quad (2.9)$$

## 2.5. Results and discussion

### 2.5.1 The impact of air pollution on well-being

I present the results of the first-stage 2SLS regression that models the relationship between PBLH changes and NO<sub>2</sub> changes in Panel B of Table 2-1. As suggested by previous studies (Schwartz et al., 2017; Xiang et al., 2019), there is a significant negative relationship between PBLH and NO<sub>2</sub> concentration ( $p < 0.01$ ). PBLH can be regarded as a



robust instrumental variable, as its F-score is a statistically significant number superior to 10 (Stock & Yogo, 2005). I further investigate the validity of PBLH as an instrument by examining its dissociation with respondents' stochastic error via Sargan tests (Panel B). Despite the test score questioning the use of PBLH in quadratic form at the 10% confidence level, I assume that the instruments are adequate, because Sargan tests are often stricter than required in real-world situations (Angrist & Jorn-Steffen, 2009).

Following the prediction of changes in air pollution exclusively resulting from changes in PBLH, I model SWB as a function of air pollution, climatic variables, individual characteristics, and country indicator variables, using equation (2.1) (see Section 4). The coefficients of the factors that affect the relationship between pollution and well-being in Panel A of Table 2-1 and those of the control variables are presented in Tables 2-6 and 2-7 in the Appendix.

OLS regressions show that monthly changes in NO<sub>2</sub> statistically harm SWB ( $p < 0.01$ , Column 1 of Table 2-1). However, the estimate of the impact of changes in NO<sub>2</sub> that are solely the results of PBLH variations, estimated via 2SLS, is three times larger than the estimate of OLS ( $p < 0.01$ , Column 2 of Table 2-1). The lower influence of emissions by OLS is likely to be linked to the correlation between air pollution and economic activities. To aid interpretation of my results, I refer to a  $10^5$  molecules/cm<sup>2</sup> change in NO<sub>2</sub> as a unit change in NO<sub>2</sub>. A unit change in NO<sub>2</sub> represents approximately a variation of pollution levels from such a city as Dallas (3.9 NO<sub>2</sub> units) to a city like New York (5 NO<sub>2</sub> units).

The 2SLS regression indicates that living in an area where the NO<sub>2</sub> density column increases by a unit within a month is related to a lower well-being of 0.256 points out of

10. Hence, it appears that well-being would drop by nearly 0.256 points if pollution were to change from the levels in Shanghai (6.6 NO<sub>2</sub> units) to that of Beijing (7.4 NO<sub>2</sub> units); from Osaka (4.3 NO<sub>2</sub> units) to Tokyo (5.1 NO<sub>2</sub> units); from Rio de Janeiro (2.2 NO<sub>2</sub> units) to São Paulo (3 NO<sub>2</sub> units); from Marseilles (4.2 NO<sub>2</sub> units) to Paris (5.1 NO<sub>2</sub> units); or from Rome (4.4 NO<sub>2</sub> units) to Milan (5.6 NO<sub>2</sub> units).

To inform policies, we can interpret this impact as WTP determined by the marginal rate of substitution between income and air pollution. The marginal rate of substitution of pollution for income is the percentage of income that an individual must be given to compensate for a one-unit increase in pollution, keeping his or her well-being unchanged. For this purpose, I use the formula of Welsch (2006):

$$WTP = \frac{\partial SWB / \partial NO_2}{\partial SWB / \partial \log(income)} \quad (2.10)$$

with  $\partial SWB / \partial NO_2$  representing the impact of NO<sub>2</sub> on SWB, and  $\partial SWB / \partial \log(income)$  representing the impact of a 1% increase of income on SWB. The regression from the 2SLS, in Column 2 of Table 2-1, shows that  $\partial SWB / \partial NO_2 = -0.224$  and  $\partial SWB / \partial \log(income) = 0.227$ . Using these estimates,  $WTP \approx -0.98$ : a one-unit increase in monthly NO<sub>2</sub> is equivalent, in SWB terms, to a 0.98% decrease in annual income. For an individual earning the average income in my sample (USD 3546), this represents a decrease of USD 35.

These results are robust to (1) the omission of respondents' characteristics (Column 2, Table 2-6); (2) the use of an ordered probit model instead of a linear model to account for the categorical nature of the SWB variable (Column 3, Table 2-1); and (3) the inclusion of average monthly NO<sub>2</sub> in the regression to account for eventual correlation

of variables (Column 4, Table 2-1). Since air pollution changes harms well-being, we fail to reject  $H_0$ . Thus, our model appears to be consistent enough to investigate the heterogeneous burden of air pollution with regard to income at the individual and country levels.

Table 2-1 Estimated impact of monthly changes in air pollution on well-being

| <b>Panel A</b>                                                                 | Dependent variable: subjective well-being (second-stage regression of 2SLS) |                                                          |                                                          |                                                          |
|--------------------------------------------------------------------------------|-----------------------------------------------------------------------------|----------------------------------------------------------|----------------------------------------------------------|----------------------------------------------------------|
|                                                                                | (1)<br>OLS regression                                                       | (2)<br>Baseline—2SLS<br>regression                       | (3)<br>Ordered probit<br>model 2SLS <sup>a</sup>         | (4)<br>Baseline—with average<br>monthly NO <sub>2</sub>  |
| <b>VARIABLES</b>                                                               |                                                                             |                                                          |                                                          |                                                          |
| Change in NO <sub>2</sub>                                                      | -0.0794***<br>(0.0130)                                                      | -0.224***<br>(0.0512)                                    | -0.131**<br>(0.0548)                                     | -0.225***<br>(0.0523)                                    |
| log(income)                                                                    | 0.235***<br>(0.0115)                                                        | 0.227***<br>(0.00897)                                    | 0.121***<br>(0.00727)                                    | 0.227***<br>(0.00910)                                    |
| Average NO <sub>2</sub>                                                        |                                                                             |                                                          |                                                          | -0.000751<br>(0.00631)                                   |
| Constant                                                                       | 2.777***<br>(0.272)                                                         | 2.719***<br>(0.198)                                      |                                                          | 2.722***<br>(0.202)                                      |
| Observations                                                                   | 89,266                                                                      | 89,266                                                   | 89,266                                                   | 89,266                                                   |
| R-squared                                                                      | 0.279                                                                       | 0.276                                                    | 0.074                                                    | 0.276                                                    |
| Weather variables                                                              | Yes                                                                         | Yes                                                      | Yes                                                      | Yes                                                      |
| Country dummy                                                                  | Yes                                                                         | Yes                                                      | Yes                                                      | Yes                                                      |
| Personal characteristics                                                       | Yes                                                                         | Yes                                                      | Yes                                                      | Yes                                                      |
| <b>Panel B</b>                                                                 |                                                                             |                                                          |                                                          |                                                          |
| First-stage regression of 2SLS. Dependent variable: changes in NO <sub>2</sub> |                                                                             |                                                          |                                                          |                                                          |
| PBLH changes                                                                   |                                                                             | -0.00121***<br>(0.00003)                                 | -0.00121***<br>(0.00003)                                 | -0.00121***<br>0.00003                                   |
| PBLH changes squared                                                           |                                                                             | -3.76(10 <sup>-6</sup> )***<br>(1.46(10 <sup>-7</sup> )) | -3.76(10 <sup>-6</sup> )***<br>(1.46(10 <sup>-7</sup> )) | -3.63(10 <sup>-6</sup> )***<br>(1.54(10 <sup>-7</sup> )) |
| Control variables                                                              |                                                                             | Yes                                                      | Yes                                                      | Yes                                                      |
| Instruments F statistics                                                       |                                                                             | 726.236***                                               | 726.236***                                               | 642.795***                                               |
| Wu–Hausman test                                                                |                                                                             | 8.9990***                                                | -                                                        | 12.4819***                                               |
| Sargan test                                                                    |                                                                             | 3.648 (p= 0.056)                                         | -                                                        | 3.661 (p= 0.056)                                         |

\*\*\* p<0.01, \*\* p<0.05, \* p<0.1. <sup>a</sup> Standard errors of the ordered probit model estimations do not account for the first stage

regression estimates density function. The standard errors of the coefficients are shown in parentheses. These standard errors correspond to the Huber–White standard errors, except for the OLS estimation, in which I cluster the standard errors using the second level of administrative division in each country provided by <https://gadm.org/>. This clustering adjustment is not performed for the 2SLS estimations, because the PBLH changes are assumed to be random. The Wu–Hausman endogeneity null hypothesis test assumes that pollution changes are exogenous, while that of the Sargan test for over-identification assumes that the model is not over-identified.

### *2.5.2 Heterogeneous impacts of changes in pollution with regard to income*

I interact the impact of air pollution with personal income subtracted from average income in our sample to determine whether the impact of pollution on well-being decreases as individuals earn higher income. The result shows a negative and significant coefficient of the changes in NO<sub>2</sub> (-0.339,  $p < 0.01$ ), while the interaction term is statistically superior to zero using a one-tailed test (2.618,  $p = 0.08$ ) (Column 1, Table 2-2). In this regression specification, the coefficient of the changes in NO<sub>2</sub> refers to the impact of a one-unit change in NO<sub>2</sub> on an individual whose income is equal to our sample average (USD 3,546), and the coefficient of the interaction term refers to the additional impact of a one-unit change in NO<sub>2</sub> when income rises by 1% from the average. Therefore, it appears that the higher the income, the lower the impact of air pollution on well-being. Our results differ from those of Ambrey et al. (2014) and Levinson (2012), showing that populations whose well-being is more responsive to dirty air have low income. This could be the result of this population having higher exposure to air pollution in their daily lives and the higher ability of high-income individuals to take costly measures against pollution (He et al., 2019b; J. G. Zivin & Neidell, 2013). Therefore, we reject H1.

Regarding hypothesis H2, I interchangeably use the expressions “1 unit change in NO<sub>2</sub>” and “change in pollution” as well as “average monthly pollution” and “average pollution.” The results show a positive and statistically insignificant effect of change in pollution on SWB (1.483), a statistically insignificant and negative impact of the average pollution (-0.0154), and a negative and statistically significant coefficient of the interaction of these two variables (-0.221,  $p = 0.04$ , Column 2, Table 2-2). The coefficient of the change in pollution displays the effect of changes in pollution when the average

pollution is null. The coefficient of average pollution represents the impact that an individual suffers from air pollution in an area where there is no change in pollution and the average pollution is of one unit of NO<sub>2</sub>. The interaction term displays the additional burden that this individual would experience from pollution changes if he or she resided in an area where the average pollution was one unit higher than that of the person's actual place of residence. Since this additional impact is negative, increases in pollution in relatively polluted areas reduce well-being more than increases in pollution in areas with low levels of pollution. This effect may be due to higher awareness of pollution among residents of polluted areas (S. Chen, Chen, Lei, & Tan-Soo, 2020), and to the inability to bear air pollution past some threshold (Y. Li et al., 2019). Thus, we reject H2, because individuals living in highly polluted areas may become habituated to air pollution. Moreover, since low-income individuals are the most affected by pollution (Column 1, Table 2-2) and live, on average, in areas with poor levels of air quality (Banzhaf et al., 2019), their harm from air pollution is doubled.

### *2.5.3 Heterogeneous impacts of changes in pollution with regard to national income characteristics*

At the national level, I interact the impact of pollution changes with dummy variables created from countries labeled as (1) developing, (2) unequal, (3) polluted, and (4) productive by reference to their national economic indexes to determine the heterogeneous impact of air pollution (see Section 4 and Table 2-3 for further details). The coefficient of the interaction variable reflects the impact of air pollution in the

corresponding labeled countries, while the coefficient of NO<sub>2</sub> changes refers to the impact of air pollution in countries excluded from these groups.

The results show that while monthly augmentations in air pollution have an insignificant impact on well-being in developed countries, they harm well-being in developing countries 0.988 out of 10 points more than they harm well-being in developed countries ( $p = 0.035$ ; Table 2-2, Column 3). I attribute some of these results to the increasing availability of information relative to pollution and the awareness of this issue in developing countries (Xudong Chen et al., 2019). Nevertheless, I believe that the higher availability of preventive measures against air pollution in developed countries may effectively prevent the exposure to pollution by the population in these countries and, therefore, their lower dissatisfaction with air pollution. Hence, we reject H3.

Second, air pollution reduces well-being in countries with large income inequality by 1.935 points more than it does in their counterpart countries ( $p = 0.04$ , Table 2-2, Column 4). These results can be explained when placed in conjunction with our finding that the low-income individuals suffer the most from pollution (Column 1, Table 2-1). Since national inequality indicates that a small proportion of the population has a relatively high income, we can deduce that the largest share of the population in countries with high income inequality suffer from pollution harm. Furthermore, national disparities within countries may illustrate that a large proportion of people are exposed to harmful levels of pollution because they lack economic and political power (Banzhaf et al., 2019; Boyce, 2007). Therefore, we reject H4.

Third, pollution significantly reduces well-being in both highly and slightly polluted countries, but the impact of air pollution in countries with high pollution is 1.487

points lower than in slightly polluted countries ( $p=0.04$ , Table 2-2, Column 5). Hence, although residents in polluted areas are the most affected by pollution changes within countries, this transnational investigation shows that citizens of highly polluted countries are least affected by changes in pollution (Levinson, 2020b; Y. Li et al., 2019). The fact that the influence of average emissions differs depending on whether the region comprising the pollution is defined as a country or a sub-national area suggests that certain national and cultural characteristics influence how people respond to air pollution. For instance, governments and individuals of highly polluted countries are likely to share information about pollution widely or engage in large-scale trade of pollution-related goods because of their higher national experience with pollution. Within countries however, residents of areas with large levels of pollution may experience disproportionately high levels of pollution relative to the national average and thus, may suffer significantly more than their compatriots. Consequently, we fail to reject H5.

Finally, pollution significantly reduces well-being in highly productive and less productive countries, but the harm of air pollution in productive countries is lower by 1.375 points ( $p=0.04$ , Table 2-2, Column 6). I believe that the lesser harm of pollution in productive countries should be related to their long-term habituation to large levels of pollution, as most productive countries are classified polluted countries (see Table 2-3). Hence, we fail to reject H6.



Table 2-2 Heterogeneous impact of air pollution with regard to national-level characteristics

| VARIABLES                                              | Dependent variable: Subjective well-being (Second-stage regression of 2SLS) |                                                     |                                                              |                                                      |                                                                 |                                                               |
|--------------------------------------------------------|-----------------------------------------------------------------------------|-----------------------------------------------------|--------------------------------------------------------------|------------------------------------------------------|-----------------------------------------------------------------|---------------------------------------------------------------|
|                                                        | (1)                                                                         | (2)                                                 | (3)                                                          | (4)                                                  | (5)                                                             | (6)                                                           |
|                                                        | Impact of pollution interacted with income                                  | Impact of pollution interacted with pollution level | Impact of pollution interacted with the level of development | Impact of pollution interacted with inequality level | Impact of pollution interacted with national pollution emission | Impact of pollution interacted with the level of productivity |
| Change in NO <sub>2</sub>                              | -0.339***<br>(0.120)                                                        | 1.483<br>(0.954)                                    | 0.752<br>(0.546)                                             | 1.735<br>(1.093)                                     | -1.604**<br>(0.780)                                             | -1.409**<br>(0.668)                                           |
| Variation in NO <sub>2</sub> * log(income)             | 2.617 <sup>†</sup><br>(1.912)                                               |                                                     |                                                              |                                                      |                                                                 |                                                               |
| Variation in NO <sub>2</sub> * Average NO <sub>2</sub> |                                                                             | -0.221 <sup>††</sup><br>(0.123)                     |                                                              |                                                      |                                                                 |                                                               |
| Change in NO <sub>2</sub> * developing country         |                                                                             |                                                     | -0.988 <sup>††</sup><br>(0.550)                              |                                                      |                                                                 |                                                               |
| Change in NO <sub>2</sub> * unequal country            |                                                                             |                                                     |                                                              | -1.935 <sup>††</sup><br>(1.078)                      |                                                                 |                                                               |
| Change in NO <sub>2</sub> * highly polluted country    |                                                                             |                                                     |                                                              |                                                      | 1.487 <sup>††</sup><br>(0.838)                                  |                                                               |
| Change in NO <sub>2</sub> * productive country         |                                                                             |                                                     |                                                              |                                                      |                                                                 | 1.375 <sup>††</sup><br>(0.773)                                |
| log(income)                                            | 0.386***<br>(0.117)                                                         | 0.219***<br>(0.0104)                                | 0.220***<br>(0.00992)                                        | 0.212***<br>(0.0124)                                 | 0.247***<br>(0.0143)                                            | 0.248***<br>(0.0146)                                          |
| Average NO <sub>2</sub>                                |                                                                             | -0.0154<br>(0.0106)                                 |                                                              |                                                      |                                                                 |                                                               |
| Observations                                           | 89,266                                                                      | 89,266                                              | 89,266                                                       | 89,266                                               | 89,266                                                          | 89,266                                                        |
| R-squared                                              | 0.262                                                                       | 0.261                                               | 0.268                                                        | 0.255                                                | 0.239                                                           | 0.248                                                         |
| Weather variables                                      | Yes                                                                         | Yes                                                 | Yes                                                          | Yes                                                  | Yes                                                             | Yes                                                           |
| Country dummy                                          | Yes                                                                         | Yes                                                 | Yes                                                          | Yes                                                  | Yes                                                             | Yes                                                           |
| Personal characteristics                               | Yes                                                                         | Yes                                                 | Yes                                                          | Yes                                                  | Yes                                                             | Yes                                                           |

The significance levels relative to the test of the coefficients being different from zero, using a two-tailed hypothesis, are displayed as

\*\*\* p<0.01, \*\* p<0.05, \* p<0.1; and the significance levels relative to the one-tailed hypothesis testing are displayed as <sup>††</sup>p<0.01,

<sup>††</sup>p<0.05, <sup>†</sup>p<0.1. Definitions of countries' taxonomies are discussed in Section 2.3 and their classes is displayed in Table 2-3

#### *2.5.4 Policy implications*

Overall, although I cannot fully claim a causal impact due to the cross-sectional nature of our survey, our findings offer perspectives on the impact of air pollution on well-being critical for research and policy purposes. First, we can deduce that life in developing countries, and in countries with high income inequality reduces access to preventive measures against air pollution, because only a small share of the population can afford them. Hence, these results oppose the findings of studies suggesting that populations in wealthy countries and wealthy individuals are more responsive to air pollution (Franzen & Vogl, 2013). Since the impact of air pollution on well-being is lowest for high-income individuals, these results suggest that stronger actions to deal with pollution by high-income individuals and countries are likely due to their greater ability to pay for clean air. Thus, governments that focus on increasing their economic growth through national productivity may significantly reduce the well-being of their least wealthy citizens, particularly when the prosperity of this economic growth is not evenly shared.

Consequently, our results suggest that the unique use of the economic valuation of clean air to create atmospheric policies may be ineffective to tackle the heterogeneous burden of air pollution. Within countries, low-income populations that live in highly polluted areas bear the highest impact of air pollution on both their well-being and their health (Ash & Boyce, 2018). To avoid associating their inability to pay with indifference to pollution, I suggest a continual measure of well-being as a policy index. Such an index would incorporate the views of the population with low economic and political leverage in policies tackling pollution (Stiglitz, Sen, & Jean-Paul, 2009). In addition, continual

monitoring of the relationship between a well-being index and local pollution changes could provide good incentive for individuals, governments, and companies to act against it.

Moreover, considering that work productivity is the highest for individuals with high well-being (He et al., 2019b), I posit that non-habituation to outdoor air pollution within countries suggests that residential sorting with respect to pollution creates a downward spiral of poverty. This process could be understood as a market failure, because it places additional burdens on the low-income populations although they pollute significantly less (e.g., Oswald et al., 2020). Thus, for the process of pollution-related residential sorting stemming from individual choices, it is necessary to provide further information on the levels of pollution (e.g. consistent monitoring) and its danger to population (the low-income population in particular) who disregard it. However, for the process of pollution-related residential sorting related to firms relocating in poor neighborhoods, it is necessary to materialize their impact on well-being as a cost.

Finally, our results show that populations living in highly polluted and productive countries may be reluctant to engage in international measures against air pollution because national environment participates in habituation to high levels of pollution on their well-being. This habituation to pollution may be a threat to international collaboration against pollution. Thus, to fight air pollution at the international level effectively, incentives that extend beyond well-being are necessary.

## **2.6. Conclusions**

This study investigated how factor related to individual and national income factors influence the harm from air pollution on well-being based on a multinational sample. On the one hand, the results show that air pollution has the highest impact on the well-being of low income individuals and those living in highly polluted areas, and thus, suggest the ineffectiveness of policies that uniquely focus on economic forces to reduce pollution burdens. On the other hand, while populations in developing and countries with high income inequality are the most affected by air pollution, those in highly polluted countries and productive countries suffer the least. Hence, the international perspective of our analysis reveals some habituation to air pollution, related to national environments, which may hinder global cooperation to reduce air pollution.

Nevertheless, our study has some limitations that require additional research. First, our sample may over-represent the wealthiest and youngest ranges of populations within countries owing to the use of an internet survey. Furthermore, our survey lacks important temporal representativeness within nations, as it was conducted for a single month for every country. Therefore, further research using a broader range of population repeatedly monitored over time is necessary to confirm our results.

Furthermore, since air pollution reduces productivity, and because low well-being is linked to lower productivity, further studies are paramount to value the cost of higher exposure to pollution and the economic tradeoff that lower-income populations face.

Finally, our estimate of the impact of pollution on well-being is not comparable with previous research because pollution is measured by satellite monitored NO<sub>2</sub> VCD.

Thus, there is a need to convert NO<sub>2</sub> VCD in  $\mu/m^3$  to increase the pertinence of satellite data in the current topic.

## References

- Abatzoglou, J.T., S.Z. Dobrowski, S.A. Parks, and K.C. Hegewisch. 2018. "Terraclimate, a High-Resolution Global Dataset of Monthly Climate and Climatic Water Balance from 1958-2015." Scientific Data. 2018.
- Acemoglu, Daron, and James Robinson. 2013. *Why Nations Fail: The Origins of Power, Prosperity, and Poverty*. 1st Edition. New York: Crown Business.
- Ambrey, Christopher L., Christopher M. Fleming, and Andrew Yiu-Chung Chan. 2014. "Estimating the Cost of Air Pollution in South East Queensland: An Application of the Life Satisfaction Non-Market Valuation Approach." *Ecological Economics* 97 (January): 172–81. <https://doi.org/10.1016/j.ecolecon.2013.11.007>.
- Angrist, Joshua D., and Pischke Jorn-Steffen. 2009. "Over-Identification and 2SLS Minimand." In *Mostly Harmless Econometrics: An Empiricist's Companion*, 105. Princeton university press.
- Ash, Michael, and James K. Boyce. 2018. "Racial Disparities in Pollution Exposure and Employment at US Industrial Facilities." *Proceedings of the National Academy of Sciences* 115 (42): 10636–41. <https://doi.org/10.1073/pnas.1721640115>.
- Balakrishnan, Uttara, and Magda Tsaneva. 2021. "Air Pollution and Academic Performance: Evidence from India." *World Development* 146 (October): 105553. <https://doi.org/10.1016/j.worlddev.2021.105553>.
- Banzhaf, Spencer, Lala Ma, and Christopher Timmins. 2019. "Environmental Justice: The Economics of Race, Place, and Pollution." *Journal of Economic Perspectives* 33 (1): 185–208. <https://doi.org/10.1257/jep.33.1.185>.
- Barwick, Panle Jia, Shanjun Li, Liguang Lin, and Eric Zou. 2019. "From Fog to Smog: The Value of Pollution Information." 26541. Cambridge, MA. <https://doi.org/10.3386/w26541>.
- Beja, Edsel L. 2011. "Subjective Well-Being Approach to Environmental Valuation: Evidence for Greenhouse Gas Emissions." *Social Indicators Research* 109 (2): 243–66. <https://doi.org/10.1007/s11205-011-9899-1>.
- Boyce, James K. 2007. "Is Inequality Bad for the Environment?" *Research in Social Problems and Public Policy* 15 (135): 267–88. [https://doi.org/10.1016/S0196-1152\(07\)15008-0](https://doi.org/10.1016/S0196-1152(07)15008-0).
- Chapman, Andrew, Hidemichi Fujii, and Shunsuke Managi. 2019. "Multinational Life Satisfaction, Perceived Inequality and Energy Affordability." *Nature Sustainability* 2 (6): 506–14. <https://doi.org/10.1038/s41893-019-0303-5>.
- Chen, Shuai, Yuyu Chen, Ziteng Lei, and Jie-Sheng Tan-Soo. 2020. "Impact of Air Pollution on Short-Term Movements: Evidence from Air Travels in China." *Journal of Economic Geography* 20 (4): 939–68. <https://doi.org/10.1093/jeg/lbaa005>.
- Chen, Shuai, Paulina Oliva, and Peng Zhang. 2018. "Air Pollution and Mental Health: Evidence from China." Cambridge, MA. <https://doi.org/10.3386/w24686>.

- Chen, Xudong, Bihong Huang, and Chin-Te Lin. 2019. "Environmental Awareness and Environmental Kuznets Curve." *Economic Modelling* 77 (March): 2–11.  
<https://doi.org/10.1016/j.econmod.2019.02.003>.
- Dell, By Melissa, Benjamin F Jones, and Benjamin A Olken. 2012. "Temperature Shocks and Economic Growth: Evidence from the Last Half Century." *American Economic Journal: Macroeconomics* 4 (3): 66–95.
- Deschênes, By Olivier, Michael Greenstone, and Joseph S Shapiro. 2017. "Defensive Investments and the Demand for Air Quality : Evidence from the NOx Budget Program." *American Economic Review* 107 (10): 2958–89.
- Deschênes, Olivier, Michael Greenstone, and Joseph S Shapiro. 2017. "Defensive Investments and the Demand for Air Quality: Evidence from the NOx Budget Program." *American Economic Review* 107 (10): 2958–89.  
<https://doi.org/10.1257/aer.20131002>.
- Diener, Ed, Daniel Kahneman, William Tov, and Raksha Arora. 2010. "Income's Association with Judgements of Life versus Life Feelings." In *International Differences in Well-Being*, edited by Ed Diener, John F. Helliwell, and Daniel Kahneman, 2nd ed., 3–33. New York: Oxford University Press.
- Diener, Ed, Shigehiro Oishi, and Louis Tay. 2018. "Advances in Subjective Well-Being Research." *Nature Human Behaviour* 2 (4): 253–60.  
<https://doi.org/10.1038/s41562-018-0307-6>.
- ESA. n.d. "GOME-2." Accessed January 7, 2021.  
<https://www.eumetsat.int/website/home/Satellites/CurrentSatellites/Metop/MetopD esign/GOME2/index.html>.
- Ferreira, Susana, Alpaslan Akay, Finbarr Brereton, Juncal Cuñado, Peter Martinsson, Mirko Moro, and Tine F Ningal. 2013. "Life Satisfaction and Air Quality in Europe." *Ecological Economics* 88: 1–10.  
<https://doi.org/10.1016/j.ecolecon.2012.12.027>.
- Franzen, Axel, and Dominikus Vogl. 2013. "Two Decades of Measuring Environmental Attitudes: A Comparative Analysis of 33 Countries." *Global Environmental Change* 23 (5): 1001–8. <https://doi.org/10.1016/j.gloenvcha.2013.03.009>.
- Grainger, Corbett, and Andrew Schreiber. 2019. "Discrimination in Ambient Air Pollution Monitoring?" *AEA Papers and Proceedings* 109 (May): 277–82.  
<https://doi.org/10.1257/pandp.20191063>.
- Greenstone, Michael, and B. Kelsey Jack. 2015. "Envirodevonomics: A Research Agenda for an Emerging Field." *Journal of Economic Literature* 53 (1): 5–42.  
<https://doi.org/10.1257/jel.53.1.5>.
- Hausman, Catherine, and Samuel Stolper. 2020. "Inequality, Information Failures, and Air Pollution." *NBER Working Paper*. <https://doi.org/10.3386/w26682>.
- Haybron, Daniel M. 2008. "Philosophy and the Science of Subjective Well-Being." In *The Science of Subjective Well-Being*, edited by Eid Michael and Larsen J. Randy,

- 17–43. New York: the Guilford Press.
- He, Jiaxiu, Haoming Liu, and Alberto Salvo. 2019. “Severe Air Pollution and Labor Productivity: Evidence from Industrial Towns in China.” *American Economic Journal: Applied Economics* 11 (1): 173–201.
- Hickman, Jonathan E., Niels Andela, Kostas Tsigaridis, Corinne Galy-Lacaux, Money Ossouhou, and Susanne E. Bauer. 2021. “Reductions in NO<sub>2</sub> Burden over North Equatorial Africa from Decline in Biomass Burning in Spite of Growing Fossil Fuel Use, 2005 to 2017.” *Proceedings of the National Academy of Sciences of the United States of America* 118 (7). <https://doi.org/10.1073/pnas.2002579118>.
- Jebb, Andrew T, Louis Tay, Ed Diener, and Shigehiro Oishi. 2018. “Happiness , Income Satiation and Turning Points around the World.” *Nature Human Behaviour* 2 (January). <https://doi.org/10.1038/s41562-017-0277-0>.
- Kahneman, Daniel, Schakade A. David, Fisher Claude, Krueger B. Alan, and Krilla Amy. 2010. “The Structure of Well-Being in Two Cities.” In *International Differences in Well-Being*, edited by Ed Diener, John F. Helliwell, and Daniel Kahneman, 2nd ed., 16–33. New York: Oxford University Press.
- Kahneman, Daniel, Peter P. Wakker, and Rakesh Sarin. 1997. “Back to Bentham? Explorations of Experienced Utility.” *Quarterly Journal of Economics* 112 (2): 375–405. <https://doi.org/10.1162/003355397555235>.
- Levinson, Arik. 2012. “Valuing Public Goods Using Happiness Data : The Case of Air Quality.” *Journal of Public Economics* 96 (9–10): 869–80. <https://doi.org/10.1016/j.jpubeco.2012.06.007>.
- Levinson, Arik. 2020a. “Happiness and Air Pollution.” In *Handbook on Wellbeing, Happiness and the Environment*, edited by Madison David, Rehdanz Katrin, and Welsch Heinz, 164–65. Northampton: Edward Elgar Publishing.
- Levy, Ilan, Cristian Mihele, Gang Lu, Julie Narayan, and Jeffrey R. Brook. 2014. “Evaluating Multipollutant Exposure and Urban Air Quality: Pollutant Interrelationships, Neighborhood Variability, and Nitrogen Dioxide as a Proxy Pollutant.” *Environmental Health Perspectives* 122 (1): 65–72. <https://doi.org/10.1289/ehp.1306518>.
- Li, Xiangdong, Ling Jin, and Haidong Kan. 2019. “Air Pollution: A Global Problem Needs Local Fixes.” *Nature* 570 (7762): 437–39. <https://doi.org/10.1038/d41586-019-01960-7>.
- Li, Yuan, Dabo Guan, Shu Tao, Xuejun Wang, and Kebin He. 2018. “A Review of Air Pollution Impact on Subjective Well-Being : Survey versus Visual Psychophysics.” *Journal of Cleaner Production* 184: 959–68. <https://doi.org/10.1016/j.jclepro.2018.02.296>.
- Li, Yuan, Dabo Guan, Yanni Yu, Stephen Westland, Xuejun Wang, Kebin He, Shu Tao, Daoping Wang, and Jing Meng. 2019. “A Psychophysical Measurement on Subjective Well-Being and Air Pollution.” *Nature Communications* 10 (1): 1–8. <https://doi.org/10.1038/s41467-019-13459-w>.



- Liu, Xiaojun, Hui Zhu, Yongxin Hu, Sha Feng, Yuanyuan Chu, Yanyan Wu, Chiyu Wang, Yuxuan Zhang, Zhaokang Yuan, and Yuanan Lu. 2016. "Public's Health Risk Awareness on Urban Air Pollution in Chinese Megacities: The Cases of Shanghai, Wuhan and Nanchang." *International Journal of Environmental Research and Public Health* 13 (9): 845. <https://doi.org/10.3390/ijerph13090845>.
- NACIS (North American Cartographic Information Society). n.d. "Coastline Boundaries." Accessed January 7, 2021. <http://www.naturalearthdata.com/downloads/10m-physical-vectors/10m-coastline/>.
- NASA. and. "MERRA-2." Accessed January 7, 2021a. <https://gmao.gsfc.nasa.gov/reanalysis/MERRA-2/>.
- NASA, Earth. n.d. "Global Roads Open Access 2010." Accessed January 7, 2021b. <https://sedac.ciesin.columbia.edu/data/set/groads-global-roads-open-access-v1>.
- Oishi, Shigehiro. 2010. "Culture and Well-Being: Conceptual and Methodological Issues." In *International Differences in Well-Being*, edited by Ed Diener, F. Helliwell John, and Kahneman Daniel, 34–69. New York, USA: Oxford University Press.
- Oltra, Christian, and Roser Sala. 2018. "Perception of Risk from Air Pollution and Reported Behaviors: A Cross-Sectional Survey Study in Four Cities." *Journal of Risk Research* 21 (7): 869–84. <https://doi.org/10.1080/13669877.2016.1264446>.
- Oswald, Yannick, Anne Owen, and Julia K. Steinberger. 2020. "Large Inequality in International and Intranational Energy Footprints between Income Groups and across Consumption Categories." *Nature Energy* 5 (3): 231–39. <https://doi.org/10.1038/s41560-020-0579-8>.
- Richter, A., M. Begoin, A. Hilboll, and J. P. Burrows. 2011. "An Improved NO<sub>2</sub> Retrieval for the GOME-2 Satellite Instrument." *Atmospheric Measurement Techniques* 4 (6): 1147–59. <https://doi.org/10.5194/amt-4-1147-2011>.
- Schwartz, Joel, Marie-abele Abele Bind, and Petros Koutrakis. 2017. "Estimating Causal Effects of Local Air Pollution on Daily Deaths: Effect of Low Levels." *Environmental Health Perspectives* 125 (1): 23–29. <https://doi.org/10.1289/EHP232>.
- Stiglitz, Joseph, Amartya Sen, and Fitoussi Jean-Paul. 2009. "Report by the Commission on the Measurement of Economic Performance and Social Progress."
- Stock, James H., and Motohiro Yogo. 2005. "Testing for Weak Instruments in Linear IV Regression." In *Identification and Inference for Econometric Models: Essays in Honor of Thomas Rothenberg*, 80–108. New York, USA: Cambridge University Press. <https://doi.org/10.1017/CBO9780511614491.006>.
- Tessum, Christopher W., Joshua S. Apte, Andrew L. Goodkind, Nicholas Z. Muller, Kimberley A. Mullins, David A. Paoletta, Stephen Polasky, et al. 2019. "Inequity in Consumption of Goods and Services Adds to Racial-Ethnic Disparities in Air Pollution Exposure." *Proceedings of the National Academy of Sciences of the United States of America*. <https://doi.org/10.1073/pnas.1818859116>.

- Welsch, Heinz. 2006. "Environment and Happiness: Valuation of Air Pollution Using Life Satisfaction Data." *Ecological Economics* 58 (4): 801–13.  
<https://doi.org/10.1016/j.ecolecon.2005.09.006>.
- Xiang, Yan, Tianshu Zhang, Jianguo Liu, Lihui Lv, Yunsheng Dong, and Zhenyi Chen. 2019. "Atmosphere Boundary Layer Height and Its Effect on Air Pollutants in Beijing during Winter Heavy Pollution." *Atmospheric Research* 215 (May 2018): 305–16. <https://doi.org/10.1016/j.atmosres.2018.09.014>.
- Zhang, Xin, Xiaobo Zhang, and Xi Chen. 2017. "Happiness in the Air: How Does a Dirty Sky Affect Mental Health and Subjective Well-Being?" *Journal of Environmental Economics and Management* 85 (September): 81–94.  
<https://doi.org/10.1016/j.jeem.2017.04.001>.
- Zivin, Joshua Graff, and Matthew Neidell. 2009. "Days of Haze : Environmental Information Disclosure and Intertemporal Avoidance Behavior." *Journal of Environmental Economics & Management* 58 (2): 119–28.  
<https://doi.org/10.1016/j.jeem.2009.03.001>.
- Zivin, Joshua Graff, and Matthew Neidell. 2012. "The Impact of Pollution on Worker Productivity." *American Economic Review* 102 (7): 3652–73.  
<https://doi.org/10.1257/aer.102.7.3652>.
- Zivin, Joshua Graff, and Matthew Neidell. 2013. "Environment , Health , and Human Capital." *Journal of Economic Literature* 51 (3): 689–730.

## Appendix 2

### Appendix 2 - Section A: Supplementary tables and figures

Table 2-3 Countries summary statistics

| Country name | Survey period |           | Observations | Frequency | National indices in 2016 |            |                 |                      |                    |                                      |                  |
|--------------|---------------|-----------|--------------|-----------|--------------------------|------------|-----------------|----------------------|--------------------|--------------------------------------|------------------|
|              | Start         | End       |              |           | Level of development     | GINI index | Unequal country | National GDP (USD B) | Productive Country | National NO2 VCD (10 <sup>15</sup> ) | Polluted Country |
| Australia    | 2/10/2016     | 2/22/2016 | 2026         | 2.1%      | Developed                | 34.4       | No              | 1175.337             | No                 | 27100                                | Yes              |
| Brazil       | 7/23/2015     | 7/26/2015 | 2258         | 2.5%      | Developing               | 53.3       | Yes             | 2978.317             | No                 | 18900                                | Yes              |
| Canada       | 9/1/2016      | 9/13/2016 | 1330         | 1.5%      | Developed                | 38.8       | No              | 1678                 | No                 | 22700                                | Yes              |
| Chile        | 7/24/2015     | 7/28/2015 | 1178         | 1.3%      | Developing               | 44.4       | Yes             | 431.9483             | No                 | 3450                                 | No               |
| China        | 1/12/2016     | 2/29/2016 | 19928        | 22.3%     | Developing               | 38.5       | Yes             | 18595.36             | Yes                | 40900                                | Yes              |
| Colombia     | 7/24/2015     | 7/27/2015 | 1089         | 1.2%      | Developing               | 50.8       | Yes             | 690.7023             | No                 | 2140                                 | No               |
| Czech        | 3/8/2017      | 3/16/2017 | 1178         | 1.3%      | Developed                | 25.4       | No              | 390.7661             | No                 | 437                                  | No               |
| France       | 8/26/2016     | 9/7/2016  | 2133         | 2.4%      | Developed                | 31.9       | No              | 2948.053             | No                 | 2100                                 | No               |
| Germany      | 8/26/2016     | 9/7/2016  | 3168         | 3.6%      | Developed                | 31.9       | No              | 4276.38              | No                 | 1950                                 | No               |
| Greece       | 8/31/2016     | 9/12/2016 | 1357         | 1.5%      | Developed                | 35         | No              | 308.2042             | No                 | 469                                  | No               |
| Hungary      | 3/8/2017      | 3/15/2017 | 1357         | 1.5%      | Developed                | 30.3       | No              | 277.0553             | No                 | 467                                  | No               |
| India        | 7/25/2015     | 8/11/2015 | 6559         | 7.4%      | Developing               | 37.8       | No              | 7736.024             | Yes                | 11700                                | No               |
| Indonesia    | 7/18/2015     | 7/23/2015 | 2365         | 2.7%      | Developing               | 39.9       | Yes             | 2754.479             | No                 | 3580                                 | No               |
| Italy        | 8/29/2016     | 9/10/2016 | 2106         | 2.4%      | Developed                | 35.2       | No              | 2488.007             | No                 | 1370                                 | No               |
| Japan        | 7/14/2015     | 8/5/2015  | 10102        | 11.3%     | Developed                | 32.9       | No              | 5070.386             | Yes                | 1560                                 | No               |
| Malaysia     | 7/23/2015     | 7/29/2015 | 1080         | 1.2%      | Developing               | 41         | Yes             | 783.8876             | No                 | 615                                  | No               |
| Myanmar      | 8/19/2015     | 9/3/2015  | 1,079        | 1.2%      | Developing               | 30.7       | No              | 238.09               | No                 | 1861                                 | No               |
| Mexico       | 7/24/2015     | 7/27/2015 | 1669         | 1.9%      | Developing               | 46.3       | Yes             | 2418.842             | No                 | 5730                                 | No               |
| Netherlands  | 8/29/2016     | 9/10/2016 | 1374         | 1.5%      | Developed                | 28.2       | No              | 1115.57              | No                 | 206                                  | No               |

|               |           |           |       |       |            |      |     |          |     |       |     |
|---------------|-----------|-----------|-------|-------|------------|------|-----|----------|-----|-------|-----|
| Philippines   | 7/15/2015 | 7/22/2015 | 1669  | 1.9%  | Developing | 44.4 | Yes | 798.7352 | No  | 597   | No  |
| Poland        | 3/8/2017  | 3/17/2017 | 2222  | 2.5%  | Developed  | 31.2 | No  | 1090.629 | No  | 1610  | No  |
| Romania       | 3/8/2017  | 3/18/2017 | 286   | 0.3%  | Developed  | 34.4 | No  | 497.2495 | No  | 1100  | No  |
| Russia        | 8/31/2015 | 9/14/2015 | 2118  | 2.2%  | Developing | 36.8 | No  | 3783.818 | No  | 54500 | Yes |
| South Africa  | 7/15/2015 | 7/23/2015 | 1107  | 1.2%  | Developing | 63   | Yes | 714.0011 | No  | 5280  | No  |
| Spain         | 8/26/2016 | 9/7/2016  | 2115  | 2.4%  | Developed  | 35.8 | No  | 1792.128 | No  | 1680  | No  |
| Sweden        | 8/31/2016 | 9/12/2016 | 1330  | 1.5%  | Developed  | 29.6 | No  | 517.153  | No  | 1330  | No  |
| Thailand      | 7/18/2015 | 7/23/2015 | 1116  | 1.3%  | Developing | 36.9 | No  | 1158.565 | No  | 1420  | No  |
| Turkey        | 3/7/2017  | 3/20/2017 | 1954  | 2.2%  | Developing | 41.9 | Yes | 2108.026 | No  | 2790  | No  |
| United States | 8/16/2016 | 8/28/2016 | 10683 | 12.0% | Developed  | 41.4 | Yes | 18596.69 | Yes | 30200 | Yes |
| Vietnam       | 7/18/2015 | 7/28/2015 | 1330  | 1.5%  | Developing | 35.3 | No  | 667.7238 | No  | 910   | No  |

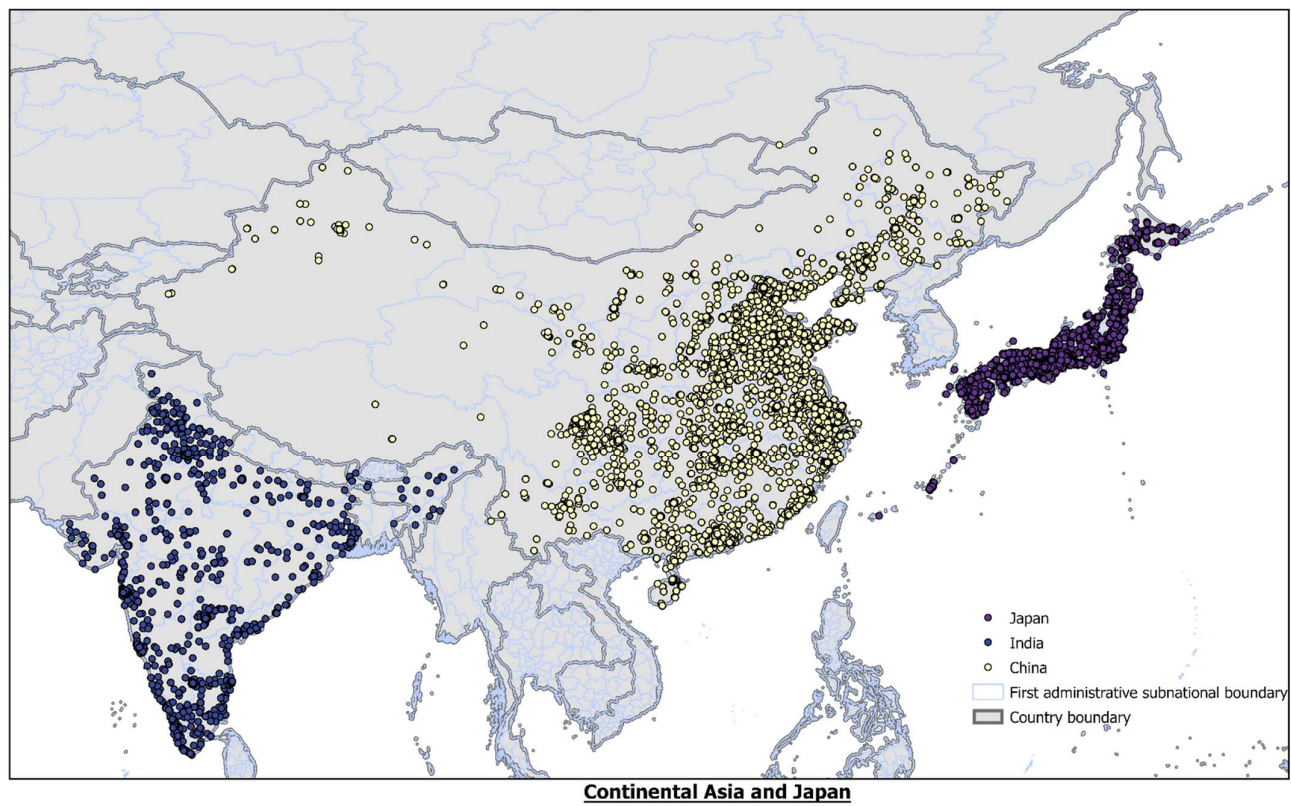
Table 2-4 Summary statistics for categorical variables

| <b>Marital status</b>              | <b>Observations</b> | <b>Frequency</b> |
|------------------------------------|---------------------|------------------|
| Single                             | 26,431              | 29.6%            |
| Married/ Cohabitation              | 54,440              | 58.17%           |
| Separate/Divorce/Widowed           | 10,903              | 12.21%           |
| Total                              | 89,266              |                  |
| <b>Health</b>                      |                     |                  |
| Very poor                          | 1,180               | 1.32%            |
| Poor                               | 7,602               | 8.52%            |
| Neither                            | 17,086              | 19.14%           |
| Good                               | 44,889              | 50.29%           |
| Very good                          | 18,509              | 20.73%           |
| Total                              | 89,266              |                  |
| <b>Education</b>                   |                     |                  |
| No education                       | 730                 | 0.82%            |
| Dropped out of primary school      | 534                 | 0.60%            |
| Completed primary school           | 1,423               | 1.59%            |
| Completed junior high school       | 4,597               | 5.15%            |
| Completed high school              | 17,423              | 19.52%           |
| Completed professional school      | 9,184               | 10.29%           |
| Completed junior college           | 10,954              | 12.27%           |
| Completed university               | 33,027              | 37.00%           |
| Completed master's education       | 8,454               | 9.47%            |
| Completed doctorate education      | 1,676               | 1.88%            |
| Other                              | 1,255               | 1.41%            |
| Do not know                        | 9                   | 0.01%            |
| Total                              | 89,266              |                  |
| <b>Occupation</b>                  |                     |                  |
| Full time                          | 36,966              | 41.41%           |
| Part-time                          | 7,181               | 8.04%            |
| Company Owner                      | 1,901               | 2.13%            |
| Government employee                | 2,809               | 3.15%            |
| Professional (Doctor / Lawyer ...) | 3,146               | 3.52%            |
| Self-employed                      | 7,078               | 7.93%            |
| Student                            | 5,983               | 6.70%            |
| Housewife / husband                | 8,005               | 8.97%            |
| Unemployed                         | 8,258               | 9.25%            |
| Other                              | 7,939               | 8.89%            |
| Total                              | 89,266              |                  |

Table 2-5 Summary statistics

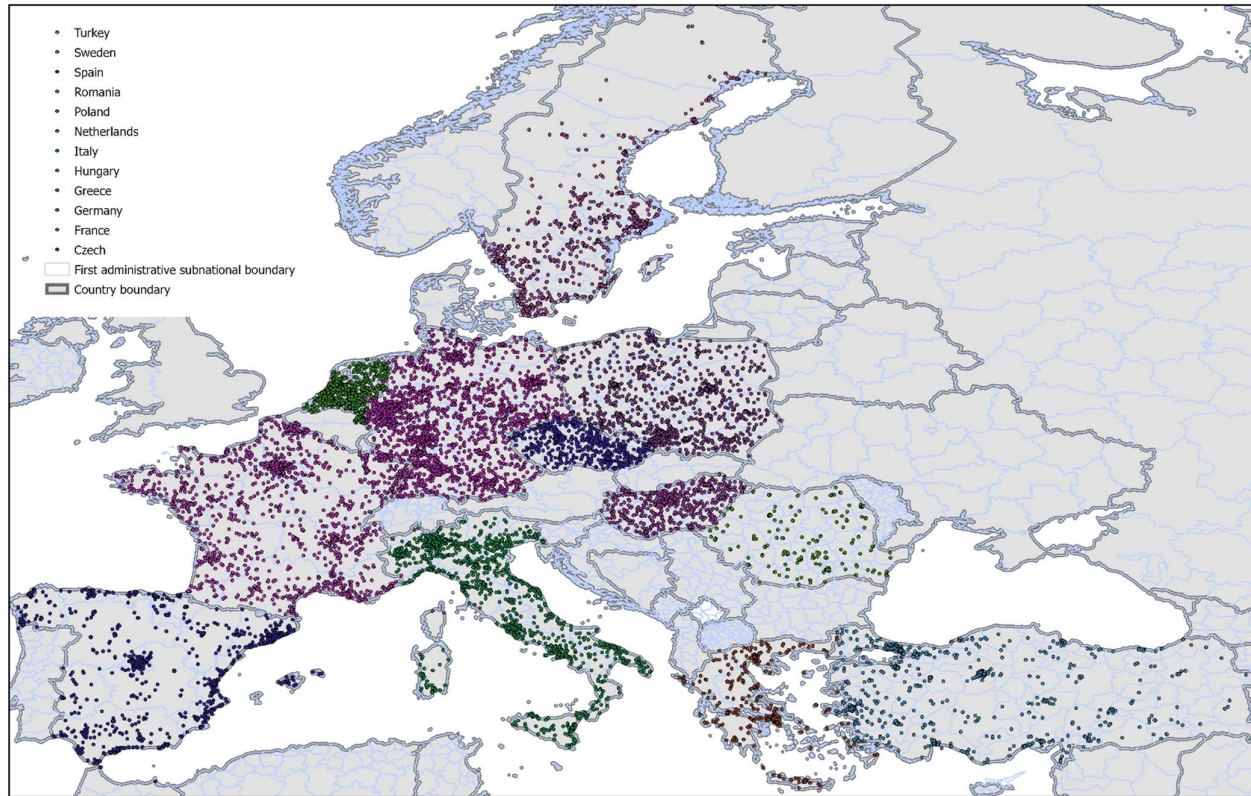
| Variable                                                      | Observations | Mean       | Standard Deviation | Minimum   | Maximum  |
|---------------------------------------------------------------|--------------|------------|--------------------|-----------|----------|
| Average levels of Pollution ( $10^{15}$ mol/cm <sup>2</sup> ) | 89,266       | 3.917838   | 1.724635           | 0.7822267 | 13.1228  |
| Change in pollution ( $10^{15}$ mol/cm <sup>2</sup> )         | 89,266       | -0.1034272 | 0.9108634          | -4.242116 | 3.653321 |
| Average PBLH (meter)                                          | 89,266       | 731.1385   | 203.7959           | 181.6043  | 2222.081 |
| Change in pollution in PBLH height (meters)                   | 89,266       | 30.51531   | 119.4541           | -812.3903 | 857.3494 |
| Income (in USD PPP)                                           | 89,266       | 3546.768   | 3370.478           | 208.7     | 57365.39 |
| Logarithm of income                                           | 89,266       | 7.759118   | 1.019166           | 2.564702  | 10.9572  |
| Minimum temperature (°C)                                      | 89,266       | 12.70824   | 10.72742           | -32.6     | 29.1     |
| Maximum temperature (°C)                                      | 89,266       | 21.78399   | 10.76969           | -21.7     | 42.6     |
| Precipitation (mm)                                            | 89,266       | 43.43647   | 256.7782           | 0         | 1145     |
| Average wind speed (m/s)                                      | 89,266       | 2.601603   | .861818            | .4        | 6.9      |
| Inverse distance to the sea                                   | 89,266       | 0.0054513  | 0.4820699          | 0         | 45.53007 |
| Distance to the closest road (km)                             | 89,266       | 1517.818   | 9472.33            | 0         | 570620.4 |
| Gender (1 as male and 2 as female)                            | 89,266       | 1.495081   | 0.4999785          | 1         | 2        |
| Number of kids                                                | 89,266       | 1.175566   | 1.19982            | 0         | 10       |
| Age                                                           | 89,266       | 42.73494   | 14.95702           | 18        | 97       |

Figure 2-4 Locations of the respondents (1)



India (6559 observations across 33 regions)  
China (19,928 observations across 35 regions)  
Japan (10,102 observations across 47 regions)

Figure 2-5 Locations of the respondents (2)



**Sample distribution in Europe**

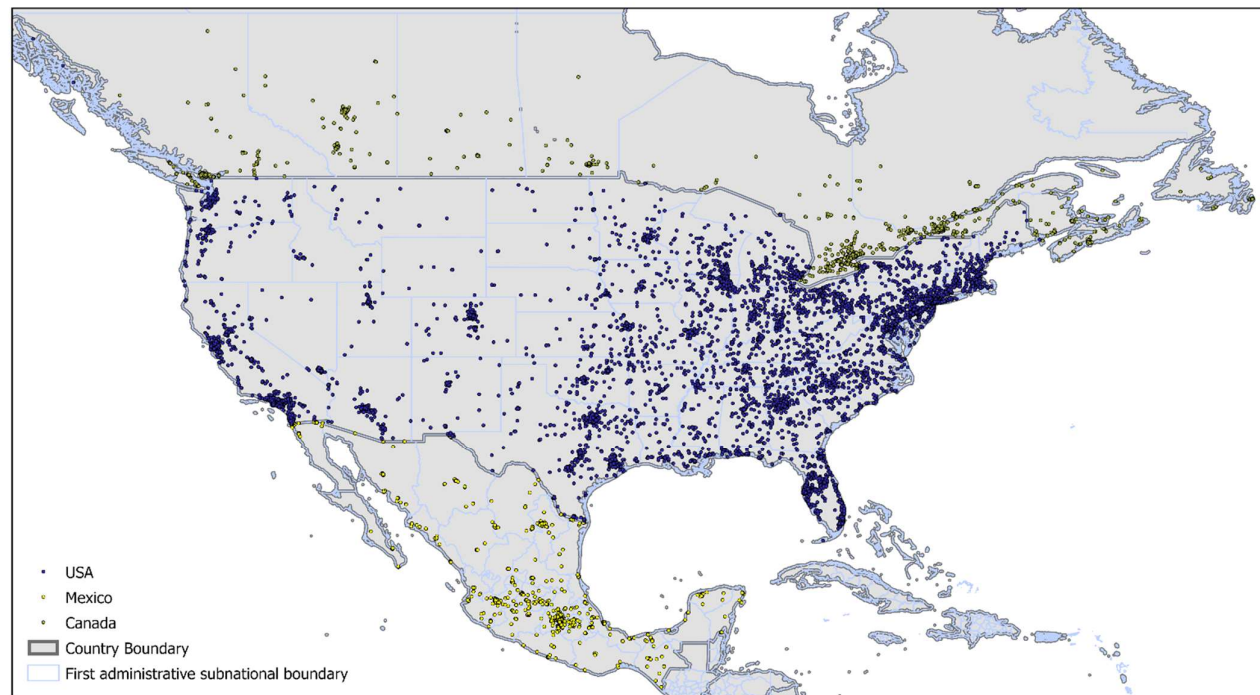
Czech Republic (1,178 observations across 14 regions)  
 Germany (3,168 observations across 16 regions)  
 Spain (2,115 observations across 19 regions)  
 France (2,133 observations across 14 regions)

Greece (1,357 observations across 8 regions)  
 Hungary (1,357 observations across 20 regions)  
 Italy (2,106 observations across 20 regions)  
 Netherlands (1,374 observations across 12 regions)

Romania (286 observations across 40 regions)  
 Sweden (1330 observations across 21 regions)  
 Turkey (1954 observations across 75 regions)  
 Poland (2218 observations across 17 regions)



Figure 2-6 Locations of the respondents (3)



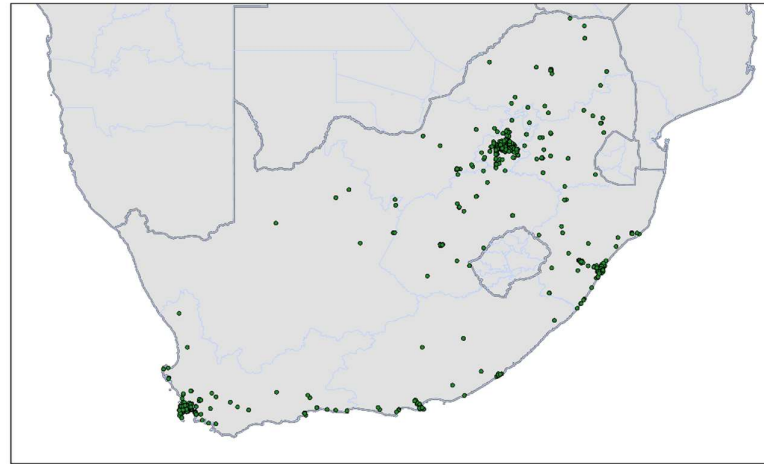
**Sample distribution in North America**

Canada (1,332 observations across 12 regions)  
The United States of America (10,684 observations across 51 states)  
Mexico (1,667 observations across 32 regions)

Figure 2-7 Locations of the respondents (4)

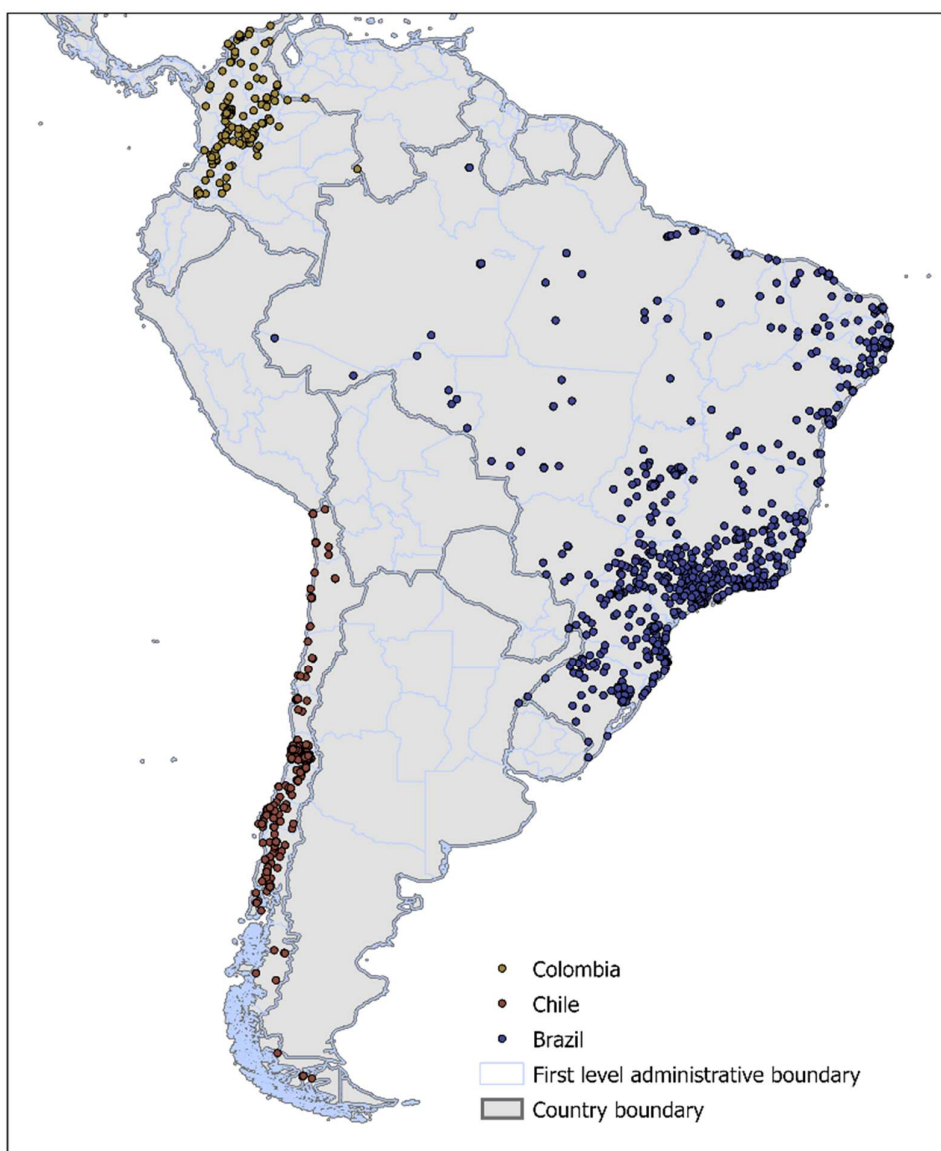


Russia  
Russia (2,118 observations across 74 regions)



South Africa  
South Africa (1,107 observations across 9 regions)

Figure 2-8 Locations of the respondents (5)

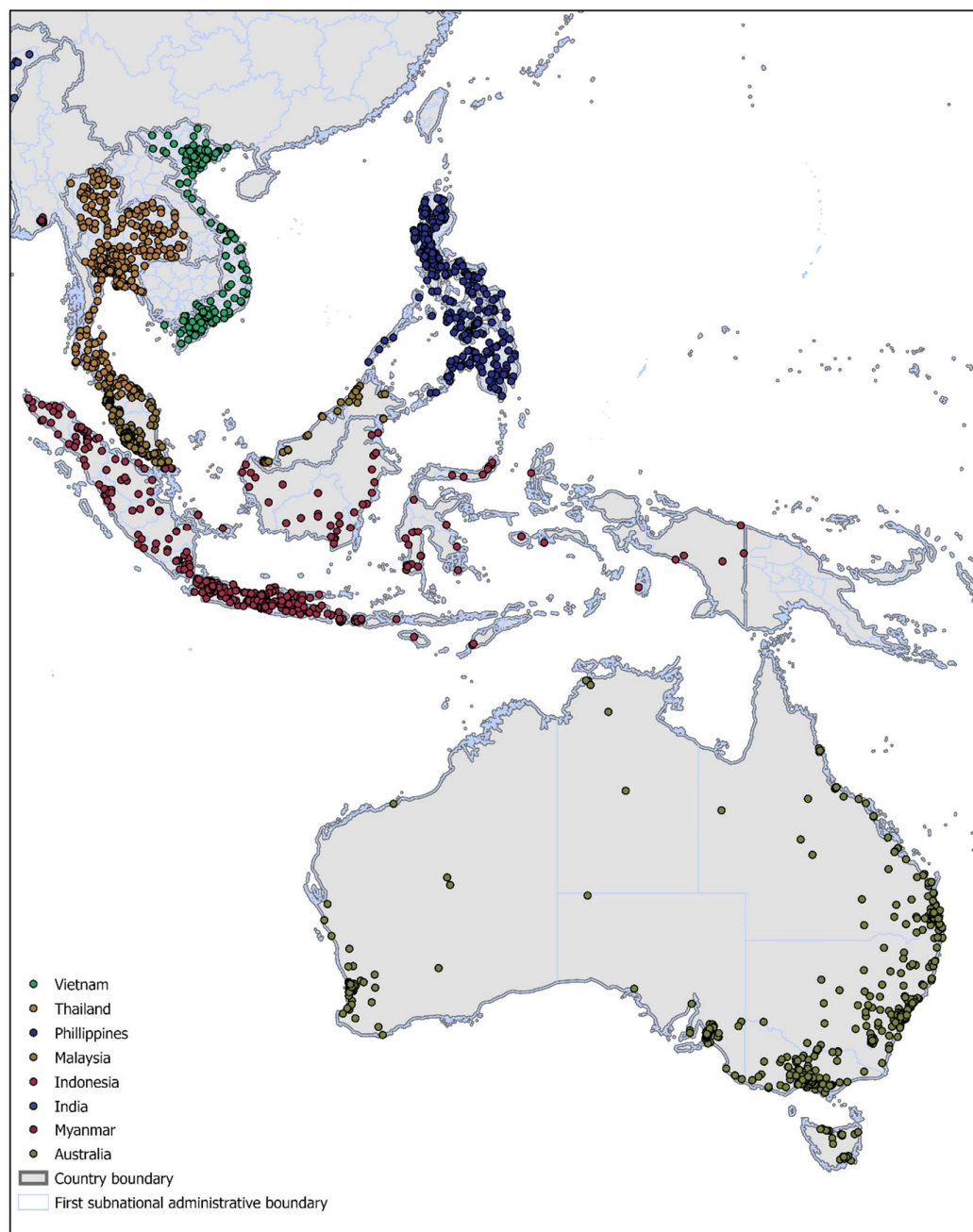


**Sample distribution in South America**

Brazil (2,258 observations across 27 regions)  
Chile (1,174 observations across 16 regions)  
Colombia (1,088 observations across 25 regions)

**Note:** Exact statistics per province is available pon request to author

Figure 2-9 Locations of the respondents (6)



#### **South East Asia and Australia**

Indonesia (2,265 observations across 32 regions)  
 India (6559 observations across 33 regions)  
 Malaysia (1,180 observations across 16 regions)  
 Myanmar (1,079 observations in 2 regions)

Philippines (1,669 observations across 69 regions)  
 Thailand (1,116 observations across 74 regions)  
 Vietnam (1,330 observations across 61 regions)  
 Australia (2026 observations across 9 regions)

Figure 2-10 Histogram of subjective well-being

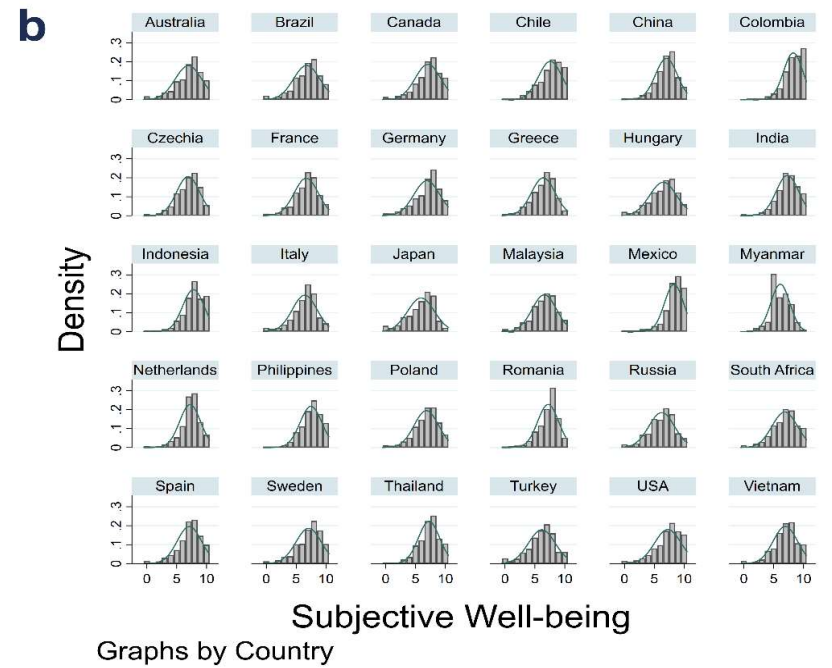
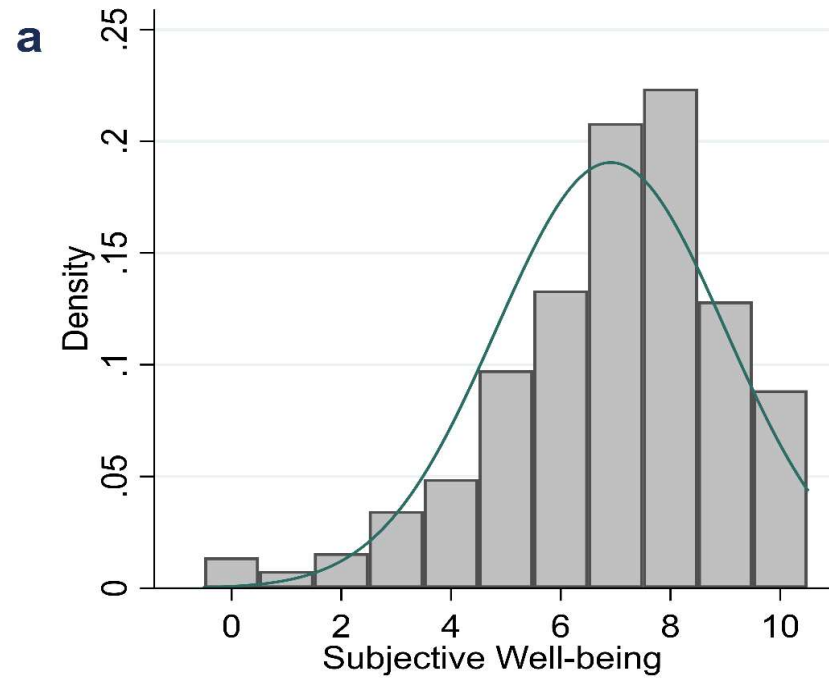


Figure 2-11 Well-being and pollution

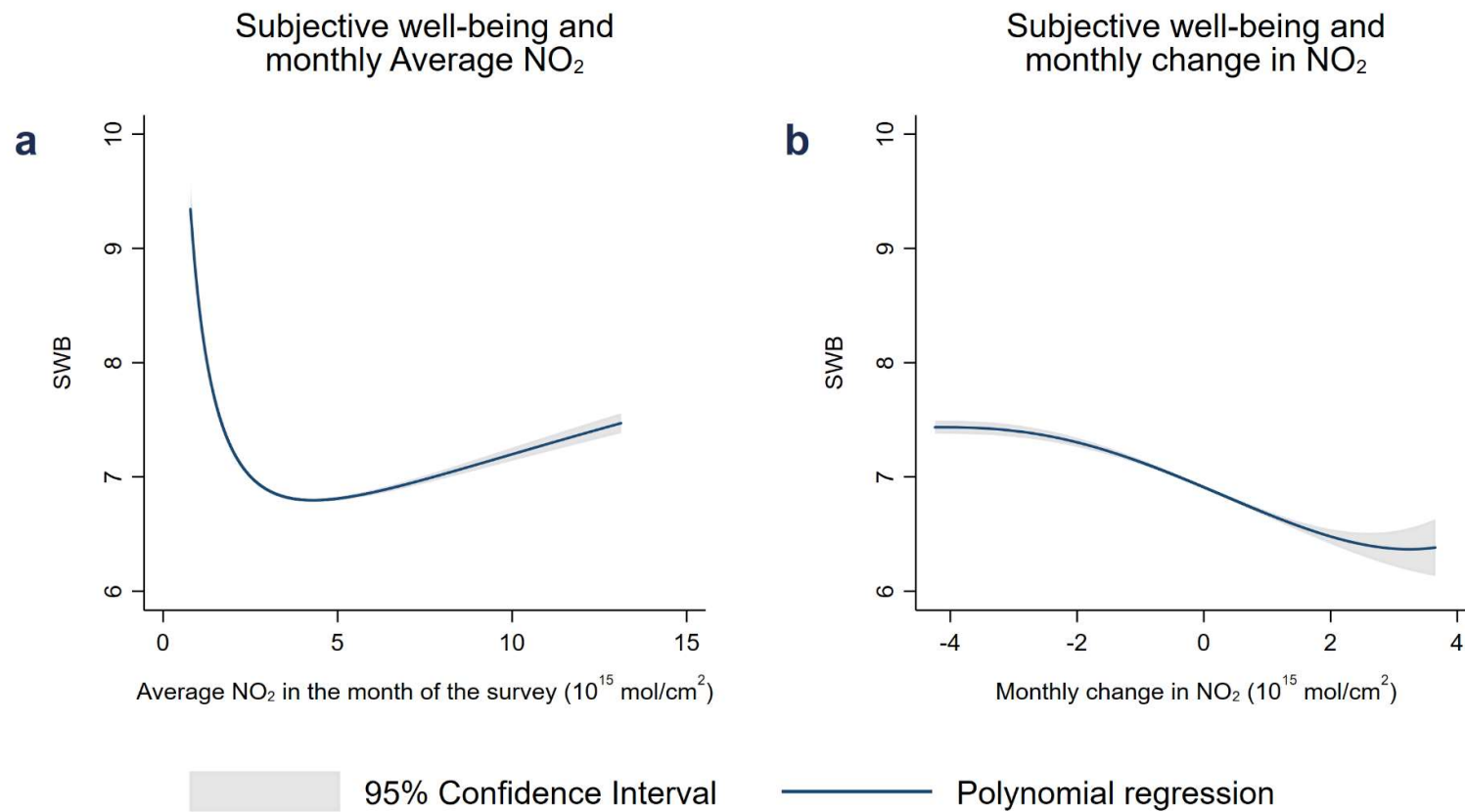


Figure 2-12 Wellbeing and pollution by country

Investigating the impact of pollution on well-being by using the 1-month changes in pollution rather than the average pollution provide sensible results for each country taken separately. Although, this impact does not match the common sense in Sweden and in Russia, I believe that the consistency of this approach in most countries bolster the identification strategy in of the analysis.

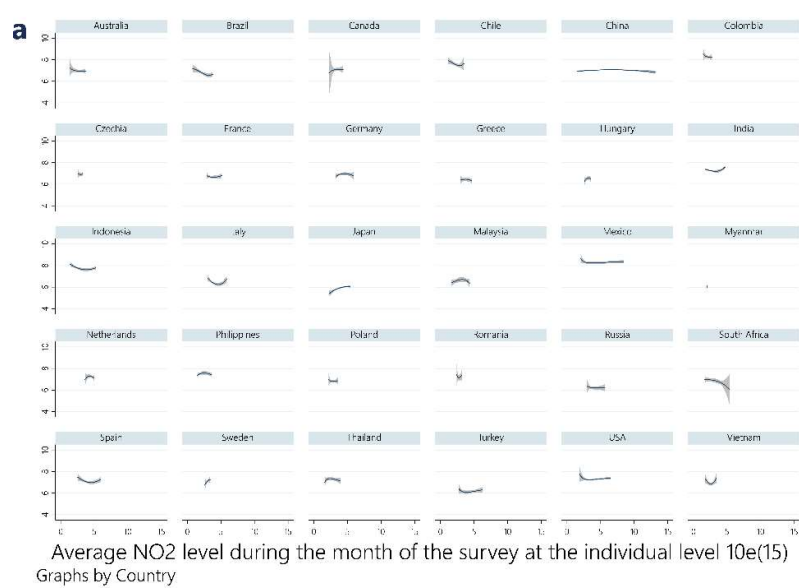


Table 2-6 Full estimation results of the estimated impact of monthly changes in air pollution on well-being

| Panel A                      | The dependent variable is subjective well-being (Second stage regression of 2SLS) |                           |                            |                            |                            |                                                                |
|------------------------------|-----------------------------------------------------------------------------------|---------------------------|----------------------------|----------------------------|----------------------------|----------------------------------------------------------------|
|                              | (1)                                                                               | (2)                       | (3)                        | (4)                        | (5)                        | (6)                                                            |
| VARIABLES                    | Regressions OLS                                                                   | Regressions 2SLS          | Baseline OLS               | Baseline 2SLS              | Ordered Probit Model 2SLS  | Regressions 2SLS – controlling average monthly NO <sub>2</sub> |
| Change in NO <sub>2</sub>    | -0.0899***<br>(0.0221)                                                            | -0.254***<br>(0.0614)     | -0.0794***<br>(0.0130)     | -0.224***<br>(0.0512)      | -0.131**<br>(0.0548)       | -0.225***<br>(0.0523)                                          |
| Average NO <sub>2</sub>      |                                                                                   |                           |                            |                            |                            | -0.000751<br>(0.00631)                                         |
| log(income)                  |                                                                                   | 2.719***<br>(0.198)       | 0.235***<br>(0.0115)       | 0.227***<br>(0.00897)      | 0.121***<br>(0.00727)      | 0.227***<br>(0.00910)                                          |
| Temperature Min              | 0.0125<br>(0.0151)                                                                | 0.0299***<br>(0.00904)    | 0.0205**<br>(0.00896)      | 0.0363***<br>(0.00794)     | 0.0235***<br>(0.00692)     | 0.0364***<br>(0.00810)                                         |
| Temperature Min squared      | 0.000187<br>(0.000486)                                                            | -0.000351<br>(0.000273)   | -0.000693**<br>(0.000294)  | -0.00117***<br>(0.000236)  | -0.000747***<br>(0.000227) | -0.00118***<br>(0.000233)                                      |
| Temperature Max              | -0.00426<br>(0.0309)                                                              | -0.0262**<br>(0.0133)     | -0.0343*<br>(0.0177)       | -0.0541***<br>(0.0117)     | -0.0365***<br>(0.0121)     | -0.0543***<br>(0.0115)                                         |
| Temperature Max squared      | 2.10e-05<br>(0.000566)                                                            | 0.000566*<br>(0.000297)   | 0.000922***<br>(0.000342)  | 0.00141***<br>(0.000259)   | 0.000930***<br>(0.000255)  | 0.00142***<br>(0.000256)                                       |
| Precipitation level          | 9.28e-05<br>(0.000312)                                                            | 0.000462<br>(0.000281)    | 0.000694***<br>(0.000262)  | 0.00102***<br>(0.000258)   | 0.000569***<br>(0.000186)  | 0.00102***<br>(0.000273)                                       |
| Precipitation level squared  | -6.47e-07<br>(4.74e-07)                                                           | -1.14e-06**<br>(4.53e-07) | -1.18e-06***<br>(4.19e-07) | -1.63e-06***<br>(4.41e-07) | -9.71e-07***<br>(2.93e-07) | -1.62e-06***<br>(4.56e-07)                                     |
| Wind speed                   | -0.136**<br>(0.0645)                                                              | -0.105**<br>(0.0445)      | -0.143***<br>(0.0454)      | -0.116***<br>(0.0408)      | -0.0973***<br>(0.0297)     | -0.116***<br>(0.0408)                                          |
| Wind speed squared           | 0.0249**<br>(0.0104)                                                              | 0.0225***<br>(0.00725)    | 0.0235***<br>(0.00745)     | 0.0216***<br>(0.00661)     | 0.0173***<br>(0.00476)     | 0.0216***<br>(0.00663)                                         |
| Inverse distance to the sea  |                                                                                   |                           | -0.0225***<br>(0.00145)    | -0.0220<br>(0.0171)        | -0.0144***<br>(0.000927)   | -0.0220<br>(0.0171)                                            |
| Distance to the closest road |                                                                                   |                           | 1.26e-07<br>(3.39e-06)     | 2.65e-06<br>(2.26e-06)     | 1.45e-06<br>(1.99e-06)     | 2.64e-06<br>(2.34e-06)                                         |



|                                    |                           |                           |                           |                           |
|------------------------------------|---------------------------|---------------------------|---------------------------|---------------------------|
| Female                             | 0.129***<br>(0.0163)      | 0.125***<br>(0.0135)      | 0.0720***<br>(0.00996)    | 0.125***<br>(0.0135)      |
| Number of kids                     | 0.0412***<br>(0.00688)    | 0.0409***<br>(0.00682)    | 0.0258***<br>(0.00418)    | 0.0408***<br>(0.00682)    |
| Age                                | -0.0609***<br>(0.00390)   | -0.0604***<br>(0.00326)   | -0.0329***<br>(0.00226)   | -0.0604***<br>(0.00326)   |
| Age square                         | 0.000803***<br>(4.16e-05) | 0.000801***<br>(3.55e-05) | 0.000439***<br>(2.33e-05) | 0.000801***<br>(3.55e-05) |
| <b>MARITAL STATUS DUMMY</b>        |                           |                           |                           |                           |
| Married/ Cohabitation              | 0.451***<br>(0.0318)      | 0.464***<br>(0.0209)      | 0.252***<br>(0.0169)      | 0.464***<br>(0.0209)      |
| Separate/Divorce/Widowed           | 0.0657*<br>(0.0338)       | 0.0749***<br>(0.0280)     | 0.0326*<br>(0.0179)       | 0.0750***<br>(0.0280)     |
| <b>HEALTH DUMMIES (Very bad as</b> |                           |                           |                           |                           |
| Poor                               | 1.439***<br>(0.129)       | 1.439***<br>(0.0943)      | 0.627***<br>(0.0587)      | 1.439***<br>(0.0943)      |
| Neither                            | 2.193***<br>(0.127)       | 2.192***<br>(0.0915)      | 0.970***<br>(0.0593)      | 2.192***<br>(0.0915)      |
| Good                               | 3.210***<br>(0.116)       | 3.208***<br>(0.0907)      | 1.539***<br>(0.0529)      | 3.208***<br>(0.0907)      |
| Very good                          | 4.125***<br>(0.114)       | 4.127***<br>(0.0916)      | 2.198***<br>(0.0535)      | 4.127***<br>(0.0916)      |
| <b>EDUCATION DUMMIES (No</b>       |                           |                           |                           |                           |
| Dropped out of primary school      | 0.0525<br>(0.130)         | 0.0576<br>(0.114)         | 0.0292<br>(0.0844)        | 0.0576<br>(0.114)         |
| Completed primary school           | 0.107<br>(0.126)          | 0.111<br>(0.121)          | 0.0236<br>(0.0777)        | 0.111<br>(0.121)          |
| Completed junior high school       | -0.0701<br>(0.106)        | -0.0683<br>(0.0999)       | -0.0775<br>(0.0618)       | -0.0683<br>(0.0999)       |
| Completed high school              | -0.0773<br>(0.0860)       | -0.0742<br>(0.0890)       | -0.0802<br>(0.0508)       | -0.0742<br>(0.0890)       |
| Completed professional school      | -0.115<br>(0.0815)        | -0.110<br>(0.0845)        | -0.101**<br>(0.0485)      | -0.110<br>(0.0845)        |

|                                   |                     |                     |                        |                        |                        |                        |
|-----------------------------------|---------------------|---------------------|------------------------|------------------------|------------------------|------------------------|
| Completed junior college          |                     |                     | -0.0879<br>(0.0857)    | -0.0854<br>(0.0858)    | -0.0879*<br>(0.0500)   | -0.0854<br>(0.0858)    |
| Completed university              |                     |                     | -0.135*<br>(0.0809)    | -0.121<br>(0.0853)     | -0.123**<br>(0.0487)   | -0.121<br>(0.0853)     |
| Completed master's education      |                     |                     | -0.0513<br>(0.0798)    | -0.0471<br>(0.0838)    | -0.0748<br>(0.0471)    | -0.0471<br>(0.0838)    |
| Completed doctorate education     |                     |                     | -0.00378<br>(0.0812)   | 0.00227<br>(0.0856)    | -0.0293<br>(0.0482)    | 0.00234<br>(0.0856)    |
| Other                             |                     |                     | 0.0756<br>(0.0995)     | 0.0771<br>(0.0956)     | 0.0477<br>(0.0635)     | 0.0772<br>(0.0956)     |
| <b>OCCUPATION DUMMIES (full</b>   |                     |                     |                        |                        |                        |                        |
| Part-time employee                |                     |                     | -0.0947***<br>(0.0292) | -0.0955***<br>(0.0265) | -0.0522***<br>(0.0173) | -0.0955***<br>(0.0265) |
| Company owner                     |                     |                     | 0.0193<br>(0.0485)     | -0.00371<br>(0.0457)   | 0.0197<br>(0.0308)     | -0.00406<br>(0.0456)   |
| Government employee               |                     |                     | 0.0310<br>(0.0385)     | 0.0331<br>(0.0363)     | 0.0145<br>(0.0226)     | 0.0330<br>(0.0364)     |
| Professional (Doctor / Lawyer...) |                     |                     | -0.0148<br>(0.0393)    | -0.00613<br>(0.0343)   | -0.00606<br>(0.0242)   | -0.00615<br>(0.0343)   |
| Self-employed                     |                     |                     | -0.176***<br>(0.0323)  | -0.175***<br>(0.0261)  | -0.0976***<br>(0.0191) | -0.175***<br>(0.0261)  |
| Student                           |                     |                     | -0.0515<br>(0.0378)    | -0.0318<br>(0.0363)    | -0.0351<br>(0.0224)    | -0.0316<br>(0.0364)    |
| Housewife / husband               |                     |                     | -0.0249<br>(0.0326)    | -0.0254<br>(0.0270)    | -0.0209<br>(0.0209)    | -0.0255<br>(0.0270)    |
| Unemployed                        |                     |                     | -0.406***<br>(0.0359)  | -0.404***<br>(0.0309)  | -0.203***<br>(0.0197)  | -0.404***<br>(0.0309)  |
| Other                             |                     |                     | -0.0446<br>(0.0364)    | -0.0477<br>(0.0316)    | 0.0118<br>(0.0226)     | -0.0478<br>(0.0316)    |
| Constant                          | 6.885***<br>(0.332) | 6.784***<br>(0.141) | 2.777***<br>(0.272)    | 2.719***<br>(0.198)    |                        | 2.722***<br>(0.202)    |
| Observations                      | 89,266              | 89,266              | 89,266                 | 89,266                 | 89,266                 | 89,266                 |
| R-squared                         | 0.065               | 0.062               | 0.279                  | 0.276                  | 0.074                  | 0.276                  |
| Weather variables                 | Yes                 | Yes                 | Yes                    | Yes                    | Yes                    | Yes                    |
| Country dummy                     | Yes                 | Yes                 | Yes                    | Yes                    | Yes                    | Yes                    |

| Personal characteristics                                                     | No | No                                                       | Yes | Yes                                                      | Yes                                                      | Yes                                                      |
|------------------------------------------------------------------------------|----|----------------------------------------------------------|-----|----------------------------------------------------------|----------------------------------------------------------|----------------------------------------------------------|
| <b>First stage regression for 2SLS. Dependent variable is changes in NO2</b> |    |                                                          |     |                                                          |                                                          |                                                          |
| PBLH change                                                                  |    | -0.00112***<br>(0.00003)                                 |     | -0.00121***<br>(0.00003)                                 | -0.00121***<br>(0.00003)                                 | -0.0012***<br>0.00003                                    |
| PBLH change squared                                                          |    | -3.22(10 <sup>-6</sup> )***<br>(1.24(10 <sup>-7</sup> )) |     | -3.76(10 <sup>-6</sup> )***<br>(1.46(10 <sup>-7</sup> )) | -3.76(10 <sup>-6</sup> )***<br>(1.46(10 <sup>-7</sup> )) | -3.63(10 <sup>-6</sup> )***<br>(1.54(10 <sup>-7</sup> )) |
| Control variables of the second stage                                        |    | Yes                                                      |     | Yes                                                      | Yes                                                      | Yes                                                      |
| F statistics of the instruments                                              |    | 783.554***                                               |     | 726.236***                                               | 726.236***                                               | 642.795***                                               |
| Wooldridge endogeneity test (H <sub>0</sub> :                                |    | 7.30458***                                               |     | 8.188***                                                 | -                                                        | 8.168***                                                 |
| Sargan test for over identification                                          |    | 2.6948 (p=0.101)                                         |     | 2.623 (p= 0.105)                                         | -                                                        | 2.601 (p=0.106)                                          |

\*\*\* p<0.01, \*\* p<0.05, \* p<0.1. <sup>a</sup> The standard errors of the ordered probit model estimations do not account for the first stage regression estimates density function. The standard errors of the coefficients

are shown in parentheses. These standard errors correspond to the Huber–White standard errors, except for the OLS estimation, where I cluster the standard errors using the second level of administrative division in each country provided by <https://gadm.org/>. This clustering adjustment is not necessary for the 2SLS estimations, because the PBHL changes are assumed to be random across areas. The Wu–Hausman endogeneity null hypothesis test assumes that pollution changes are exogenous, meanwhile, that of the Sargan test for over-identification assumed that the model is not over-identified.

Table 2-7 Full estimation results of the heterogeneous impact of air pollution

| VARIABLES                                              | (1)<br>Pollution<br>variation<br>interacted with<br>income | (2)<br>Pollution<br>changes<br>interacted<br>pollution level | (3)<br>Pollution<br>changes<br>interacted with<br>development | (4)<br>Average<br>pollution<br>interacted with<br>inequality | (5)<br>Average<br>pollution<br>interacted with<br>average pollution<br>at the country<br>level | (6)<br>Average<br>pollution<br>interacted with<br>national<br>productivity |
|--------------------------------------------------------|------------------------------------------------------------|--------------------------------------------------------------|---------------------------------------------------------------|--------------------------------------------------------------|------------------------------------------------------------------------------------------------|----------------------------------------------------------------------------|
| Change in NO <sub>2</sub>                              | -0.339***<br>(0.120)                                       | 1.483<br>(0.954)                                             | 0.752<br>(0.546)                                              | 1.735<br>(1.093)                                             | -1.604**<br>(0.780)                                                                            | -1.409**<br>(0.668)                                                        |
| Variation in NO <sub>2</sub> * log(income)             | 2.617 <sup>†</sup><br>(1.912)                              |                                                              |                                                               |                                                              |                                                                                                |                                                                            |
| Variation in NO <sub>2</sub> * Average NO <sub>2</sub> |                                                            | -0.221 <sup>††</sup><br>(0.123)                              |                                                               |                                                              |                                                                                                |                                                                            |
| Change in NO <sub>2</sub> * developing country         |                                                            |                                                              | -0.988 <sup>††</sup><br>(0.550)                               |                                                              |                                                                                                |                                                                            |
| Change in NO <sub>2</sub> * unequal country            |                                                            |                                                              |                                                               | -1.935 <sup>††</sup><br>(1.078)                              |                                                                                                |                                                                            |
| Change in NO <sub>2</sub> * country pollution          |                                                            |                                                              |                                                               |                                                              | 1.487 <sup>††</sup><br>(0.838)                                                                 |                                                                            |
| Change in NO <sub>2</sub> * productive country         |                                                            |                                                              |                                                               |                                                              |                                                                                                | 1.375 <sup>††</sup><br>(0.773)                                             |
| log(income)                                            | 0.386***<br>(0.117)                                        | 0.219***<br>(0.0104)                                         | 0.220***<br>(0.00992)                                         | 0.212***<br>(0.0124)                                         | 0.247***<br>(0.0143)                                                                           | 0.248***<br>(0.0146)                                                       |
| Average NO <sub>2</sub>                                |                                                            | -0.0154<br>(0.0106)                                          |                                                               |                                                              |                                                                                                |                                                                            |
| <b>WEATHER DATA</b>                                    |                                                            |                                                              |                                                               |                                                              |                                                                                                |                                                                            |
| Temperature Min                                        | 0.0872**<br>(0.0387)                                       | -0.0264<br>(0.0360)                                          | 0.0303***<br>(0.00864)                                        | 0.0337***<br>(0.00813)                                       | 0.0771***<br>(0.0244)                                                                          | 0.0682***<br>(0.0197)                                                      |
| Temperature Min squared                                | -0.00208***<br>(0.000761)                                  | 0.000689<br>(0.00107)                                        | -0.000788**<br>(0.000321)                                     | -0.00130***<br>(0.000247)                                    | -0.00150***<br>(0.000303)                                                                      | -0.00129***<br>(0.000249)                                                  |
| Temperature Max                                        | -0.134**<br>(0.0609)                                       | 0.0485<br>(0.0585)                                           | -0.0331**<br>(0.0166)                                         | -0.0521***<br>(0.0118)                                       | -0.125***<br>(0.0417)                                                                          | -0.111***<br>(0.0343)                                                      |
| Temperature Max squared                                | 0.00187***<br>(0.000520)                                   | -0.000838<br>(0.00129)                                       | 0.000548<br>(0.000547)                                        | 0.00120***<br>(0.000286)                                     | 0.00265***<br>(0.000745)                                                                       | 0.00217***<br>(0.000498)                                                   |
| Precipitation level                                    | -0.00184<br>(0.00208)                                      | -0.000181<br>(0.000719)                                      | 0.000557<br>(0.000362)                                        | 0.00114***<br>(0.000271)                                     | 0.000302<br>(0.000478)                                                                         | -0.000194<br>(0.000727)                                                    |

|                                                   |                          |                           |                           |                            |                           |                           |
|---------------------------------------------------|--------------------------|---------------------------|---------------------------|----------------------------|---------------------------|---------------------------|
| Precipitation level squared                       | 2.36e-06<br>(2.92e-06)   | -1.59e-07<br>(9.29e-07)   | -1.07e-06**<br>(5.35e-07) | -1.72e-06***<br>(4.52e-07) | -4.89e-07<br>(7.72e-07)   | 2.50e-07<br>(1.14e-06)    |
| Wind speed                                        | 0.173<br>(0.218)         | -0.160***<br>(0.0482)     | -0.0433<br>(0.0574)       | -0.0289<br>(0.0637)        | -0.0764<br>(0.0473)       | -0.147***<br>(0.0450)     |
| Wind speed squared                                | -0.0255<br>(0.0354)      | 0.0311***<br>(0.00854)    | 0.0177**<br>(0.00699)     | 0.0100<br>(0.00931)        | 0.0135*<br>(0.00813)      | 0.0206***<br>(0.00672)    |
| Inverse distance to the sea                       | -0.0113<br>(0.0216)      | -0.0261<br>(0.0173)       | -0.0208<br>(0.0171)       | -0.0206<br>(0.0170)        | -0.0192<br>(0.0173)       | -0.0205<br>(0.0173)       |
| Distance to the closest road                      | -3.18e-06<br>(5.13e-06)  | 5.54e-07<br>(2.89e-06)    | 3.07e-06<br>(2.28e-06)    | 3.62e-06<br>(2.33e-06)     | -1.13e-06<br>(3.26e-06)   | -2.63e-06<br>(3.87e-06)   |
| Female                                            | 0.127***<br>(0.0219)     | 0.122***<br>(0.0138)      | 0.123***<br>(0.0136)      | 0.123***<br>(0.0138)       | 0.135***<br>(0.0149)      | 0.137***<br>(0.0152)      |
| Number of kids                                    | 0.0331***<br>(0.0112)    | 0.0424***<br>(0.00695)    | 0.0434***<br>(0.00705)    | 0.0436***<br>(0.00712)     | 0.0402***<br>(0.00697)    | 0.0406***<br>(0.00693)    |
| Age                                               | -0.0859***<br>(0.0193)   | -0.0601***<br>(0.00327)   | -0.0604***<br>(0.00327)   | -0.0596***<br>(0.00333)    | -0.0623***<br>(0.00352)   | -0.0614***<br>(0.00338)   |
| Age square                                        | 0.00105***<br>(0.000188) | 0.000796***<br>(3.57e-05) | 0.000789***<br>(3.61e-05) | 0.000789***<br>(3.65e-05)  | 0.000819***<br>(3.81e-05) | 0.000810***<br>(3.67e-05) |
| <b>MARITAL STATUS DUMMY</b> (single as reference) |                          |                           |                           |                            |                           |                           |
| Married/ Cohabitation                             | 0.498***<br>(0.0418)     | 0.460***<br>(0.0212)      | 0.469***<br>(0.0213)      | 0.467***<br>(0.0214)       | 0.461***<br>(0.0215)      | 0.451***<br>(0.0224)      |
| Separate/Divorce/Widowed                          | 0.0310<br>(0.0513)       | 0.0590**<br>(0.0295)      | 0.0652**<br>(0.0286)      | 0.0824***<br>(0.0287)      | 0.0726**<br>(0.0287)      | 0.0659**<br>(0.0289)      |
| <b>HEALTH DUMMIES</b> (Very bad as reference)     |                          |                           |                           |                            |                           |                           |
| Poor                                              | 1.588***<br>(0.165)      | 1.454***<br>(0.0949)      | 1.462***<br>(0.0953)      | 1.467***<br>(0.0969)       | 1.421***<br>(0.0954)      | 1.433***<br>(0.0951)      |
| Neither                                           | 2.361***<br>(0.174)      | 2.196***<br>(0.0919)      | 2.208***<br>(0.0921)      | 2.202***<br>(0.0931)       | 2.194***<br>(0.0920)      | 2.198***<br>(0.0923)      |
| Good                                              | 3.388***<br>(0.178)      | 3.219***<br>(0.0912)      | 3.227***<br>(0.0915)      | 3.219***<br>(0.0923)       | 3.206***<br>(0.0912)      | 3.213***<br>(0.0915)      |
| Very good                                         | 4.324***<br>(0.189)      | 4.159***<br>(0.0935)      | 4.153***<br>(0.0928)      | 4.137***<br>(0.0931)       | 4.119***<br>(0.0923)      | 4.119***<br>(0.0925)      |

**EDUCATION DUMMIES** (No education as reference)

|                               |                  |                      |                     |                     |                     |                      |
|-------------------------------|------------------|----------------------|---------------------|---------------------|---------------------|----------------------|
| Completed primary school      | 0.719<br>(0.497) | 0.0845<br>(0.124)    | 0.117<br>(0.122)    | 0.120<br>(0.123)    | 0.133<br>(0.122)    | 0.138<br>(0.123)     |
| Completed junior high school  | 0.361<br>(0.360) | -0.0830<br>(0.102)   | -0.0568<br>(0.101)  | -0.0519<br>(0.102)  | -0.0713<br>(0.102)  | -0.0667<br>(0.101)   |
| Completed high school         | 0.446<br>(0.413) | -0.0969<br>(0.0912)  | -0.0715<br>(0.0895) | -0.0494<br>(0.0912) | -0.0748<br>(0.0903) | -0.0694<br>(0.0902)  |
| Completed professional school | 0.326<br>(0.355) | -0.118<br>(0.0859)   | -0.123<br>(0.0853)  | -0.0788<br>(0.0872) | -0.117<br>(0.0859)  | -0.111<br>(0.0856)   |
| Completed junior college      | 0.415<br>(0.399) | -0.0796<br>(0.0870)  | -0.0820<br>(0.0862) | -0.0465<br>(0.0895) | -0.0997<br>(0.0875) | -0.103<br>(0.0876)   |
| Completed university          | 0.317<br>(0.358) | -0.0847<br>(0.0889)  | -0.121<br>(0.0857)  | -0.0938<br>(0.0877) | -0.140<br>(0.0873)  | -0.140<br>(0.0871)   |
| Completed master's education  | 0.349<br>(0.329) | -0.0527<br>(0.0850)  | -0.0459<br>(0.0842) | -0.0315<br>(0.0853) | -0.0505<br>(0.0850) | -0.0523<br>(0.0849)  |
| Completed doctorate education | 0.520<br>(0.411) | -0.00319<br>(0.0869) | 0.0101<br>(0.0862)  | 0.0131<br>(0.0870)  | 0.00361<br>(0.0870) | -0.00447<br>(0.0869) |
| Other                         | 0.852<br>(0.593) | 0.0875<br>(0.0970)   | 0.128<br>(0.100)    | 0.0987<br>(0.0978)  | 0.0587<br>(0.0979)  | 0.0680<br>(0.0972)   |

**OCCUPATION DUMMIES** (full time employee as reference)

|                                   |                      |                       |                        |                        |                        |                        |
|-----------------------------------|----------------------|-----------------------|------------------------|------------------------|------------------------|------------------------|
| Part-time employee                | -0.210**<br>(0.0929) | -0.104***<br>(0.0271) | -0.0887***<br>(0.0269) | -0.0931***<br>(0.0269) | -0.0898***<br>(0.0275) | -0.0898***<br>(0.0273) |
| Company owner                     | -0.242<br>(0.197)    | 0.00262<br>(0.0465)   | 0.00537<br>(0.0463)    | 0.0212<br>(0.0486)     | -0.00586<br>(0.0468)   | 0.0101<br>(0.0473)     |
| Government employee               | -0.0954<br>(0.111)   | 0.0318<br>(0.0370)    | 0.0601<br>(0.0392)     | 0.0615<br>(0.0397)     | -0.0130<br>(0.0458)    | -0.00562<br>(0.0432)   |
| Professional (Doctor / Lawyer...) | -0.214<br>(0.164)    | -0.0138<br>(0.0355)   | 0.00647<br>(0.0352)    | 0.0207<br>(0.0379)     | -0.0323<br>(0.0382)    | -0.0257<br>(0.0368)    |
| Self-employed                     | -0.305***<br>(0.102) | -0.187***<br>(0.0270) | -0.163***<br>(0.0272)  | -0.163***<br>(0.0274)  | -0.190***<br>(0.0279)  | -0.183***<br>(0.0269)  |
| Student                           | 0.00199<br>(0.0704)  | -0.0407<br>(0.0373)   | -0.0140<br>(0.0378)    | -0.0235<br>(0.0369)    | -0.0604<br>(0.0408)    | -0.0648<br>(0.0417)    |
| Housewife / husband               | -0.132<br>(0.0877)   | -0.0294<br>(0.0274)   | -0.0357<br>(0.0278)    | -0.0204<br>(0.0276)    | -0.0293<br>(0.0278)    | -0.0231<br>(0.0275)    |
| Unemployed                        | -0.571***<br>(0.132) | -0.426***<br>(0.0336) | -0.411***<br>(0.0314)  | -0.403***<br>(0.0313)  | -0.404***<br>(0.0315)  | -0.409***<br>(0.0316)  |
| Other                             | -0.228               | -0.0760**             | -0.0520                | -0.0398                | -0.0489                | -0.0481                |

|                          |                     |                     |                     |                     |                     |                     |
|--------------------------|---------------------|---------------------|---------------------|---------------------|---------------------|---------------------|
|                          | (0.140)             | (0.0353)            | (0.0318)            | (0.0324)            | (0.0323)            | (0.0322)            |
| Constant                 | 2.724***<br>(0.342) | 2.588***<br>(0.217) | 2.884***<br>(0.220) | 3.202***<br>(0.338) | 3.025***<br>(0.266) | 3.063***<br>(0.280) |
| Observations             | 89,266              | 89,266              | 89,266              | 89,266              | 89,266              | 89,266              |
| R-squared                | 0.262               | 0.259               | 0.268               | 0.255               | 0.239               | 0.248               |
| Weather variables        | Yes                 | Yes                 | Yes                 | Yes                 | Yes                 | Yes                 |
| Country dummy            | Yes                 | Yes                 | Yes                 | Yes                 | Yes                 | Yes                 |
| Personal characteristics | Yes                 | Yes                 | Yes                 | Yes                 | Yes                 | Yes                 |

The significance levels relative to the test of the coefficients being different from zero, using a two-tailed hypothesis, are displayed as \*\*\*  $p < 0.01$ , \*\*  $p < 0.05$ , \*  $p < 0.1$ ; and the significance levels relative to the tests using one-tailed hypothesis testing, are displayed as  $^{+++}p < 0.01$ ,  $^{++}p < 0.05$ ,  $^{+}p < 0.1$ . Definitions of countries taxonomy are discussed in the *Section 4* and each country class is displayed in *Table 2-3*

*Appendix 2 - Section B: The impact of air pollution on affective well-being:*

We ensure the generalization of the inferences irrespective to the definition of the well-being. Previous studies show that impact of air pollution on well-being may differ when well-being is either defined as judgmental or as hedonic assessment of life (X. Zhang et al., 2017). Thus, I use a hedonic definition of well-being, so-called, affective well-being as a robustness check of our inference. Our index of affective well-being is the answer to question “Overall, how happy are you with your life?” (Diener, Kahneman, Tov, & Arora, 2010; X. Zhang et al., 2017). The answer to this question was recorded on a 5-point Likert-type scale with 0 representing the lowest level of happiness and 5 the highest. As displayed in Table 2-8, analyses with this definition of well-being does not change the inference made in the main manuscript.



Table 2-8 Impact of pollution on affective well-being

| VARIABLES                                              | The dependent variable is happiness (affective well-being) measured on 1-5 scale level. |                                                            |                                                                                |                                                                                    |                                                              |                                                                              |                                                                |
|--------------------------------------------------------|-----------------------------------------------------------------------------------------|------------------------------------------------------------|--------------------------------------------------------------------------------|------------------------------------------------------------------------------------|--------------------------------------------------------------|------------------------------------------------------------------------------|----------------------------------------------------------------|
|                                                        | (1)<br>Average<br>pollution<br>impact on<br>happiness                                   | (2)<br>Pollution<br>variation<br>interacted with<br>income | (3)<br>Pollution<br>changes<br>interacted<br>pollution level<br>per respondent | (4)<br>Pollution<br>changes<br>interacted with<br>national<br>development<br>index | (5)<br>Pollution<br>changes<br>interacted with<br>inequality | (6)<br>Pollution changes<br>interacted with<br>national average<br>pollution | (7)<br>Pollution changes<br>interacted with<br>national income |
| Change in NO <sub>2</sub>                              | -0.0456*<br>(0.0236)                                                                    | -0.108*<br>(0.0620)                                        | 0.867*<br>(0.478)                                                              | 0.486*<br>(0.273)                                                                  | 1.021*<br>(0.547)                                            | -0.796**<br>(0.389)                                                          | -0.690**<br>(0.335)                                            |
| Variation in NO <sub>2</sub> * log(income)             |                                                                                         | 1.425 <sup>†</sup><br>(0.990)                              |                                                                                |                                                                                    |                                                              |                                                                              |                                                                |
| Variation in NO <sub>2</sub> * Average NO <sub>2</sub> |                                                                                         |                                                            | -0.118 <sup>††</sup><br>(0.0619)                                               |                                                                                    |                                                              |                                                                              |                                                                |
| Change in NO <sub>2</sub> * developing country         |                                                                                         |                                                            |                                                                                | -0.538 <sup>††</sup><br>(0.275)                                                    |                                                              |                                                                              |                                                                |
| Change in NO <sub>2</sub> * unequal country            |                                                                                         |                                                            |                                                                                |                                                                                    | -1.053 <sup>††</sup><br>(0.540)                              |                                                                              |                                                                |
| Change in NO <sub>2</sub> * country pollution          |                                                                                         |                                                            |                                                                                |                                                                                    |                                                              | 0.809 <sup>††</sup><br>(0.418)                                               |                                                                |
| Change in NO <sub>2</sub> * productive country         |                                                                                         |                                                            |                                                                                |                                                                                    |                                                              |                                                                              | 0.748 <sup>††</sup><br>(0.387)                                 |
| log(income)                                            | 0.0691***<br>(0.00399)                                                                  | 0.155**<br>(0.0605)                                        | 0.0639***<br>(0.00482)                                                         | 0.0650***<br>(0.00454)                                                             | 0.0608***<br>(0.00588)                                       | 0.0796***<br>(0.00682)                                                       | 0.0800***<br>(0.00699)                                         |
| Average NO <sub>2</sub>                                |                                                                                         |                                                            | -0.00374<br>(0.00513)                                                          |                                                                                    |                                                              |                                                                              |                                                                |
| Temperature Min                                        | 0.00249<br>(0.00362)                                                                    | 0.0302<br>(0.0200)                                         | -0.0309*<br>(0.0180)                                                           | -0.000801<br>(0.00401)                                                             | 0.00107<br>(0.00373)                                         | 0.0247**<br>(0.0120)                                                         | 0.0198**<br>(0.00970)                                          |
| Temperature Min squared                                | -0.000137<br>(0.000107)                                                                 | -0.000632<br>(0.000392)                                    | 0.000866<br>(0.000535)                                                         | 7.22e-05<br>(0.000154)                                                             | -0.000205*<br>(0.000114)                                     | -0.000316**<br>(0.000142)                                                    | -0.000200*<br>(0.000114)                                       |
| Temperature Max                                        | -0.00744<br>(0.00528)                                                                   | -0.0511<br>(0.0314)                                        | 0.0479<br>(0.0293)                                                             | 0.00399<br>(0.00795)                                                               | -0.00636<br>(0.00536)                                        | -0.0461**<br>(0.0206)                                                        | -0.0386**<br>(0.0169)                                          |
| Temperature Max squared                                | 0.000213*<br>(0.000118)                                                                 | 0.000465*<br>(0.000266)                                    | -0.000998<br>(0.000644)                                                        | -0.000257<br>(0.000271)                                                            | 9.87e-05<br>(0.000134)                                       | 0.000887**<br>(0.000366)                                                     | 0.000624**<br>(0.000243)                                       |

|                              |                           |                           |                           |                           |                            |                           |                           |
|------------------------------|---------------------------|---------------------------|---------------------------|---------------------------|----------------------------|---------------------------|---------------------------|
| Precipitation level          | 0.000326***<br>(0.000116) | -0.00123<br>(0.00108)     | -0.000286<br>(0.000356)   | 7.41e-05<br>(0.000173)    | 0.000392***<br>(0.000123)  | -6.46e-05<br>(0.000232)   | -0.000335<br>(0.000360)   |
| Precipitation level squared  | -4.73e-07**<br>(1.89e-07) | 1.70e-06<br>(1.50e-06)    | 2.74e-07<br>(4.52e-07)    | -1.71e-07<br>(2.42e-07)   | -5.23e-07***<br>(1.94e-07) | 1.46e-07<br>(3.70e-07)    | 5.48e-07<br>(5.60e-07)    |
| Wind speed                   | 0.0113<br>(0.0196)        | 0.169<br>(0.113)          | -0.0118<br>(0.0235)       | 0.0510*<br>(0.0282)       | 0.0589*<br>(0.0314)        | 0.0330<br>(0.0229)        | -0.00521<br>(0.0218)      |
| Wind speed squared           | -0.00180<br>(0.00322)     | -0.0275<br>(0.0183)       | 0.00308<br>(0.00423)      | -0.00392<br>(0.00342)     | -0.00812*<br>(0.00461)     | -0.00623<br>(0.00398)     | -0.00234<br>(0.00328)     |
| Inverse distance to the sea  | -0.0120**<br>(0.00508)    | -0.00616<br>(0.00787)     | -0.0141***<br>(0.00522)   | -0.0113**<br>(0.00509)    | -0.0112**<br>(0.00506)     | -0.0105**<br>(0.00519)    | -0.0112**<br>(0.00515)    |
| Distance to the closest road | 1.33e-06<br>(9.56e-07)    | -1.84e-06<br>(2.53e-06)   | 4.13e-07<br>(1.21e-06)    | 1.56e-06<br>(9.65e-07)    | 1.85e-06*<br>(1.02e-06)    | -7.28e-07<br>(1.46e-06)   | -1.54e-06<br>(1.80e-06)   |
| Female                       | 0.0392***<br>(0.00630)    | 0.0403***<br>(0.0113)     | 0.0377***<br>(0.00649)    | 0.0380***<br>(0.00640)    | 0.0380***<br>(0.00647)     | 0.0445***<br>(0.00707)    | 0.0454***<br>(0.00725)    |
| Number of kids               | 0.0223***<br>(0.00325)    | 0.0180***<br>(0.00575)    | 0.0232***<br>(0.00333)    | 0.0237***<br>(0.00339)    | 0.0237***<br>(0.00341)     | 0.0219***<br>(0.00335)    | 0.0221***<br>(0.00333)    |
| Age                          | -0.0281***<br>(0.00154)   | -0.0419***<br>(0.00998)   | -0.0279***<br>(0.00156)   | -0.0280***<br>(0.00155)   | -0.0276***<br>(0.00159)    | -0.0291***<br>(0.00167)   | -0.0286***<br>(0.00160)   |
| Age square                   | 0.000332***<br>(1.69e-05) | 0.000466***<br>(9.71e-05) | 0.000330***<br>(1.71e-05) | 0.000326***<br>(1.73e-05) | 0.000326***<br>(1.76e-05)  | 0.000343***<br>(1.81e-05) | 0.000337***<br>(1.74e-05) |

#### HEALTH DUMMIES (Very bad as reference)

|           |                      |                      |                      |                      |                      |                      |                      |
|-----------|----------------------|----------------------|----------------------|----------------------|----------------------|----------------------|----------------------|
| Poor      | 0.605***<br>(0.0428) | 0.686***<br>(0.0839) | 0.613***<br>(0.0434) | 0.617***<br>(0.0436) | 0.620***<br>(0.0446) | 0.595***<br>(0.0434) | 0.601***<br>(0.0432) |
| Neither   | 0.930***<br>(0.0411) | 1.022***<br>(0.0887) | 0.933***<br>(0.0416) | 0.939***<br>(0.0418) | 0.936***<br>(0.0424) | 0.931***<br>(0.0414) | 0.933***<br>(0.0416) |
| Good      | 1.406***<br>(0.0408) | 1.504***<br>(0.0910) | 1.412***<br>(0.0413) | 1.416***<br>(0.0415) | 1.411***<br>(0.0419) | 1.405***<br>(0.0410) | 1.408***<br>(0.0412) |
| Very good | 1.801***<br>(0.0411) | 1.908***<br>(0.0967) | 1.818***<br>(0.0425) | 1.815***<br>(0.0421) | 1.807***<br>(0.0423) | 1.797***<br>(0.0414) | 1.797***<br>(0.0416) |

#### MARITAL STATUS DUMMY (single as reference)

|                          |                       |                      |                       |                       |                       |                       |                       |
|--------------------------|-----------------------|----------------------|-----------------------|-----------------------|-----------------------|-----------------------|-----------------------|
| Married/ Cohabitation    | 0.255***<br>(0.00965) | 0.274***<br>(0.0215) | 0.253***<br>(0.00986) | 0.258***<br>(0.00988) | 0.257***<br>(0.00994) | 0.253***<br>(0.00995) | 0.248***<br>(0.0105)  |
| Separate/Divorce/Widowed | 0.0610***<br>(0.0133) | 0.0371<br>(0.0263)   | 0.0526***<br>(0.0142) | 0.0557***<br>(0.0137) | 0.0651***<br>(0.0137) | 0.0597***<br>(0.0136) | 0.0561***<br>(0.0138) |

#### EDUCATION DUMMIES (No education as reference)

|                                                      |                        |                       |                        |                        |                        |                        |                        |
|------------------------------------------------------|------------------------|-----------------------|------------------------|------------------------|------------------------|------------------------|------------------------|
| Completed primary school                             | 0.124**<br>(0.0494)    | 0.313*<br>(0.173)     | 0.114**<br>(0.0507)    | 0.125**<br>(0.0503)    | 0.126**<br>(0.0508)    | 0.136***<br>(0.0509)   | 0.139***<br>(0.0508)   |
| Completed junior high school                         | 0.152***<br>(0.0534)   | 0.483*<br>(0.257)     | 0.138**<br>(0.0552)    | 0.156***<br>(0.0541)   | 0.157***<br>(0.0551)   | 0.164***<br>(0.0546)   | 0.167***<br>(0.0552)   |
| Completed high school                                | 0.0981**<br>(0.0445)   | 0.332*<br>(0.186)     | 0.0901**<br>(0.0457)   | 0.104**<br>(0.0452)    | 0.107**<br>(0.0459)    | 0.0965**<br>(0.0455)   | 0.0990**<br>(0.0453)   |
| Completed professional school                        | 0.0684*<br>(0.0395)    | 0.351<br>(0.214)      | 0.0566<br>(0.0408)     | 0.0699*<br>(0.0400)    | 0.0820**<br>(0.0410)   | 0.0681*<br>(0.0402)    | 0.0710*<br>(0.0402)    |
| Completed junior college                             | 0.0359<br>(0.0372)     | 0.273<br>(0.184)      | 0.0318<br>(0.0382)     | 0.0288<br>(0.0379)     | 0.0526<br>(0.0391)     | 0.0320<br>(0.0380)     | 0.0351<br>(0.0378)     |
| Completed university                                 | 0.0463<br>(0.0378)     | 0.319<br>(0.207)      | 0.0499<br>(0.0388)     | 0.0481<br>(0.0384)     | 0.0675*<br>(0.0403)    | 0.0385<br>(0.0387)     | 0.0367<br>(0.0388)     |
| Completed master's education                         | 0.0470<br>(0.0376)     | 0.286<br>(0.186)      | 0.0666*<br>(0.0399)    | 0.0473<br>(0.0381)     | 0.0620<br>(0.0392)     | 0.0370<br>(0.0386)     | 0.0368<br>(0.0385)     |
| Completed doctorate education                        | 0.0502<br>(0.0369)     | 0.266<br>(0.170)      | 0.0475<br>(0.0378)     | 0.0509<br>(0.0374)     | 0.0587<br>(0.0380)     | 0.0484<br>(0.0376)     | 0.0474<br>(0.0375)     |
|                                                      | 0.0506                 | 0.332                 | 0.0476                 | 0.0548                 | 0.0565                 | 0.0513                 | 0.0469                 |
| OCCUPATION DUMMIES (full time employee as reference) |                        |                       |                        |                        |                        |                        |                        |
| Part-time employee                                   | -0.0310**<br>(0.0123)  | -0.0935*<br>(0.0479)  | -0.0355***<br>(0.0127) | -0.0273**<br>(0.0126)  | -0.0298**<br>(0.0126)  | -0.0280**<br>(0.0128)  | -0.0279**<br>(0.0127)  |
| Company owner                                        | 0.00489<br>(0.0211)    | -0.125<br>(0.102)     | 0.00915<br>(0.0216)    | 0.00983<br>(0.0214)    | 0.0184<br>(0.0227)     | 0.00372<br>(0.0218)    | 0.0124<br>(0.0221)     |
| Government employee                                  | 0.00180<br>(0.0164)    | -0.0682<br>(0.0577)   | 0.00219<br>(0.0170)    | 0.0165<br>(0.0181)     | 0.0172<br>(0.0184)     | -0.0233<br>(0.0215)    | -0.0193<br>(0.0202)    |
| Professional (Doctor / Lawyer...)                    | -0.00835<br>(0.0156)   | -0.121<br>(0.0846)    | -0.0119<br>(0.0163)    | -0.00149<br>(0.0161)   | 0.00627<br>(0.0177)    | -0.0226<br>(0.0179)    | -0.0190<br>(0.0171)    |
| Self-employed                                        | -0.0421***<br>(0.0122) | -0.113**<br>(0.0530)  | -0.0479***<br>(0.0128) | -0.0353***<br>(0.0128) | -0.0356***<br>(0.0129) | -0.0500***<br>(0.0132) | -0.0463***<br>(0.0127) |
| Student                                              | 0.0224<br>(0.0165)     | 0.0408<br>(0.0363)    | 0.0178<br>(0.0171)     | 0.0321*<br>(0.0174)    | 0.0269<br>(0.0169)     | 0.00681<br>(0.0189)    | 0.00444<br>(0.0194)    |
| Housewife / husband                                  | 0.0494***<br>(0.0123)  | -0.00859<br>(0.0454)  | 0.0479***<br>(0.0126)  | 0.0438***<br>(0.0128)  | 0.0521***<br>(0.0127)  | 0.0473***<br>(0.0128)  | 0.0507***<br>(0.0127)  |
| Unemployed                                           | -0.151***<br>(0.0144)  | -0.242***<br>(0.0679) | -0.163***<br>(0.0159)  | -0.155***<br>(0.0147)  | -0.150***<br>(0.0147)  | -0.151***<br>(0.0148)  | -0.154***<br>(0.0148)  |
| Other                                                | -0.00358<br>(0.0148)   | -0.102<br>(0.0723)    | -0.0183<br>(0.0169)    | -0.00595<br>(0.0150)   | 0.000688<br>(0.0154)   | -0.00423<br>(0.0152)   | -0.00382<br>(0.0152)   |
| Constant                                             | 2.300***<br>(0.0891)   | 2.302***<br>(0.176)   | 2.207***<br>(0.101)    | 2.390***<br>(0.101)    | 2.563***<br>(0.163)    | 2.467***<br>(0.125)    | 2.487***<br>(0.132)    |

|                          |        |        |        |        |        |        |        |
|--------------------------|--------|--------|--------|--------|--------|--------|--------|
| Observations             | 89,266 | 89,266 | 89,266 | 89,266 | 89,266 | 89,266 | 89,266 |
| R-squared                | 0.243  | 0.223  | 0.208  | 0.225  | 0.205  | 0.195  | 0.203  |
| Weather variables        | Yes    | Yes    | Yes    | Yes    | Yes    | Yes    | Yes    |
| Country dummy            | Yes    | Yes    | Yes    | Yes    | Yes    | Yes    | Yes    |
| Personal characteristics | Yes    | Yes    | Yes    | Yes    | Yes    | Yes    | Yes    |

The significance levels relative to the test of the coefficients being different from zero, using a two-tailed hypothesis, are displayed as \*\*\*  $p < 0.01$ , \*\*  $p < 0.05$ , \*  $p < 0.1$ ; and the significance levels relative to the tests using one-tailed hypothesis testing, are displayed as  $^{++}p < 0.01$ ,  $^{+}p < 0.05$ ,  $^{+}p < 0.1$ . Definitions of countries taxonomy are discussed in the *Section 4* and each country class is displayed in *Table 2-3*

## *Appendix 2 - Section C: Measuring well-being as a single item*

Despite certain studies arguing for the need of multi-item surveys to measure wellbeing (mostly concept of affective well-being), I believe that a single item description of SWB is more appropriate in the current study. The combination of multiple indicators to indicate well-being is classified in the *List Theories* of the measure of wellbeing (Haybron, 2008). According to these theories, well-being consist of several basic conditions, such as health and good relationships (Oishi, 2010). However, the answer to the question of “All things considered, how satisfied are you with your life these days?” is part of the *Authentic Happiness Theories*. It represents a global, evaluative judgement that respondents make about their own lives (Kahneman, David, Claude, Alan, & Amy, 2010). Despite the fact that none of these approaches prevail on the other (Haybron, 2008), using the *Authentic Happiness Theories* is preferable in this study because of the multinational nature of our survey. The *Happiness Theories* do not assume a knowledge of the formula of indices that dedicates the construction of well-being across *nations* and *culture*. I rather put the emphasis on individuals’ subjective experience and sovereignty in defining their well-being. The bias of *List Theories* in this instance is illustrated by Shigehiro Oishi (2010) when he refers to the cultural differences of item functioning of surveys. He highlights, for instance, that although researchers may be willing to associate an item such as “smiling” as a positive factor of wellbeing, cultures such as that of the Ifaluk condemns “smiling” because it is thought to lead to the neglect of one’s duties (Oishi, 2010).

I show the relevance of this argument by estimating the Cronbach’s alpha of survey items that are often used by other studies to describe an aggregated well-being (Diener et al., 2010). A Cronbach’s alpha statistic ( $\alpha$ ) is a way of evaluating the reliability of an additive scale created over several items. An  $\alpha$  of an aggregation of survey items is considered acceptable when it is higher than 0.7. It is considered poor when  $0.6 > \alpha > 0.5$  and unacceptable when  $0.5 > \alpha$ . The association of items describing the occurrences of pleasure, enjoyment, smile, anger,

sadness, and depression (the three later terms included in reverse) generate a Cronbach's alpha statistic of 0.2.

## **Chapter 3 An evaluation of the sub-national difference in inclusive wealth sustainability in Africa**

### **3.1. Introduction**

GDP per capita in African countries grew on average by 1.6% since 2000 (World Bank estimate). It enabled a significant reduction of the poverty rate and improvement in life expectancy (Roser, Ortiz-Ospina and Ritchie, 2013) and educational attainment (Lutz *et al.*, 2021). To some extent, these advances highlight efforts made by private sectors, governments, and NGOs in improving macroeconomic indices such as GDP or HDI. However, these efforts have raised concerns over the sustainability of African development because (1) most African countries are resource-intensive (UNU-IHDP & UNEP, Report and Report, 2012; UNU-IHDP & UNEP, 2014; Lange, Wodon and Carey, 2018) (e.g. development based on export revenues from fossil fuels and precious minerals) and (2) African economic growth is linked to serious environmental degradations illustrated by losses in forest covers (Brandt *et al.*, 2017) and biodiversity (Scholes and Biggs, 2005), as well as increases in air (Steckel *et al.*, 2020) and water (Yerema *et al.*, 2020; Vörösmarty *et al.*, 2021) pollutions.

To investigate this matter, international bodies created macroeconomic indices that incorporate the issue of sustainability by accounting for the value of capital stocks over time (Holden, Linnerud and Banister, 2014; Lange, Wodon and Carey, 2018; Managi, Kumar and

Pushpam, 2018; Bateman and Mace, 2020). Among these indices, inclusive wealth (IW), developed by the United Nations Environmental Program (UNEP), appears as the most holistic (UNU-IHDP & UNEP, Report and Report, 2012; UNU-IHDP & UNEP, 2014; Polasky *et al.*, 2015). IW is an aggregated monetary indicator of all capital assets that (1) can be regarded as an "economy's productive base" and/or that (2) directly participate in the betterment of human wellbeing (Polasky *et al.*, 2015). In practice, these capital assets are classified as natural (e.g. agricultural production and ecosystem services), human (e.g. education and income), or manufactured (e.g. infrastructure and savings). The IW classifies countries' development as sustainable when this development is followed by a non-declining IW over time, otherwise it is considered unsustainable.

Through the lens of the inclusive wealth framework, differences of sustainability across countries were estimated and they reveal that only 11 of a sample of 35 African countries followed a sustainable path between 1992 and 2004 (Lange, Wodon and Carey, 2018; Managi, Kumar and Pushpam, 2018). With this assessment, UNEP and the World Bank recommended a shift toward the renewable energy sector, and improvement in land productivity in African countries to create sustainable economies.

However, these assessments exist only at the national level in Africa. Thus, they neglect unequitable sustainable development paths engendered by inequalities of political powers and influences within countries. In this regard, several pieces of evidence stemming from remote



sensing data and machine learning techniques suggest unequal sustainability patterns between areas around capital cities and rural areas.

In rural areas, a study shows the presence of extractive activities whose benefits are transferred towards urban areas in the case of water resources (Showers, 2002). Taking South African rural areas as examples, other studies provide evidence of the absence of efficient policies to protect the environment and the health of rural populations (Chakravarty, 2011; Yerema *et al.*, 2020). Across countries in the continent, rural areas attract relatively few investments owing to the low productivity of their economic activities (Dorosh and Thurlow, 2014; Barrett *et al.*, 2017). A research on the economic growth in African rural areas corroborate this claim by showing that national institutions and subnational economic development suffer in locations remote from national capital cities (Growth *et al.*, 2014).

On the other hand, mismanagement of wealth in main urban areas (generally capital cities) is linked to urban pollutions (Amegah and Agyei-Mensah, 2017), and unsustainable conversion of croplands to built-up areas (Bren d'Amour *et al.*, 2017). Nevertheless, these urban centers attract the largest shares of investments on the continent because they are built around profitable industries (Dorosh and Thurlow, 2014; Padraig and Francis, 2016), and national political powers (Davis and Vernon Henderson, 2003). Overall, these findings show that issues of heterogeneous within-country sustainability performances in Africa remain to be explored.

Thus, I aim to explore differences in sustainability in 48 African countries at the provincial level (GADM dataset first administrative level). I quantify IW in 2012, 2018 and the resulting IW growth between these years across 716 provinces. Since no African government provides data enabling such work, these estimates are built using innovative disaggregated data under the IW framework developed by UNEP set out to value African natural, human, and manufactured capital stocks. This dataset include satellite-monitored maps of population (Tatem, 2017), and estimates of education, income, and life expectancy (Smits and Permanyer, 2019) from the Subnational Human Development Index (SHDI) to construct provincial human capital. For natural capital, I use satellite data of MODIS land cover (Friedl *et al.*, 2002; Sulla-Menashe and Friedl, 2018), mineral PRIO maps of mineral deposit (Taylor *et al.*, 2005) and value these assets with the willingness to pay for ecosystems (de Groot *et al.*, 2012) and international prices of minerals and fossil fuels (U.S. Geological Survey, 2018) respectively. Finally, I constructed estimates of provincial manufactured capital stocks by disseminating national manufactured capital provided by the IMF (IMF, 2019) to each province relative to its amount of satellite-monitored night light data (Elvidge *et al.*, 2021) (see Methods for further details and the robustness of this procedure).

Through estimating IW, this study's objectives are to assess the difference in per capita wealth growth between provinces where the national capital city is located, referred to as key provinces, and other provinces, referred to as secondary provinces, and the reason behind this

difference (if existing). This objective is based on evidence of higher economic growth and more efficient policies in African urban areas (Onyeji, Bazilian and Nussbaumer, 2012; Graetz *et al.*, 2018; Hilson *et al.*, 2018) and in capital cities (Baker and Therborn, 2012; Michalopoulos and Papaioannou, 2013). This evidence may suggest a divergence in sustainability performances and equity issues geographically. To reach this objective, I first describe the geographical distribution of wealth and its nature across provinces in 2018. Then, I perform regression analyses that (1) estimate the difference in IW growth between key and secondary provinces and that (2) determine the forces accounting for this difference in IW growth. To determine these forces, the analysis considers explanation based on the components of wealth including provincial population growth between 2012-2018; and the total population, as well as the amount of manufactured, human, and natural capitals in 2012. I use ordinary least-squares (OLS) to understand the relationship between assets and sustainability performances while controlling for countries' fixed effects to provide an interpretation relative to each countries' average (see Methods).

## **3.2. Methods**

### *3.2.1 Estimation of Inclusive wealth*

I use the variation of the inclusive wealth (IW) per capita from 2012 to 2018 to investigate the sustainability of wealth in African subnational provinces. For a general audience,

variations in IW provide an aggregate view of the outcome of wealth management in societies (Arrow *et al.*, 2012; Polasky *et al.*, 2015). In an illustrative model, individuals in a region may choose to exploit local oil deposits (i.e. NC) to either invest in local infrastructure (i.e. PC) and level of education (i.e. HC). Individuals may also choose to consume the revenues of the extraction of oil to purchase provisions or goods. In this context, a negative variation in IW indicates an unsustainable development path as the oil reserves used to generate the PC or HC are worth more than the new forms of capital. If the variation in IW were to be positive, however, it would be an instance of weak sustainable development because, despite the overall increase in IW, one source of capital is decreasing in the process. Finally, this example cannot exhibit a case of a strong sustainable development, wherein all three types of capital are increasing over time.

For the estimation, I assume that assets that necessary to generate economic activities are either natural, human, manufactured, implying that IW can be expressed as follows:

$$IW_{it} = P_{NC} \times NC_{it} + P_{HC} \times HC_{it} + P_{pc} \times PC_{it} \quad (3.1)$$

where  $IW_{it}$ ,  $NC_{it}$ ,  $HC_{it}$  and,  $PC_{it}$  stand for the IW, natural capital, human capital, and manufactured capital respectively, for the province  $i$  at the year  $t$ . The terms  $P_{NC}$ ,  $P_{HC}$  and  $P_{PC}$  are the prices of the natural, human, and manufactured capitals respectively and serve as valuations of these assets. Since no conventional market describes the prices or monetary

values of human and natural capitals, I use the concept of shadow price to estimate their worth (Arrow *et al.*, 2012) (*further details below*). The shadow prices represent the contributions of these capitals in the total IW both through the goods and services they help produce, and through direct enjoyment that these capital stocks offer.

For simplicity, I assume that prices (1) are uniform across areas and (2) constant over time (Arrow *et al.*, 2012). Prices of all assets are set to their international values in 2011 (UNU-IHDP & UNEP, Report and Report, 2012; UNU-IHDP & UNEP, 2014; Managi, Kumar and Pushpam, 2018). For sustainability investigations in the short-term (2012 to 2018), focusing on changes in assets rather than those in prices provide a more meaningful indication of the development trajectory of regions (Arrow *et al.*, 2012). Thus, changes in IW are caused only by changes in the capital stocks while the effects of price fluctuations are disregarded. Change in IW can be expressed as follows:

$$\Delta IW_i = \text{Inclusive Investment}_i = P_{NC} \times \Delta NC_i + P_{HC} \times \Delta HC_i + P_{PC} \times \Delta PC_i$$

(3.2)

with the symbol  $\Delta$  representing variations in time.

I describe the estimation of the variations of capital stocks in the subsequent sections. In these estimations, we owe to note that the definition of province boundaries are derived from the first level of GADM administrative areas dataset (<https://gadm.org/index.html>). In our

calculations, dataset collected as raster data are aggregated within these province boundaries using ArcGIS software. For datasets collected as shapefiles and whose province boundaries differ from that of GADM dataset, I weighted their values proportionally to the surface of the coverage of GADM shapefiles.

### 3.2.1.1 Natural capital

Natural capital refers to economic inputs that are naturally occurring within a province.

The estimation of the natural capital can be expressed as follows:

$$NCW_{it} = NCN_{it} + RCN_{it} \quad (3.3)$$

where  $NCW_{it}$  is the natural capital wealth ( $P_{NC} \times NC_{it}$ ) of the province  $i$  at year  $t$ . It is composed of the non-renewable natural capital,  $NCN_{it}$  and renewable natural capital,  $RCN_{it}$ .

#### ● Non-renewable natural capital

The underground wealth, or non-renewable natural capital, includes the wealth related to precious minerals and metals, as well as fossil fuel stocks. The underground wealth is estimated as follows:

$$NCN_{it} = \sum_{m=1}^k P_m \times rent \times reserves_{itm} \quad (3.4)$$

with  $reserves_{itm}$  representing the proven natural reserves,  $P_m$ , the price of the natural asset  $m$  in the province  $i$  at year  $t$ . The term  $rent$  is the rental rate of the mineral industry and represents the proportion of the price of the mineral that goes to profit. Data of rental rates are collected from World Bank reports on total wealth (*Table 3-7*) (Lange, Wodon and Carey, 2018). The shadow price of the underground wealth is the multiplication of the rental rate and the price of the natural asset price.

Since no dataset describes the sub-national reserves of natural endowments in Africa, I create a proxy of provincial reserves,  $reserves_{itm}$ . For this purpose, I allocate national mineral reserves of coal and precious mineral reserves across provinces proportionally to the number of deposits of these assets in each province. Data on the location and the nature of mineral deposits come from USGS (<https://mrdata.usgs.gov/major-deposits/>). For the estimation of provincial oil and gas reserves, I allocate national reserves based on the surface of coverage of oil drilling maps per province. Off-shore fuel deposits were allocated to the province whose coast is the closet to the mineral deposit. These procedures were carried out with the on-shore and off-shore oil drilling maps provided by Peace Research Institute Oslo (PRIO) (Lujala, Ketil Rod and Thieme, 2007).

Since it is prohibitively difficult to value all fossil and mineral reserves in Africa, I concentrate on a sample of resources that are either essential to the world economy or of which African countries are the main producers: bauxite, chromium, coal, cobalt, copper, diamond,

gas, gold, lead, manganese, oil, silver, tin, and zinc. The data of national proven resources were collected from various sources and checked several times to ensure their consistency (*see* Table 3-4 to 3-6 and file of resources). Reserve estimates that were not consistent across sources were considered to be missing values. Prices of natural resources and the data sources are displayed in Table 3-7 in the Appendix (U.S. Geological Survey, 2018).

- *Renewable natural capital*

Renewable natural capital represents natural assets that regenerate themselves. I estimate it as follows:

$$RNC_{it} = \sum_{n=1}^l P_n \times area_{itn} + \sum_{a=1}^{\infty} \frac{Agri_{it} \times rent_{agri}}{(1+r)^a} \quad (3.5)$$

With the renewable natural capital,  $RNC_{it}$ , estimated as the addition of the ecosystem wealth ( $\sum_{n=1}^l P_n \times area_{itn}$ ) and the agricultural wealth ( $\sum_{a=1}^{\infty} \frac{Agri_{it} \times rent_{agri}}{(1+r)^a}$ ).

The ecosystem wealth is the sum of services that each type of ecosystem in a province provides to a population. I consider six types of ecosystems: inland wetlands, tropical forest, grasslands, woodlands, coastal system and marine (de Groot *et al.*, 2012). The worth of these ecosystems is determined by their contribution to human well-being via aspects such as the provision of foods and raw materials; the preventions of erosion and climate adverse effects; or the endowment of aesthetic and recreational settings (de Groot *et al.*, 2012; Bateman and



Mace, 2020). In this regard, I use the willingness to pay (WTP) for a unit area of an ecosystem  $n$  as an index of its shadow price ( $P_n$ ). The data on WTP for ecosystems stems from the largest meta-analysis conducted in this field to our knowledge (de Groot *et al.*, 2012).

Then, the total monetary value of the ecosystem  $n$  in the province  $i$  at the year  $t$  is the product of its shadow price  $P_n$  and the area of that ecosystem ( $area_{int}$ ) in the province  $i$ . Our estimates of ecosystem areas are based on satellite images of land cover provided by the MODIS imaging sensor (<https://modis.gsfc.nasa.gov>). The MODIS land cover dataset, generated by remote sensing techniques, describes distinct land cover types at a resolution of 500m<sup>2</sup>. With this data, I made correspondence between land cover types and ecosystems as described in Table 3-3 and, thereby, estimate  $area_{itn}$ .

Further, I follow the procedure of the IWI report to estimate agricultural wealth (UNU-IHDP & UNEP, 2014). In Equation (3.5), the term  $Agri_{it}$  represents the value of agricultural productions in the province  $i$  at year  $t$ ; the term  $r$  is the discount rate, set to 8.5%; and  $rent_{agri}$  stands for rental rate from agricultural production (the part of the production that goes to profit), set to 0.23 (Lange, Wodon and Carey, 2018). Equation (5) implies that the agricultural wealth is the discounted sum over an infinite time horizon of profits from agricultural production. To proxy  $Agri_{it}$ , I first collect national agricultural production from FAO, then allocate agricultural production across each province  $i$  proportionally to its total agricultural land covers as provided by the MODIS dataset.

### 3.2.1.2. Human capital

Human capital refers to the worth of populations in producing economic activities. Similar to the IWI report, estimations of the human capital is modeled as follows (Managi, Kumar and Pushpam, 2018):

$$HCW_{it} = e^{(A_{it} \cdot \rho)} \cdot P_{it15+} \cdot \int_0^{T_i} \bar{r}_i \cdot e^{-\delta \cdot t} dt \quad (3.6)$$

where  $HCW_{it}$  stands for the human capital wealth ( $P_{HC} \times HC_{it}$ ) of the province  $i$  at year  $t$ . The term  $e^{(A_{it} \cdot \rho)}$  represents the benefit of education to the human capital. The component  $A_{it}$  of  $e^{(A_{it} \cdot \rho)}$  is the average school attainment, defined as the average years of formal education, and  $\rho$  represents the average return of education on future wages, set to 8.5% (UNU-IHDP & UNEP, 2014; Managi, Kumar and Pushpam, 2018). The term  $P_{it15+}$  stands for the total working population and is assumed to be the population older than 15 years old. I assume the individuals younger than 15 years are outside the labor force because they are either at school or substantially less productive than others. Hence,  $P_{it15+}$  may be grossly referred as the human capital stock and  $e^{(A_{it} \cdot \rho)}$  is the value of education in this population.

The term  $\int_0^{T_i} \bar{r}_i \cdot e^{-\delta \cdot t}$  represents the shadow price of the human capital. This shadow price is the discounted sum of income an individual may expect to earn in his lifetime ( $T_i$ ). It is made up of the average income in the province  $i$ ,  $\bar{r}_i$ , multiplied by the discount rate of future earnings,  $e^{-\delta \cdot t}$  (with  $\delta$  set to an interest rate of 8.5%) (Managi, Kumar and Pushpam, 2018).

The effective working lifetime,  $T_i$ , is the average life expectancy in the province  $i$  deducted to 15 years. Including the life expectancy in the human capital computation enables us to somehow account for the benefits of health in each region.

I collected the sub-national population data and age structure from Worldpop unconstrained dataset (Tatem, 2017) while the data of sub-national education attainment, income and life expectancy come from the Subnational Human Development Index 4.0 (Smits and Permanyer, 2019).

### *3.2.1.3 The manufactured capital*

Manufactured capital encompasses manufactured assets such as machinery, equipment, and physical structures and savings within a province. Few countries monitor such dataset (African Development Bank Group, 2016) inciting, previous works to use alternative indices to value it (Mumford, 2012; Zhang, Nozawa and Managi, 2021). In the USA, researchers complement (1) available data on infrastructure values with (2) data of stock ownership and (3) surveys on equities to disintegrate national capital stock at state and county levels (Mumford, 2012; Ghadimi, Harris and Warner, 2015). In Japan and China, satellite-monitored night light maps were used to decompose national capital stock at the 1km grid level, under the assumption that the most valuable infrastructure within countries are the most lit at night (e.g. infrastructures in urban centers).

Hence, building on these studies, I disaggregate national manufactured capital data to the provincial level based on the distribution of satellite-monitored night light data. I use night light as index of wealth because it is consistently available in Africa at the subnational level. Our assumption with this method is based on a significant body of works suggesting that these light emissions strongly correlate with wealth nationally and subnationally (Elvidge *et al.*, 2012; Henderson, Storeygard and Weil, 2012; Liang *et al.*, 2014; Pokhriyal and Jacques, 2017; Weidmann and Schutte, 2017).

This method is relevant internationally with the data since the Pearson's correlation between capital stock and total of national luminosity is 0.859 ( $p < 0.01$ ) across countries. Within countries, I can rely on findings from several studies (Jean *et al.*, 2016; Njuguna and McSharry, 2017; Yeh *et al.*, 2020) such as Neeti Pokhriyal and Damien Jacques (2017). Neeti Pokhriyal and Damien Jacques (2017) show that night light data has by far the highest ability to predict asset (manufactured capital) ownership of respondents of a survey at the communal level in Senegal among Big data including 120 variables stemming from mobile phone records, satellite imageries, prices and goods sold in local markets, distance to hospital, schools and more (Pokhriyal and Jacques, 2017).

Thus, I first collected the valuations of national manufactured capital stocks from the (IMF, 2019). IMF data accounts national gross fixed capital formation (GFCF) and is estimated via a perpetual inventory method with collection of available data on investment since 1860.

Then, I use night light emissions (Elvidge *et al.*, 2021) as indices of the distribution of sub-national capital stocks in 2018. With these data, I estimated the provincial capital stock as follows:

$$PCW_{it} = \frac{PCW_{Kt}}{light_{Kt}} \times light_{it} \quad (3.7)$$

where  $PCW_{it}$  and  $PCW_{Kt}$  represent the manufactured capital wealth ( $P_{pc} \times PC_{it}$ ) in the province  $i$  and the country  $K$  (where province  $i$  is located) at year  $t$ . While  $light_{it}$  and  $light_{Kt}$  are the sum of night light radiance units in the province  $i$  and the country  $K$  respectively. In equation (7), I assume that the proportion of night light emission in province  $i$  over the national light emissions is equivalent to the proportion of provincial manufactured capital over national manufactured capital.

Then, I use the growth of night light between 2012 and 2018 per province to estimate their respective manufactured capital growth. It consisted in allocating national growth of manufactured capital to each province  $i$  proportionally to its growth in night light during this period. This approach ensures that provinces whose total light grow are allocated an increase in capital.

These estimations use the data of yearly median value of night light developed by (Elvidge *et al.*, 2021).

### 3.2.2 *Difference in IW growth between key and secondary provinces and factors underlying it*

To estimate the difference in IW growth between key provinces and secondary provinces, I regress the following model:

$$\Delta IWpc_i = \beta_0 + \beta_1 key_1 + K_i + \varepsilon_i \quad (3.8)$$

with  $\Delta IWpc_i$  representing the change in IW per capita from 2012 to 2018. The term  $K_i$  is a country factor variable which equals 1 if province  $i$  is within the country K and 0 otherwise. This term captures national differences in wealth across provinces. Therefore, it enables interpreting the coefficients of other variables as average variations relative to the national averages. The term  $\varepsilon_i$  is the idiosyncratic error term of the province  $i$ .

The term  $key_1$  is the variable of interest. It is factor variable indicating that province  $i$  host a national capital city. I include Abidjan, Bujumbura, Cotonou, Dar es Salaam and Lagos in the group of capital cities because they have an important governmental presence (for instance the residence of the president) in their countries in 2012. The coefficient  $\beta_1$  is the estimate of the difference in IW growth between provinces where there is capital cities and other provinces within the same country.

Then, I test the hypothesis that the difference in IW growth between key provinces and other provinces reflects variables that are omitted from Equation (3.8). Hence, I carried out the

regressions including potential variables  $X_i$  that may explain the difference in IW growth between provinces:

$$\Delta IWpc_i = \beta_0 + \beta_1 key_i + \beta_2 X_i + K_i + \varepsilon_i \quad (3.9)$$

After performing separate iterations for each province's variable  $X_i$ , I assess which variable attenuates the estimated difference in IW growth between key and secondary provinces  $\beta_1$ . For this purpose, the province's characteristics I focus on are the population growth and level of per capita manufactured capital, human capital, natural capital, average income, average years of school attainment, ecosystem wealth, underground wealth, and the total population in 2012. Among these variables, all but population variation was used in logarithm form because they are right-skewed (Figure 3-5 and Figure 3-6). I ran OLS regressions to estimate the coefficients of the regressions.

### 3.3. Results

#### 3.3.1 *The spatial distribution of African wealth*

Estimates of the IW per capita and the total IW at the provincial level in 2018 are shown in Figure 3-1 and Figure 3-7 in the Appendix respectively. These maps report the wealth in key provinces, in red if that wealth is less than the national average wealth; otherwise, it is reported in blue. Figure 3-2 provides an overview of the components of wealth per province. Figure 3-2 uses a ternary diagram with percentages of human, natural, and manufactured capitals on the

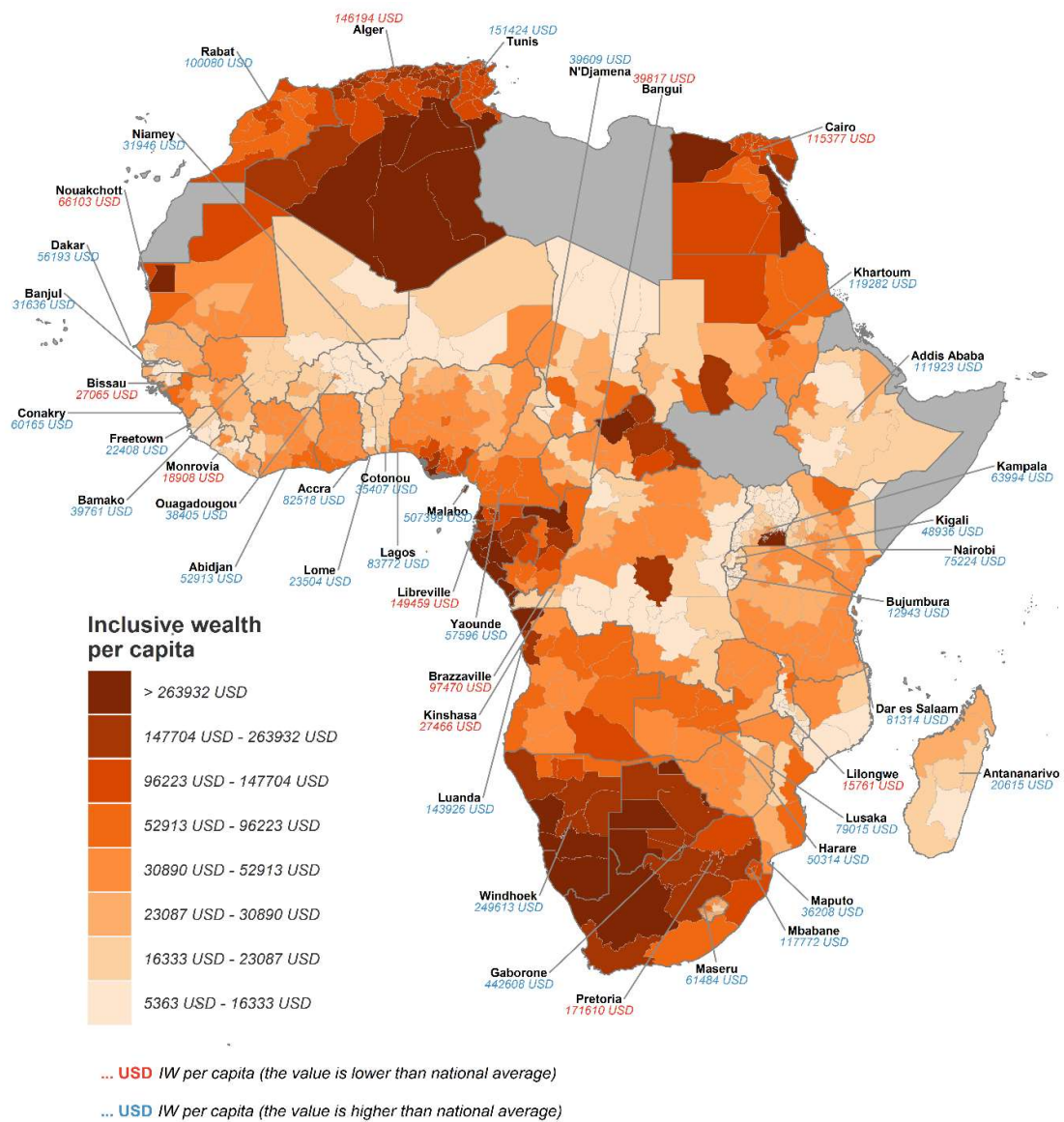
three axes and a tricolor scheme on an African map to illustrate the average composition of wealth per province (Kashnitsky and Schöley, 2018).

Distribution of per capita IW in 2018 show geographical disparities ranging from USD 5,518 in Rutana, Burundi to USD 8,412,311 in Illizi, Algeria. The average composition of wealth of African provinces is 44% natural capital, 37% human capital, and 19% manufactured capital (Figure 2). Despite this composition of wealth being similar for both wealthy and least wealthy provinces, it is dissimilar in key provinces where IW mostly comes from human capital.

The distribution of the total IW Figure 3-7 in the Appendix shows a strong correlation between the most populous provinces and the wealthiest of the continent (Pearson's correlation between population and wealth is 0.86). Yet, the distribution of IW per capita discloses more transnational patterns that provide more insightful view of wealth distribution (Figure 3-1). Distinctive patterns of per capita wealth distribution are prominent throughout continental geographical areas (see Figure 3-5).

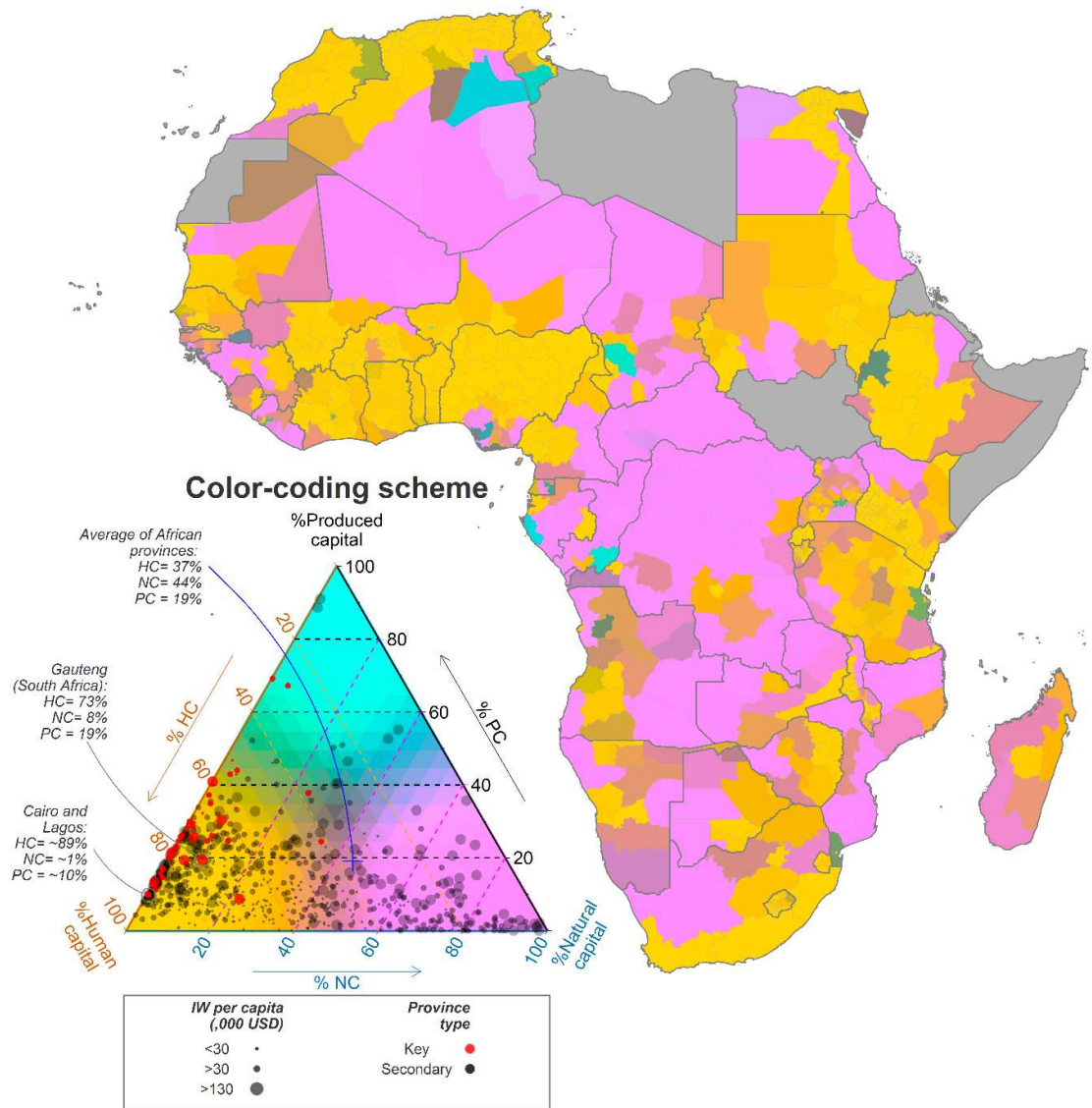


Figure 3-1 Levels of capital per capita



**Caption:** This map displays estimates of the IW per capita in 2018. The lines in light gray represents provincial boundaries and the lines in bold gray represent the national boundaries. Areas with the darkest colors have the highest IW per capita, those with the lightest color have the lowest IW per capita. The name of each key province is displayed with the value of its IW per capita. The value of IW per capita of key provinces is written in red if that wealth is less than the national average wealth; otherwise, it is reported in blue.

Figure 3-2 Nature of capital



**Caption:** This map provides an overview of the components of wealth per province. It uses the tricolor packages developed by Kashnitsky and Schöley (2018) in R. The map uses a ternary diagram with percentages of human, natural, and manufactured capitals on the three axes and a tricolor scheme on an African map to illustrate the average composition of wealth per province.

Starting with Northern and Southern Africa, provinces with the highest IW per capita are the least populated (Figure 3-1). In Egypt, Algeria, and Tunisia, the provinces with the highest IW per capita are overlapping the Sahara Desert, while in South Africa, Namibia and Botswana, they are located near the Kalahari Desert. In South Africa, for instance, the average IW per capita of provinces over the Kalahari Desert (Northern Cape and North West) is USD 341,752 while other South African provinces' average is USD 170,343. The concentration of wealth in these provinces may show that despite important natural resources (mainly fossil fuels in the North and precious minerals in the South), harsh living conditions prevent population settlements and therefore prevent the consumption and the transformation of this wealth locally.

In West Africa, however, the distribution of wealth is different since the highest levels of wealth per capita is concentrated around populous capital cities. This type of distribution of wealth is salient in Burkina Faso where the IW per capita in the Kadiogo (the key province) is USD 38,395, almost twice as large as the wealth in the Haut-Bassins (USD 23,050), the second wealthiest provinces of the country. All around West Africa, the main source of wealth is made up of human capital. Wealth distribution in East Africa is similar to that of the West in the sense that the highest IW per capita is concentrated in key provinces and the levels of wealth are equivalent. Despite being less pronounced than in the West, the sources of wealth in East

Africa are mostly made up of human capital except for Mozambique where the main source of IW in most of the country comes from natural endowment.

Finally, Central Africa is the geographical region where the disparity in wealth composition between key and secondary provinces is the greatest. While on average, wealth in key provinces are composed of 66% of human, 27% of manufactured and 7% natural capitals, these ratios in secondary provinces are respectively 33%, 7% and 60% on average in Central Africa. The importance of natural capital in Central African secondary provinces can be attributed to the important forest cover, oil reserves (in Gabon, Equatorial Guinea, Congo Brazzaville), and precious minerals reserves (Cameroon, Congo Kinshasa, Central Africa).

Acknowledging the geographical discrepancies of wealth in the continent, two broad tendencies emerge. First, national boundaries restrict the repartition of wealth per capita across African sub-national provinces. There is a concentration of wealth per capita in Northern, Southern, and Central-western countries and strong discontinuities of wealth distribution at almost all international boundaries. These discontinuities in the distribution of wealth may suggest that international boundaries significantly restrict migration of population in enjoying wealth; it may also suggest the importance of national (cultural and business) environment, government, and institutions in wealth creation, although it may simply suggest a difference in transparency of wealth information across countries.

Second, IW figures of key provinces are (1) mostly composed of human capital and are (2) higher than their respective national average aggregately and on a per capita basis. Despite the inability of the data to ascertain the origin of this trend, previous works may help us interpret this result. Building on works focusing on economic growth theories, one can attribute this concentration of wealth in capital cities to the importance and often primacy of human capital in the creation of wealth (Manuelli and Seshadri, 2014; Kraay, 2018). However, historical analyses have linked contemporary concentrations of wealth in capital cities to African colonial history. After their colonial episodes, several colony centers turned into capital cities and continue to host important political powers and to field significant investment as a legacy of this past (Baker and Therborn, 2012; Growth *et al.*, 2014).

Irrespective of the initial cause of this concentration of wealth and population size, the results portray key provinces as summits of agglomeration economies, which may allow them to attract larger inclusive investment (Growth *et al.*, 2014; Padraig and Francis, 2016). Consequently, I investigate this pattern in the analysis of wealth sustainability across provinces (Dustmann and Okatenko, 2014).

### *3.3.2 Sustainability of wealth in African provinces and the difference between key and secondary provinces*

#### *3.3.2.1. Descriptive analysis*

I display the distribution of the per capita IW growth per provinces in Africa from 2012 to 2018 in Figure 3-3. The variations of IW components are displayed in figures 3-8 – 3-12 in the Appendix. One can notice that these figures (Figure 3-3 and figures 3-8 – 3-12) exhibit some spatial distributions of growth in wealth per capita that are highly clustered by national boundaries. In this regard, several countries in West and East Africa experienced an increase in IW per capita in almost all their provinces (in Benin, Côte d’Ivoire, Morocco, Swaziland, Rwanda, and Zimbabwe, all provinces increased their IW per capita). These results highlight the importance of national environment in wealth variation. Taking Rwanda as an example, one may link its sustainability performance to the strength of its government and its ability to finance economic growth, and capital creation with growing public debt from 2012 to 2018 (see IMF publications (International Monetary Fund. African Dept., 2017) and the World Bank dataset (THE WORLD BANK, 2021)).

On the contrary, most provinces in DRC, Chad, Cameroon, and Equatorial Guinea experienced a decrease in wealth per capita. In these countries, low sustainability performance is prominent in provinces far from capital cities. The situation in the Central African Republic

is the most concerning as it is the only country where IW per capita is decreasing in all provinces. We may consider that the nationwide civil war in Central Africa Republic which started from 2013 and lasted eight years now contribute to this situation.

In Table 3-1, I present an aggregated view of provincial growths via the summary statistics of growths in IW, human, natural and manufactured capitals as well as growths in education attainment, income, ecosystem services and underground wealth. This table also provides the summary statistics for key and secondary provinces separately as well as the average difference of each variables between these two types of province.

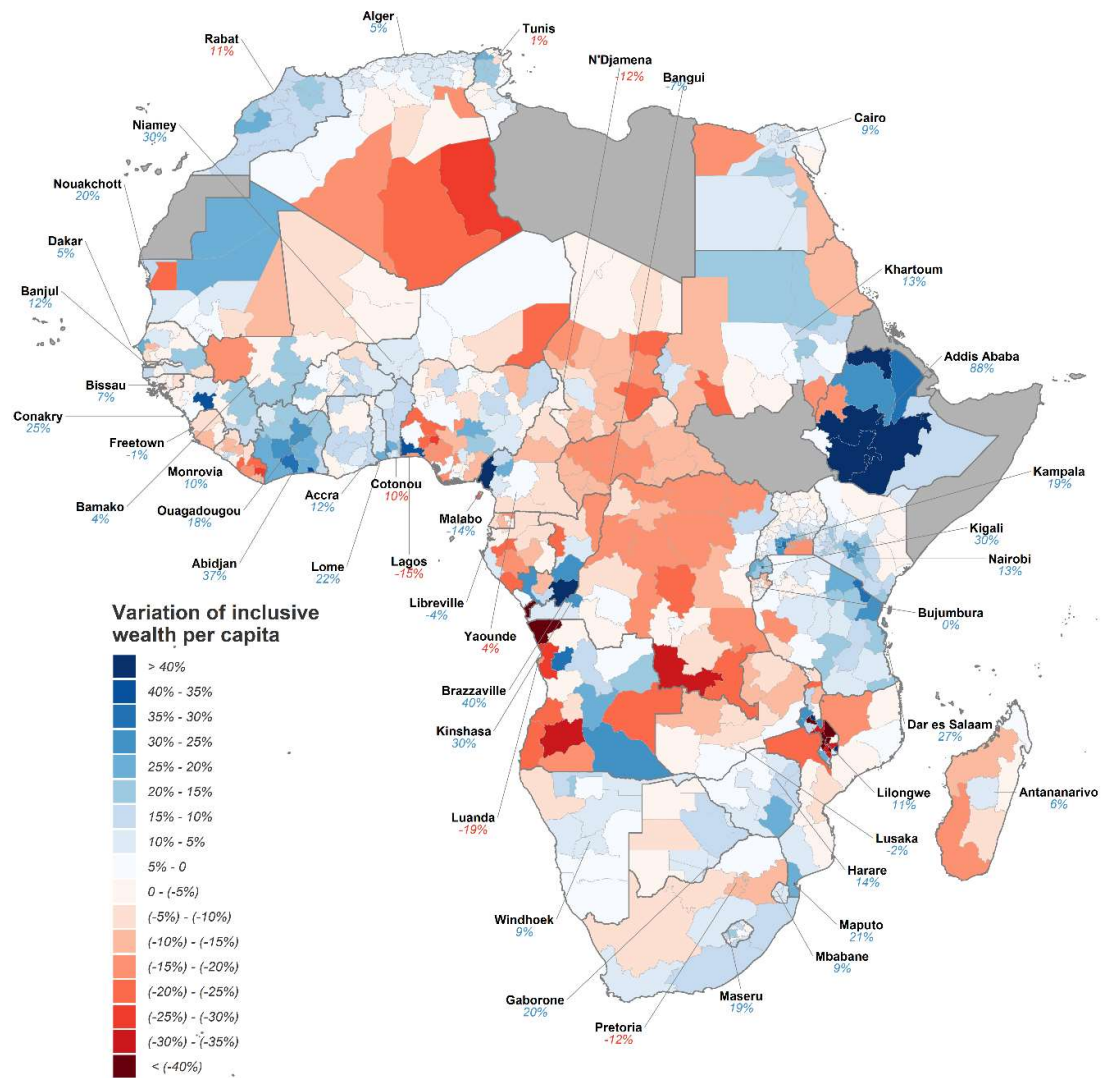
Our data shows that African provinces' IW per capita increased by 3% (t-test, 95% confidence interval (CI): 1.82% - 4.19 %) and 60% of them experienced a positive variation of wealth per capita (see Table 3-1). Per capita human, manufactured, and natural capitals varied by 10%, 17%, and -3% respectively between 2012 and 2018. Since African provinces compensate their exploitation of natural capital with higher investment in manufactured and human capitals, the data shows that these provinces follow a weak sustainability path on average.

Furthermore, except from the growth in natural capital per capita being positive in key provinces and negative in the secondary provinces, summary statistics shows that growth in IW, human capital, manufactured capital and population have the same sign in the two types

of provinces (Table 3-1). However, statistically significant estimate shows that population growth in key provinces is 3.81% higher than that of others. Notwithstanding, manufactured human, and natural capitals are 9%, 10% and 62% higher in key provinces than in others. The accumulation of population growth and IW per capita growth contributes to an inequity of wealth creation as total IW growth in key provinces is on average 14.9% higher than in secondary provinces. I further investigate this inequity of wealth growth in regression analyses that differentiate out the average performances of national level sustainability since the maps show that national environments play an important role in sustainability performances.



Figure 3-3 Variations of capital per capita



**Caption:** This map displays estimates of the variation of IW per capita from 2012 to 2018. The lines in light gray represents provincial boundaries and the lines in bold gray represent the national boundaries. Decrease in wealth are represented in blue and increases are displayed in red. I displayed the estimates of IW per capita variations of key provinces in red if that variations in wealth is lower than the national average variations; otherwise, it is reported in blue.

### *3.3.2.2.Differences in IW growth between key and secondary provinces*

In Table 3-2, I estimate the difference in growths of IW, human, natural and manufactured capitals between key and secondary provinces from 2012 to 2018. The coefficients of the regression analyses should be interpreted with reference to national averages because I control for countries' fixed effects. These estimates show that when compared with provinces within the same country, key provinces had an IW growth 8.41% ( $p < 0.01$ ) higher than that of other provinces. The difference in IW growth between these two types of province differ from their difference in income growth (a proxy of regional GDP growth) since the data show that there was no difference in income growth between key and secondary provinces (see table 3-8). This nonalignment of results suggests that changes in wealth and income disclose different information for African provinces. Thus, focusing on variations policy index based on income such as GDP may mask sustainability performance of provinces.

Our regressions further reveal a statistically significant difference in manufactured capital growth between key and secondary provinces of 9% ( $p < 0.01$ ) but a weak difference in human and natural capital growths of 60% ( $p > 0.1$ ) and 3% ( $p = 0.07$ ) respectively. Overall, these regressions suggest that the difference in IW growth between key and secondary provinces is composed by the higher investment in manufactured capital in key provinces.

Table 3-1 Growth in IWI and its components from 2012 to 2018

| Variables  |                             | Average<br>in whole<br>sample<br><i>N</i> = 714 | Average in<br>key<br>provinces<br><i>N</i> = 48 | Average in<br>secondary<br>provinces<br><i>N</i> = 669 | Difference<br>between<br>key and<br>secondary<br>provinces<br>(t-test) |
|------------|-----------------------------|-------------------------------------------------|-------------------------------------------------|--------------------------------------------------------|------------------------------------------------------------------------|
| Per capita | Population growth           | 16.7%                                           | 20.2%                                           | 16.4%                                                  | 3.81%*                                                                 |
|            | IW growth                   | 2.4%                                            | 11.8%                                           | 1.7%                                                   | 10.1%***                                                               |
|            | Human capital growth        | 10.1%                                           | 8.5%                                            | 10.2%                                                  | -1.69%                                                                 |
|            | Manufactured capital growth | 17.8%                                           | 27.7%                                           | 17.1%                                                  | 10.6%**                                                                |
|            | Natural capital growth      | -3.1%                                           | 54.6%                                           | -7.3%                                                  | 61.9%***                                                               |
|            | Income growth               | 9.4%                                            | 7.9%                                            | 9.5%                                                   | -1.66%                                                                 |
|            | Education growth            | 12.6%                                           | 9.5%                                            | 12.8%                                                  | -3.27%                                                                 |
|            | Ecosystem capital growth    | 9.4%                                            | 61.7%                                           | 5.7%                                                   | 56%***                                                                 |
|            | Underground capital growth  | -13.4%                                          | -16.3%                                          | -13.2%                                                 | -3.04%                                                                 |
| Total      | IW growth                   | 19.0%                                           | 34.1%                                           | 17.9%                                                  | 16.2%***                                                               |
|            | Human capital growth        | 28.5%                                           | 30.1%                                           | 28.4%                                                  | 17.1%                                                                  |
|            | Manufactured capital growth | 36.6%                                           | 53.2%                                           | 35.4%                                                  | 17.8%***                                                               |
|            | Natural capital growth      | 12.5%                                           | 87.5%                                           | 7.1%                                                   | 80.4%***                                                               |
|            | Income growth               | 27.7%                                           | 29.4%                                           | 27.6%                                                  | 1.74%                                                                  |
|            | Education growth            | 12.6%                                           | 9.5%                                            | 12.8%                                                  | -3.27%                                                                 |
|            | Ecosystem capital growth    | 25.9%                                           | 96.0%                                           | 20.9%                                                  | 75.1%                                                                  |
|            | Underground capital growth  | -10.2%                                          | -12.3%                                          | -10.0%                                                 | -2.3%                                                                  |

Level of statistical significance \*\*\*  $p < 0.01$ , \*\*  $p < 0.05$ , \*  $p < 0.1$

Table 3-2 Difference in capital formation within countries between key and secondary provinces

| VARIABLES               | (1)<br>Growth in<br>IW  | (2)<br>Growth in<br>Natural Capital | (3)<br>Growth in<br>Human<br>Capital | (4)<br>Growth in<br>Manufactured<br>Capital |
|-------------------------|-------------------------|-------------------------------------|--------------------------------------|---------------------------------------------|
| Key provinces           | 0.0922***<br>(0.0251)   | 0.661<br>(0.556)                    | -0.0317*<br>(0.0174)                 | 0.0994***<br>(0.0190)                       |
| Constant                | -0.0832***<br>(0.00139) | -0.134***<br>(0.0309)               | 0.0142***<br>(0.000969)              | 0.169***<br>(0.00106)                       |
| Observations            | 714                     | 714                                 | 714                                  | 714                                         |
| R-squared               | 0.354                   | 0.093                               | 0.636                                | 0.762                                       |
| Country Factor variable | Yes                     | Yes                                 | Yes                                  | Yes                                         |

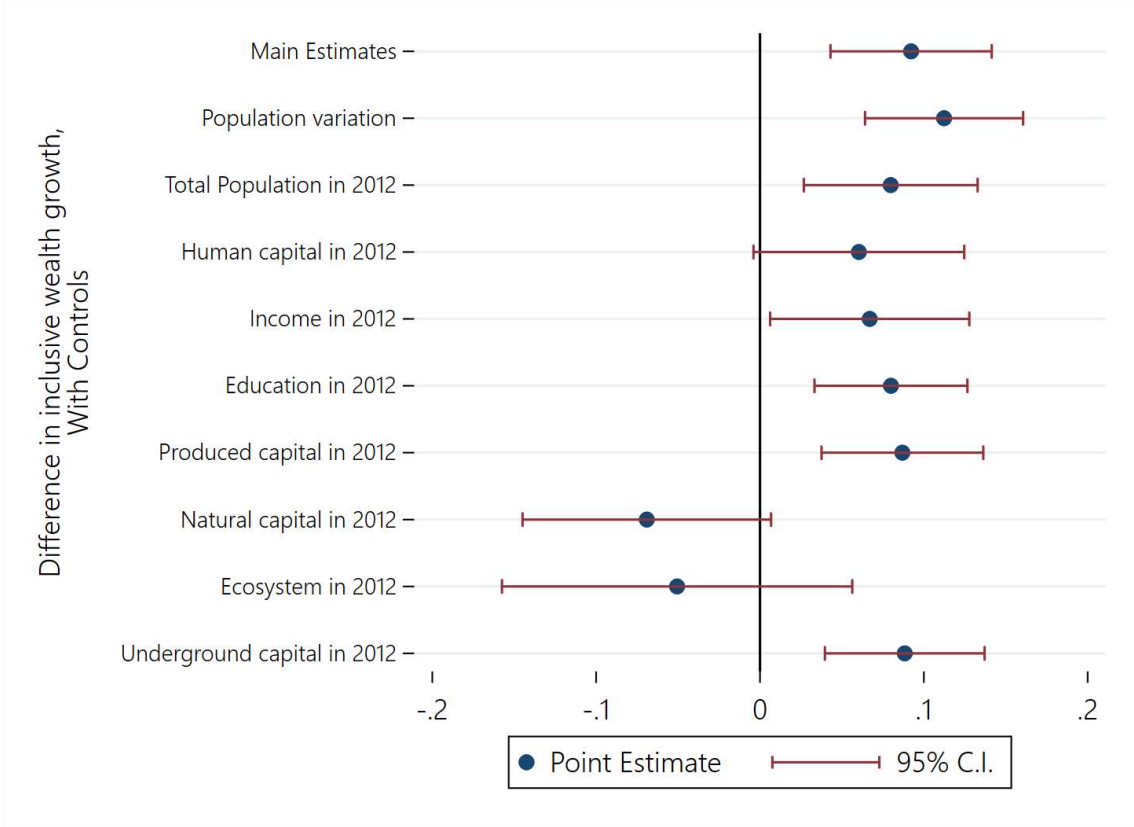
Standard errors are clustered at the county level \*\*\* p<0.01, \*\* p<0.05, \* p<0.1

I attempt to determine the reasons behind the higher IW growth in key provinces in Figure 3-4, where I display the estimate of the difference in IW growth between key and secondary provinces from seven different regression analyses. In this figure, each blue circle is the estimate of the difference of IW growth between key and secondary provinces. The red horizontal line shows a 95% confidence interval. The “Main estimate” restates the results from Table 3-2 column (1). Each other lines controls for one additional variable. These variables are provincial population growth between 2012-2018; and the total population, as well as the manufactured, human, and natural capitals in 2012 to determine the forces that account for the difference in growth in IW per capita (see Methods). These variables are used in logarithm form except for population variation.

The results in Figure 3-4 show that controlling for total population, produced capital and population growth one by one produces little change to the estimated difference in IW growth between key and secondary provinces. Controlling out the amount of human and natural capitals in 2012, separately, reveals that IW growth in key and secondary provinces is not significantly different. Regarding the role of human capital in this outcome, the results suggest that level of income per capita accounts for an important share of the difference of IW growth between key provinces and the other provinces. On the other hand, splitting natural capital into ecosystem services wealth and underground wealth shows that the comparison of key provinces with provinces with similar amount of ecosystem wealth in 2012 renders a difference in IW

growth that is statistically indifferent from 0. This result may suggest an unsustainable management of ecosystem wealth in secondary provinces since their IW originates mostly from natural capital.

Figure 3-4 Explanations of the difference in IW growth between key and secondary provinces



**Caption:** Each blue circle represents the coefficient on the dummy variable for key provinces from a regression of Inclusive wealth growth as dependent variable. The red bar depicts the robust 95% confidence interval. Each regression contains control countries fixed effects by the addition of countries dummy variables. Each regression contains a feature of the IW composition, indicated on the left of the graph. These features are accounted for in logarithm form except for the population variation. Standard errors are clustered at the country level.

### **3.4. Discussions and conclusion**

Overall, the results show significant differences in subnational sustainability performances in Africa at the provincial level. First, it appears that the wealth of African provinces increased on average by 3% between 2012 and 2018, an increase that is composed of growths in human and manufactured capitals and a decline in natural capital per capita. Moreover, wealth in provinces with African national capital cities increases faster than that of secondary provinces by 8%, on a per capita basis. This higher-than-national-average IW growth is connected to a difference in ecosystem wealth and in part to a difference in levels of income. Thus, it does not imply governmental preference for investment in these provinces.

Regarding the implication, the results mapping the levels of IW, its nature, and the sustainability of wealth at the provincial level permit the identification of areas where wealth management needs to be improved. On this matter, there is a potential for sustainability improvement within nations since I find that 37 out of 44 countries have at least 1 province that is sustainable, and at least 1 province that is not. Although local economies may be different, it may be an opportunity for national governments to replicate successful management of wealth to provinces with lower performances. Examples such as (1) the low sustainability path in Central African Republic, affected by its long-lasting civil war, (2) the high sustainability path in Rwanda governed through public debt, and (3) the significant



differences of sustainability at international boundaries highlight the strength of national environments and institutions in affecting sustainability performances.

In certain countries however, the differences in sustainability performance is highly influenced by the nature of wealth across provinces. For instance, the fall in IW per capita along Angola's coast is linked to the country's extensive offshore oil production, which peaked between 2008 and 2015. This decline in IW per capita is an intriguing instance of an unsustainable development path, as Angola's GDP per capita benefits from the export revenues of its oil production (OECD/IEA, 2006). In this case, there is a national need to depart from this short-sighted development where investment in other forms of capital cannot compensate for the depletion of oil reserves at the country level nor at the provincial level.

Moreover, the results can be used, in general, to improve accountability in need-based investment strategies from national to local level. Regarding the emphasis of this study, the fast IW growth in provinces with a national capital city may inform the types of investment needed to deal with population growth or to achieve several SDGs. Regarding population growth, the fast IW growth in provinces with capital cities represent a natural incentive for population migration toward them. This will lead to significant population density which can be linked to undesired effects such as slum formations and pollution. Thus, it will be important to invest in smart technologies in these areas to reduce the ills of population growth while leveraging its benefits. Alternatively, the results may help appraising several sustainable

development goals (SDGs) developed by the UN. The higher sustainability performance in provinces with capital cities suggest that African countries follow a development path inconsistent to SDG 10 (Reduce inequality within and among countries). Thus, African countries will need to reconsider their development path to achieve these SDG.

Furthermore, the results show a need to address the widely spread depletion of natural capital across provinces in the continent. If not managed properly, this depletion may reach tipping points, beyond which further exploitation of resources produces more than proportional harm to populations and the environment. Considering that (1) the African population and its economy are expected to grow, and (2) that the continent is the home of unique ecosystems and resources, each country and its trade partners will benefit from defining sustainable uses of natural resources. Since areas with important ecosystem wealth, like key provinces, have highest IW growth the results approve an investment in ecosystems and conform to UNEP and World Bank recommendation of investment in agricultural productivity.

Nevertheless, since most of this research is based on incomplete data, it should not be seen as a conclusion on wealth accounting. Among the shortcomings of the estimates, three should be held in mind while interpreting the results. I do not formally account (1) for important natural assets such as fisheries and livestock and (2) assets that generate economic activities but whose worth is “inestimable” (for instance historical monuments such as Egypt pyramids or natural assets such as the Victoria Falls). Moreover, I only value natural mineral and fossil

fuels reserves that are proven and whose information is freely disclosed. Finally, the estimate of changes in wealth may be affected by measurement errors in the proxy of capital stocks or by the disregard of the role of technology increasing the worth of capitals because the analysis assumes that the only driver on IW growth is variations in capital stock.

Rather, this work should be seen as an effort to make a step in the right direction and as advocacy for a transparent valuation of sub-national wealth as means to assess its sustainable use. For the moment, the use of satellite data for the feasibility of this work extends the suggestions from UNEP and World Bank stemming from national sustainability estimates (Lange, Wodon and Carey, 2018; Managi, Kumar and Pushpam, 2018). As argued by (Karunanithi *et al.*, 2011), the results show that it vital to strengthen national environments and institutions for a geographically shared sustainability path in Africa and to share efficient sustainability model across provinces within countries.

## References

- African Development Bank Group (2016) *Africa Infrastructure Knowledge Program*.
- Amegah, A. K. and Agyei-Mensah, S. (2017) 'Urban air pollution in Sub-Saharan Africa: Time for action', *Environmental Pollution*, 220, pp. 738–743. doi: 10.1016/j.envpol.2016.09.042.
- Arrow, K. J. *et al.* (2012) 'Sustainability and the measurement of wealth', *Environment and Development Economics*, 17(3), pp. 317–353. doi: 10.1017/S1355770X12000137.
- Baker, S. and Therborn, G. (2012) *Power and powerlessness: capital cities in Africa*. Edited by S. Baker and G. Therborn. Cape Town: HSRC Press.
- Barrett, C. B. *et al.* (2017) *On the Structural Transformation of Rural Africa*. WPS7938. Washington DC.
- Bateman, I. J. and Mace, G. M. (2020) 'The natural capital framework for sustainably efficient and equitable decision making', *Nature Sustainability*, 3(10), pp. 776–783. doi: 10.1038/s41893-020-0552-3.
- Brandt, M. *et al.* (2017) 'Human population growth offsets climate-driven increase in woody vegetation in sub-Saharan Africa', *Nature Ecology & Evolution*, 1(4), p. 0081. doi: 10.1038/s41559-017-0081.
- Bren d'Amour, C. *et al.* (2017) 'Future urban land expansion and implications for global croplands', *Proceedings of the National Academy of Sciences*, 114(34), pp. 8939–8944. doi: 10.1073/pnas.1606036114.
- Chakravarty, A. (2011) *Dams and Infant Mortality in Africa*. <http://www.homepages.ucl.ac.uk/~uctpabc/Dams%20and%20Infant%20Mortality%20in%20Africa201>.
- Davis, J. C. and Vernon Henderson, J. (2003) 'Evidence on the political economy of the urbanization process', *Journal of Urban Economics*, 53(1), pp. 98–125. doi: 10.1016/S0094-1190(02)00504-1.
- Dorosh, P. and Thurlow, J. (2014) 'Can Cities or Towns Drive African Development? Economywide Analysis for Ethiopia and Uganda', *World Development*, 63, pp. 113–123. doi: 10.1016/j.worlddev.2013.10.014.
- Dustmann, C. and Okatenko, A. (2014) 'Out-migration, wealth constraints, and the quality of local amenities', *Journal of Development Economics*, 110, pp. 52–63. doi: 10.1016/j.jdeveco.2014.05.008.
- Elvidge, C. D. *et al.* (2012) 'The Night Light Development Index (NLDI): a spatially explicit

measure of human development from satellite data', *Social Geography*, 7(1), pp. 23–35. doi: 10.5194/sg-7-23-2012.

Elvidge, C. D. *et al.* (2021) 'Annual Time Series of Global VIIRS Nighttime Lights Derived from Monthly Averages: 2012 to 2019', *Remote Sensing*, 13(5), p. 922. doi: 10.3390/rs13050922.

Friedl, M. . *et al.* (2002) 'Global land cover mapping from MODIS: algorithms and early results', *Remote Sensing of Environment*, 83(1–2), pp. 287–302. doi: 10.1016/S0034-4257(02)00078-0.

Ghadimi, H., Harris, T. and Warner, T. (2015) 'A geospatial approach to measuring regional and sub-regional inclusive wealth: the case of resource rich West Virginia, USA', *Regional Science Policy & Practice*, 7(4), pp. 199–224. doi: 10.1111/rsp3.12067.

Graetz, N. *et al.* (2018) 'Mapping local variation in educational attainment across Africa', *Nature*, 555(7694), pp. 48–53. doi: 10.1038/nature25761.

de Groot, R. *et al.* (2012) 'Global estimates of the value of ecosystems and their services in monetary units', *Ecosystem Services*, 1(1), pp. 50–61. doi: 10.1016/j.ecoser.2012.07.005.

Growth, E. *et al.* (2014) 'National Institutions and Subnational', *The Quarterly Journal of Economics*, (January), pp. 151–213. doi: 10.1093/qje/qjt029.Advance.

Henderson, J. V., Storeygard, A. and Weil, D. N. (2012) 'Measuring Economic Growth from Outer Space Measuring Economic Growth from Outer Space', *American Economic Review*, 102(2), pp. 994–1028.

Hilson, G. *et al.* (2018) 'Formalizing artisanal gold mining under the Minamata convention: Previewing the challenge in Sub-Saharan Africa', *Environmental Science & Policy*, 85, pp. 123–131. doi: 10.1016/j.envsci.2018.03.026.

Holden, E., Linnerud, K. and Banister, D. (2014) 'Sustainable development: Our Common Future revisited', *Global Environmental Change*, 26, pp. 130–139. doi: 10.1016/j.gloenvcha.2014.04.006.

IMF (2019) *FAD investment and capital stock database 2019: manual & FAQ - Estimating public, private, and PPP capital stocks*. Washington DC.

International Monetary Fund. African Dept. (2017) *The Rwandan government is implementing a policy adjustment program, supported*. Washington DC.

Jean, N. *et al.* (2016) 'Combining satellite imagery and machine learning to predict poverty', *Science*, 353(6301), pp. 790–794. doi: 10.1126/science.aaf7894.

- Karunanithi, A. T. *et al.* (2011) ‘The characterization of socio-political instability, development and sustainability with Fisher information’, *Global Environmental Change*, 21(1), pp. 77–84. doi: 10.1016/j.gloenvcha.2010.11.002.
- Kashnitsky, I. and Schöley, J. (2018) ‘Regional population structures at a glance’, *The Lancet*, 392(10143), pp. 209–210. doi: 10.1016/S0140-6736(18)31194-2.
- Kraay, A. (2018) *Methodology for a World Bank Human Capital Index*. WPS8593.
- Lange, G.-M., Wodon, Q. and Carey, K. (2018) *The Changing Wealth of Nations 2018: Building a Sustainable Future*. Washington, DC: World Bank. doi: 10.1596/978-1-4648-1046-6.
- Liang, H. *et al.* (2014) ‘Modeling In-Use Steel Stock in China’s Buildings and Civil Engineering Infrastructure Using Time-Series of DMSP/OLS Nighttime Lights’, *Remote Sensing*, 6(6), pp. 4780–4800. doi: 10.3390/rs6064780.
- Lujala, P., Ketil Rod, J. and Thieme, N. (2007) ‘Fighting over Oil: Introducing a New Dataset’, *Conflict Management and Peace Science*, 24(3), pp. 239–256. doi: 10.1080/07388940701468526.
- Lutz, W. *et al.* (2021) ‘Skills-adjusted human capital shows rising global gap’, *Proceedings of the National Academy of Sciences*, 118(7). doi: 10.1073/pnas.2015826118.
- Managi, S., Kumar, P. and Pushpam, K. (2018) *Inclusive Wealth Report 2018, Inclusive Wealth Report 2018*. Edited by S. Managi and P. Kumar. New York, USA: Routledge. doi: 10.4324/9781351002080.
- Manuelli, R. E. and Seshadri, A. (2014) ‘Human Capital and the Wealth of Nations’, *American Economic Review*, 104(9), pp. 2736–2762. doi: 10.1257/aer.104.9.2736.
- Michalopoulos, S. and Papaioannou, E. (2013) ‘Pre-Colonial Ethnic Institutions and Contemporary African Development’, *Econometrica*, 81(1), pp. 113–152. doi: 10.3982/ecta9613.
- Mumford, K. J. (2012) ‘Measuring inclusive wealth at the state level in the United States’, in UNEP and UNU-IHDP (ed.) *Inclusive Wealth Report 2012*. Cambridge University Press, pp. 69–86.
- Njuguna, C. and McSharry, P. (2017) ‘Constructing spatiotemporal poverty indices from big data’, *Journal of Business Research*, 70, pp. 318–327. doi: 10.1016/j.jbusres.2016.08.005.
- OECD/IEA (2006) *Angola: Towards an Energy Strategy*. Paris.
- Onyeji, I., Bazilian, M. and Nussbaumer, P. (2012) ‘Contextualizing electricity access in sub-

- Saharan Africa', *Energy for Sustainable Development*, 16(4), pp. 520–527. doi: 10.1016/j.esd.2012.08.007.
- Padraig, C. and Francis, O. (2016) 'Neoliberalism, Urbanization and Change in Africa: the Political Economy of Heterotopias', *Journal of African Development*, 18(1), pp. 61–73.
- Pokhriyal, N. and Jacques, D. C. (2017) 'Combining disparate data sources for improved poverty prediction and mapping', *Proceedings of the National Academy of Sciences*, 114(46), pp. E9783–E9792. doi: 10.1073/pnas.1700319114.
- Polasky, S. *et al.* (2015) 'Inclusive Wealth as a Metric of Sustainable Development', *Annual Review of Environment and Resources*, 40(1), pp. 445–466. doi: 10.1146/annurev-environ-101813-013253.
- Roser, M., Ortiz-Ospina, E. and Ritchie, H. (2013) 'Life expectancy', *Our World in Data*.
- Scholes, R. J. and Biggs, R. (2005) 'A biodiversity intactness index', *Nature*, 434(7029), pp. 45–49. doi: 10.1038/nature03289.
- Showers, K. B. (2002) 'Water Scarcity and Urban Africa: An Overview of Urban–Rural Water Linkages', *World Development*, 30(4), pp. 621–648. doi: 10.1016/S0305-750X(01)00132-2.
- Smits, J. and Permanyer, I. (2019) 'The Subnational Human Development Database', *Scientific Data*, 6(1), p. 190038. doi: 10.1038/sdata.2019.38.
- Steckel, J. C. *et al.* (2020) 'Coal and carbonization in sub-Saharan Africa', *Nature Climate Change*, 10(1), pp. 83–88. doi: 10.1038/s41558-019-0649-8.
- Sulla-Menashe, D. and Friedl, M. A. (2018) *User Guide to Collection 6 MODIS Land Cover (MCD12Q1 and MCD12C1) Product*. Reston, VA, USA. doi: 10.5067/MODIS/MCD12Q1.
- Tatem, A. J. (2017) 'WorldPop, open data for spatial demography', *Scientific Data*, 4(1), p. 170004. doi: 10.1038/sdata.2017.4.
- Taylor, C. D. *et al.* (2005) *Geology and Nonfuel Mineral Deposits of Africa and the Middle East*. US Geological Survey.
- THE WORLD BANK (2021) *External debt stocks*.
- U.S. Geological Survey (2018) *Mineral Commodity Summaries*. Washington DC.
- UNU-IHDP & UNEP (2014) *Inclusive Wealth Report 2014 – Measuring Progress Toward Sustainability*, *International Journal of Sustainability in Higher Education*. doi: 10.1108/ijshe.2012.24913daa.006.

UNU-IHDP & UNEP, Report, I. W. and Report, I. W. (2012) *Inclusive Wealth Report 2012 – Measuring Progress Toward Sustainability, International Journal of Sustainability in Higher Education*. New York: Cambridge University Press. doi: 10.1108/ijshe.2012.24913daa.006.

Vörösmarty, C. J. *et al.* (2021) ‘A green-gray path to global water security and sustainable infrastructure’, *Global Environmental Change*, 70, p. 102344. doi: 10.1016/j.gloenvcha.2021.102344.

Weidmann, N. B. and Schutte, S. (2017) ‘Using night light emissions for the prediction of local wealth’, *Journal of Peace Research*, 54(2), pp. 125–140. doi: 10.1177/0022343316630359.

Yeh, C. *et al.* (2020) ‘Using publicly available satellite imagery and deep learning to understand economic well-being in Africa’, *Nature Communications*, 11(1), p. 2583. doi: 10.1038/s41467-020-16185-w.

Yerema, C. T. C. T. *et al.* (2020) ‘Differences in Water Policy Efficacy across South African Water Management Areas’, *Ecological Economics*, 175, p. 106707. doi: 10.1016/j.ecolecon.2020.106707.

Zhang, B., Nozawa, W. and Managi, S. (2021) ‘Spatial inequality of inclusive wealth in China and Japan’, *Economic Analysis and Policy*, 71, pp. 164–179. doi: 10.1016/j.eap.2021.04.014.



### **Appendix 3**

Since satellite data have a 500 m<sup>2</sup> resolution, there are errors in the land cover classification of pixels where these pixels are snapshots of a combination of small land areas. Therefore, the MODIS dataset classify land covers based on some thresholds and intervals of coverages (see Table 3-3 column 1). These thresholds imply that pixels of land covers are not a 100 % representation of that land types in areas. Thus, for each land cover I assume the effective coverages of land cover to be the midpoint of their interval of coverage (see Table 3-3 column 2). To value the ecosystem of savannah, I assumed that they are a mix between woodlands and grasslands. The proportion of woodlands and grasslands in the value of savannah is determined by the proportion of tree covers in the savannah.

Table 3-3 Correspondence between land cover and ecosystem values

| <b>MODIS land covers types (FAO-Land Cover Classification System)</b> | <b>Land cover assumed coverage</b> | <b>Mean ecosystem value<sup>32</sup></b>                     | <b>Final value in USD/ ha*</b> |
|-----------------------------------------------------------------------|------------------------------------|--------------------------------------------------------------|--------------------------------|
| Water bodies (area at 60 to 100%)                                     | 80%                                | Inland wetlands (\$25,682 / ha)                              | 22290                          |
| Evergreen need leaf forest (area at 60 to 100%)                       | 80%                                | Forest tropical (\$5,264 / ha)                               | 4569                           |
| Evergreen broadleaf Forest (area at 60 to 100%)                       | 80%                                |                                                              | 4569                           |
| Deciduous need leaf forest (area at 60 to 100%)                       | 80%                                |                                                              | 4569                           |
| Deciduous broadleaf forest (area at 60 to 100%)                       | 80%                                |                                                              | 4569                           |
| Mixed forests (area at 60 to 100%)                                    | 80%                                |                                                              | 4569                           |
| Closed Shrublands (area at 60 to 100%)                                | 80%                                | Grassland (\$2,871 / ha)                                     | 2492                           |
| Open Shrublands (10 - 60 % of tree cover)                             | 35%                                |                                                              | 1090                           |
| Woody Savannas (30 – 60% of tree cover)                               | 80%                                | 55% Grassland (\$2,871 / ha) and 45% Woodland (\$1,588 / ha) | 2488                           |
| Savannas (10 – 30% of area covered)                                   | 80%                                | 80% Grassland (\$2,871 / ha) and 20% Woodland (\$1,588 / ha) | 2836                           |
| Grasslands (at least 50 % of area)                                    | 50%                                | Grassland (\$2,871 / ha)                                     | 1557                           |
| Open ocean (2 km <sup>2</sup> of coverage)                            | 100%                               | Coastal system (\$28,917/ha)                                 | 31372                          |
| Open ocean (98 km <sup>2</sup> of coverage)                           | 100%                               | Marine (\$491/ha)                                            | 533                            |

\*The values collected for 2007 are actualised to 2011 dollars by multiplying them by 1.0849

Table 3-4 Oil reserves per country in 2018

| Country           | Oil reserves in barrels |
|-------------------|-------------------------|
| Libya             | 48,363,000,000          |
| Nigeria           | 37,070,000,000          |
| Algeria           | 12,200,000,000          |
| Angola            | 8,423,000,000           |
| Sudan             | 5,000,000,000           |
| Egypt             | 4,400,000,000           |
| Uganda            | 2,500,000,000           |
| Gabon             | 2,000,000,000           |
| Congo             | 1,600,000,000           |
| Chad              | 1,500,000,000           |
| Equatorial Guinea | 1,100,000,000           |
| Ghana             | 660,000,000             |
| Tunisia           | 425,000,000             |
| Cameroon          | 200,000,000             |
| DR Congo          | 180,000,000             |
| Niger             | 150,000,000             |
| Côte d'Ivoire     | 100,000,000             |
| Mauritania        | 20,000,000              |
| South Africa      | 15,000,000              |
| Benin             | 8,000,000               |
| Morocco           | 684,000                 |
| Ethiopia          | 428,000                 |

Sources: British Petroleum and U.S. Energy Information Administration

Deposit in 2012 is estimated by subtracting oils reserves in 2018 by cumulative production of oils between these two years.

Table 3-5 Gas reserves per country in 2018

| <b>Country</b>    | <b>Gas reserves (cubic meters)</b> |
|-------------------|------------------------------------|
| Nigeria           | 5,110,908,256,500                  |
| Algeria           | 4,503,908,259,900                  |
| Mozambique        | 2,831,685,000,000                  |
| Egypt             | 2,186,060,820,000                  |
| Libya             | 1,505,975,033,550                  |
| Angola            | 274,984,930,350                    |
| Cameroon          | 135,071,374,500                    |
| Congo             | 90,613,920,000                     |
| Sudan             | 84,950,550,000                     |
| Tunisia           | 65,128,755,000                     |
| Namibia           | 62,297,070,000                     |
| Rwanda            | 56,633,700,000                     |
| Equatorial Guinea | 36,811,905,000                     |
| Gabon             | 28,316,850,000                     |
| Mauritania        | 28,316,850,000                     |
| Côte d'Ivoire     | 28,316,850,000                     |
| Ethiopia          | 24,918,828,000                     |
| Ghana             | 22,653,480,000                     |
| Uganda            | 14,158,425,000                     |
| Tanzania          | 6,512,875,500                      |
| Somalia           | 5,663,370,000                      |
| Morocco           | 1,444,159,350                      |
| Benin             | 1,132,674,000                      |
| DR Congo          | 991,089,750                        |

Sources: British Petroleum and U.S. Energy Information Administration

Deposit in 2012 is estimated by subtracting oils reserves in 2018 by cumulative production of gas between these two years.

Table 3-6 Coal reserves per country in 2018

| <b>Country</b>           | <b>Coal reserves (tons)</b> |
|--------------------------|-----------------------------|
| South Africa             | 35,053,458,000              |
| Mozambique               | 1,975,339,520               |
| Zimbabwe                 | 553,359,620                 |
| Nigeria                  | 379,194,640                 |
| Tanzania                 | 296,521,390                 |
| Eswatini                 | 158,732,640                 |
| Botswana                 | 106,434,000                 |
| DR Congo                 | 97,003,280                  |
| Algeria                  | 65,036,290                  |
| Zambia                   | 49,603,950                  |
| Egypt                    | 17,636,960                  |
| Morocco                  | 15,432,340                  |
| Niger                    | 6,613,860                   |
| Central African Republic | 3,306,930                   |
| Malawi                   | 2,204,620                   |

Sources: British Petroleum and U.S. Energy Information Administration

Deposit in 2012 is estimated by subtracting oils reserves in 2018 by cumulative production of coal between these two years.

Table 3-7 Commodities' price

| Commodity                  | Price in 2011 USD | Sources of price                                                                                                                                                                                                  | Rental rates (World Bank estimates) |
|----------------------------|-------------------|-------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|-------------------------------------|
| Oil (crude barrel)         | 111.26            | <a href="http://www.bp.com/statisticalreview">http://www.bp.com/statisticalreview</a>                                                                                                                             | 0.7                                 |
| Natural gas (millions Btu) | 4.01              | <a href="http://www.bp.com/statisticalreview">http://www.bp.com/statisticalreview</a>                                                                                                                             | 0.35                                |
| Coal (ton)                 | 84.75             | <a href="http://www.bp.com/statisticalreview">http://www.bp.com/statisticalreview</a>                                                                                                                             | 0.5                                 |
| Bauxite (ton)              | 28                | <a href="http://www.Mining.com">http://www.Mining.com</a>                                                                                                                                                         | 0.5                                 |
| Chromium (ton)             | 149               | <a href="https://assets.kpmg/content/dam/kpmg/xx/pdf/2018/11/kpmg-commodity-insights-bulletin-chromite.pdf">https://assets.kpmg/content/dam/kpmg/xx/pdf/2018/11/kpmg-commodity-insights-bulletin-chromite.pdf</a> | 0.5                                 |
| Cobalt (ton)               | 39.66             | <a href="http://www.bp.com/statisticalreview">http://www.bp.com/statisticalreview</a>                                                                                                                             | 0.5                                 |
| Copper (ton)               | 9739              | 2010 Minerals Yearbook                                                                                                                                                                                            | 0.5                                 |
| Gold (g)                   | 43.07             | <a href="http://www.Mining.com">http://www.Mining.com</a>                                                                                                                                                         | 0.5                                 |
| Lead (ton)                 | 2585              | <a href="http://www.Mining.com">http://www.Mining.com</a>                                                                                                                                                         | 0.5                                 |
| Ore manganese (ton)        | 10                | Mineral Industry Surveys                                                                                                                                                                                          | 0.5                                 |
| Ore Zinc (ton)             | 2334              | <a href="http://www.Mining.com">http://www.Mining.com</a>                                                                                                                                                         | 0.5                                 |
| Diamond (carats)           | 100               | 2010 Minerals Yearbook                                                                                                                                                                                            | 0.5                                 |
| Tin (tons)                 | 26940             | 2010 Minerals Yearbook                                                                                                                                                                                            | 0.5                                 |
| Silver (gram)              | 1                 | 2010 Minerals Yearbook                                                                                                                                                                                            | 0.5                                 |

Figure 3-5 Histogram of variation of IWI

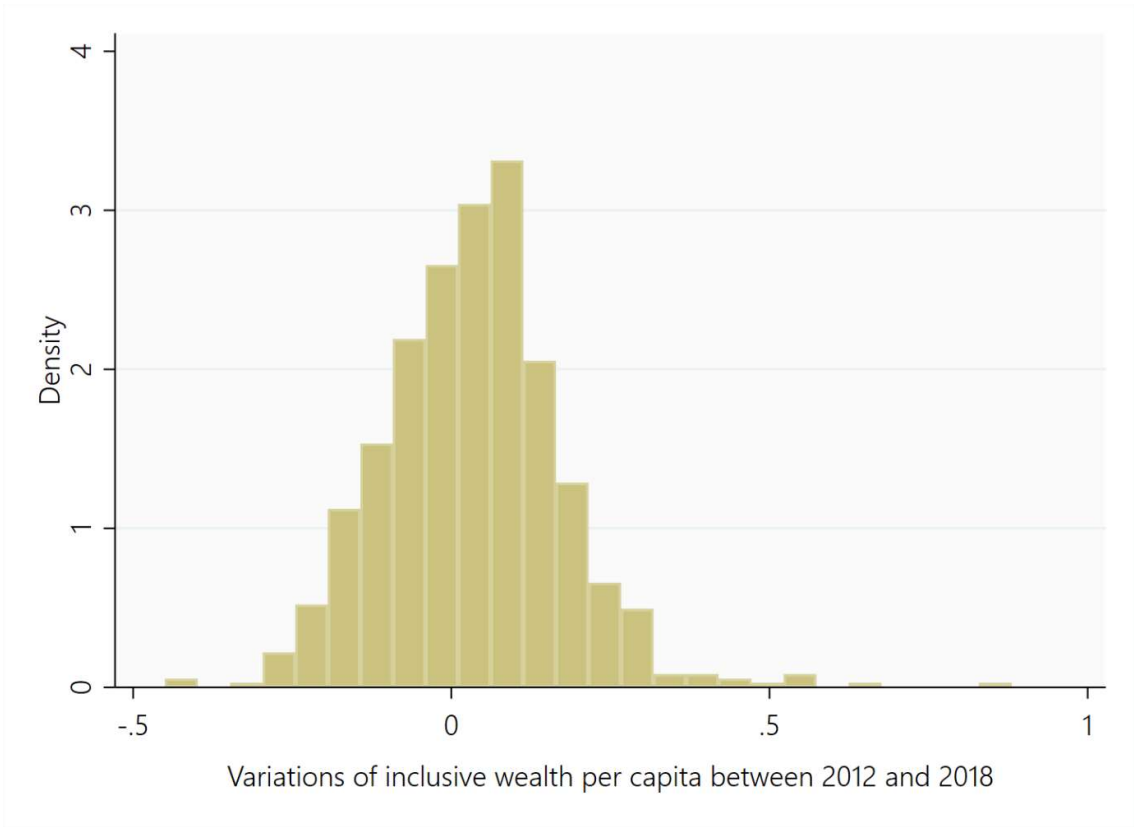


Figure 3-6 Histogram of the natural, human, and produced capital per capita in 2012

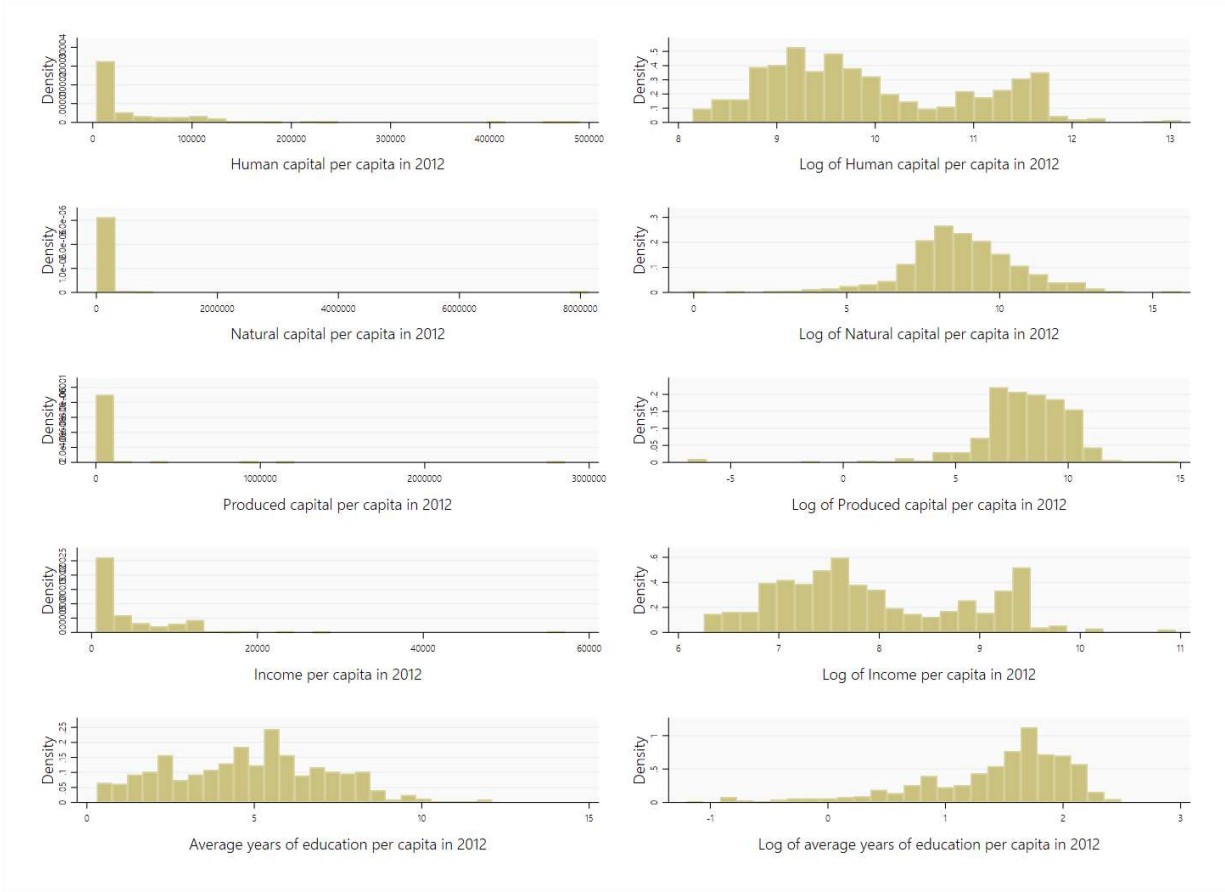




Figure 3-7 Total IWI

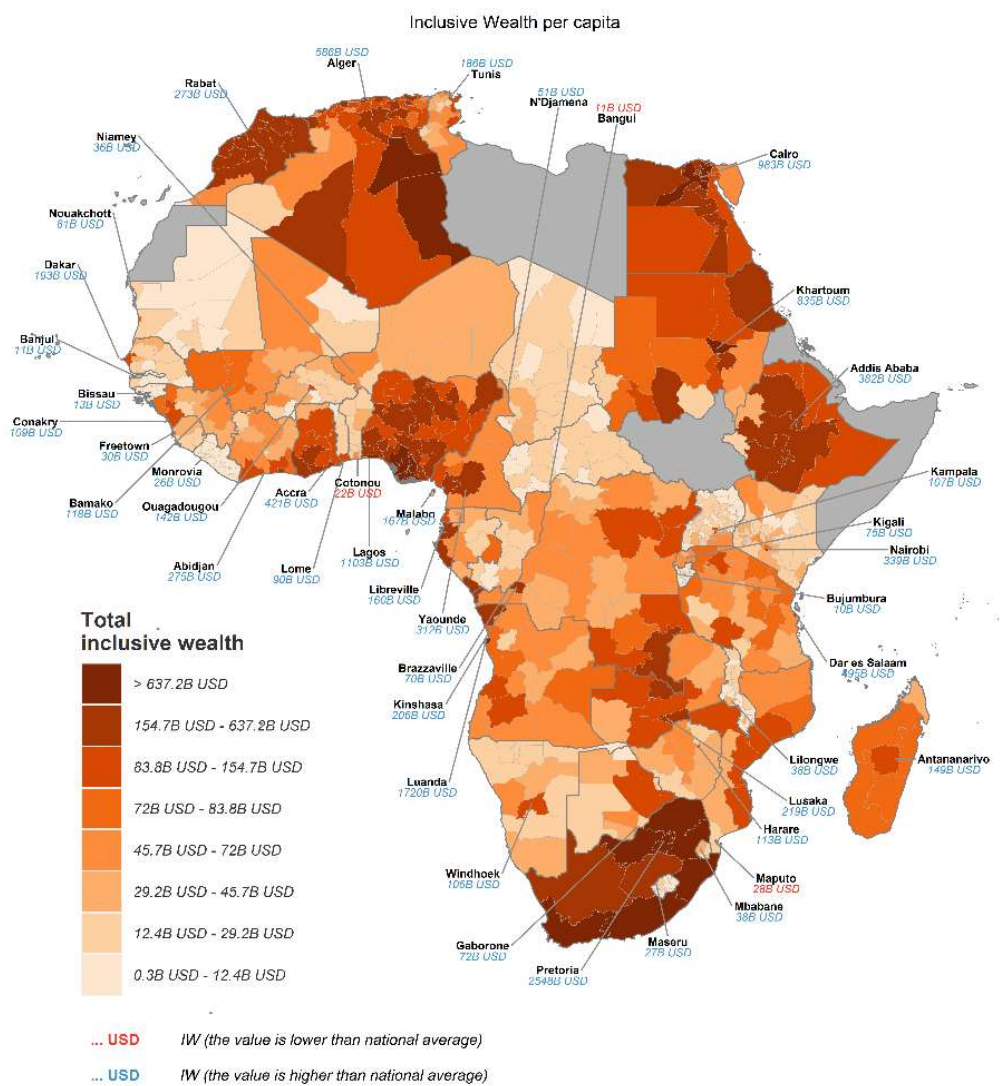


Figure 3-8 Variations of all forms of capital

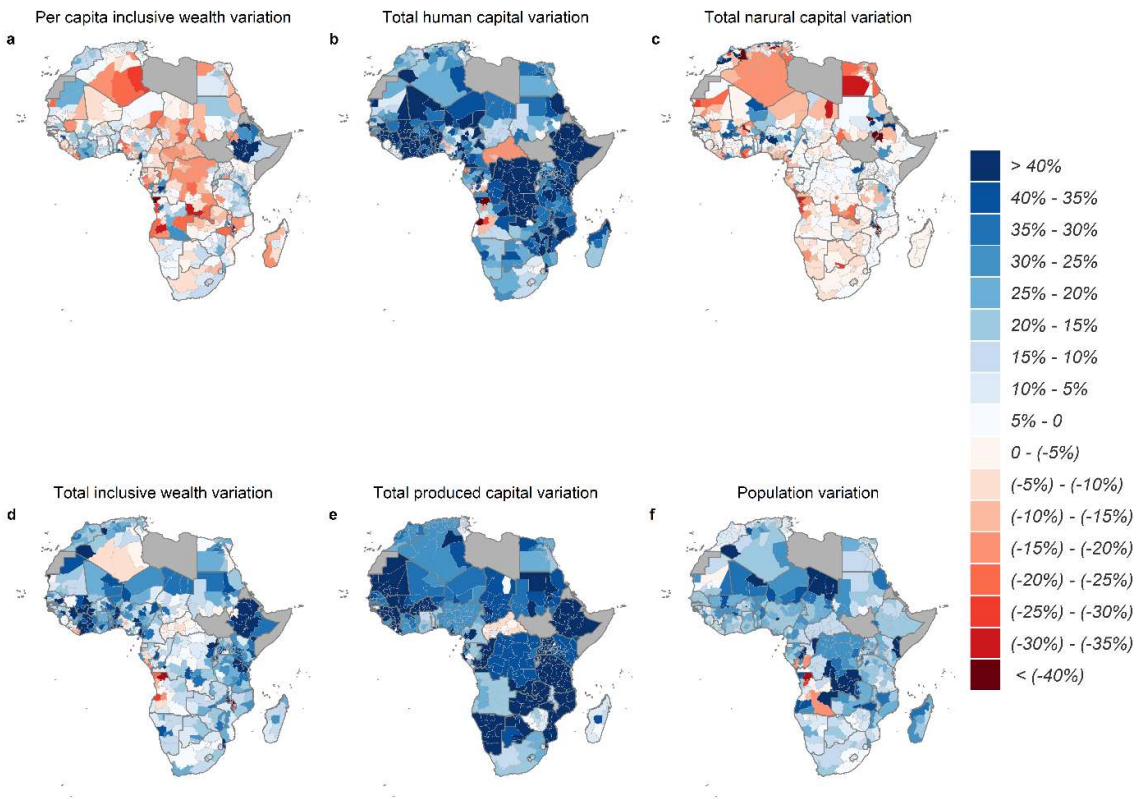


Figure 3-9 Variation in the corpusants of human capital

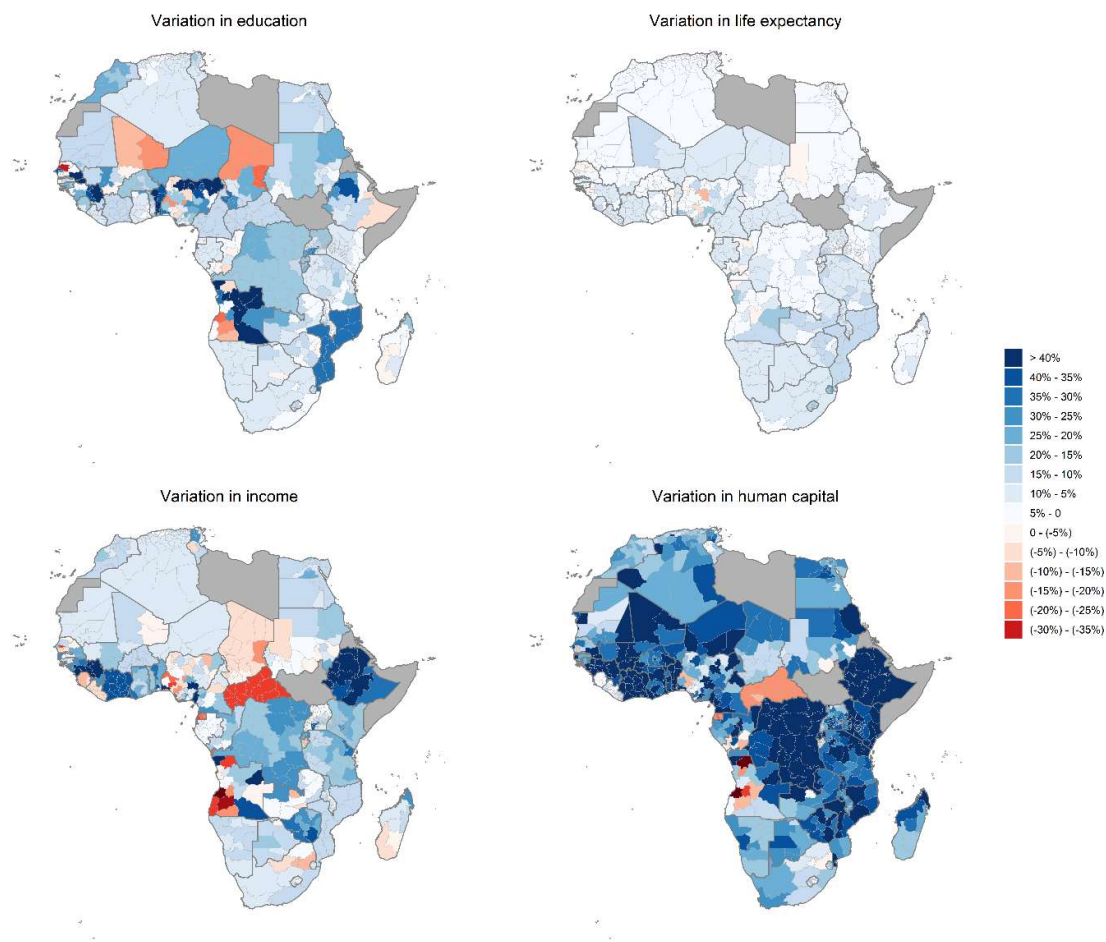


Figure 3-10 Variation in the corposants of produced capitals

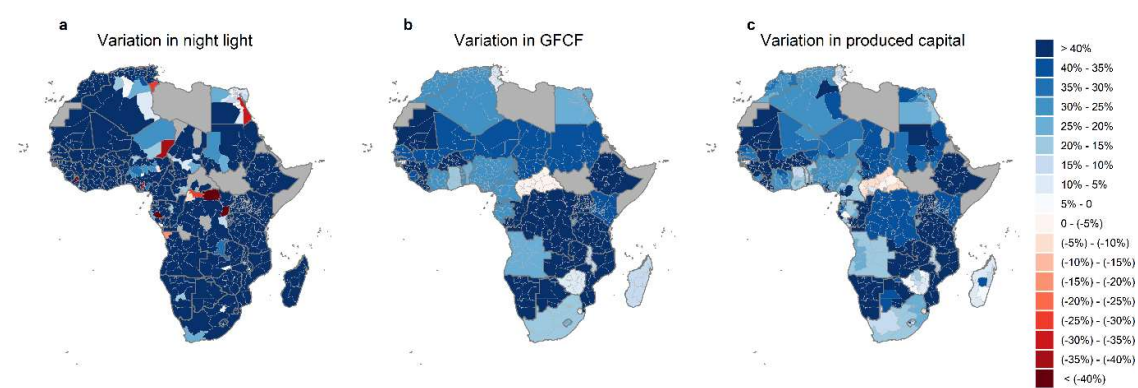


Figure 3-11 Variation in the composants of renewable natural capital

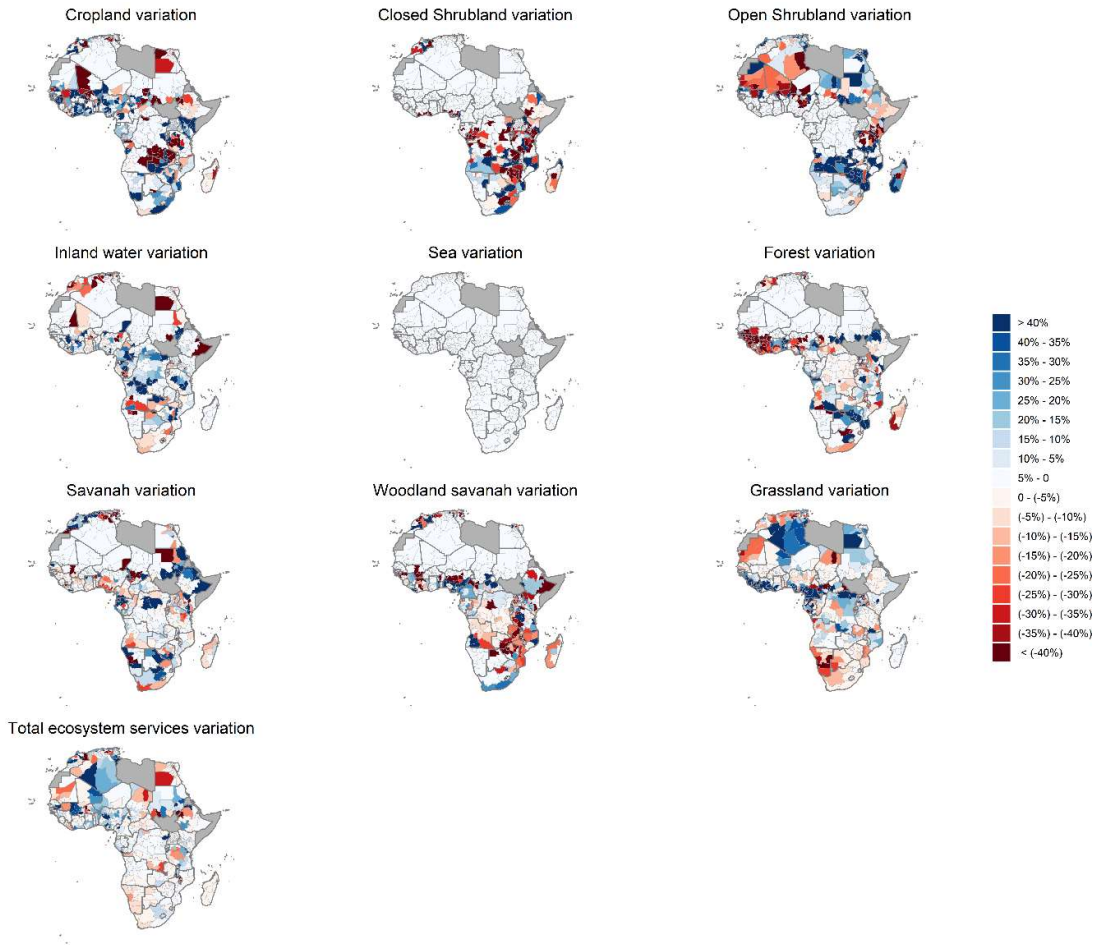




Figure 3-12 Variation in the corposants of non-renewable natural capital

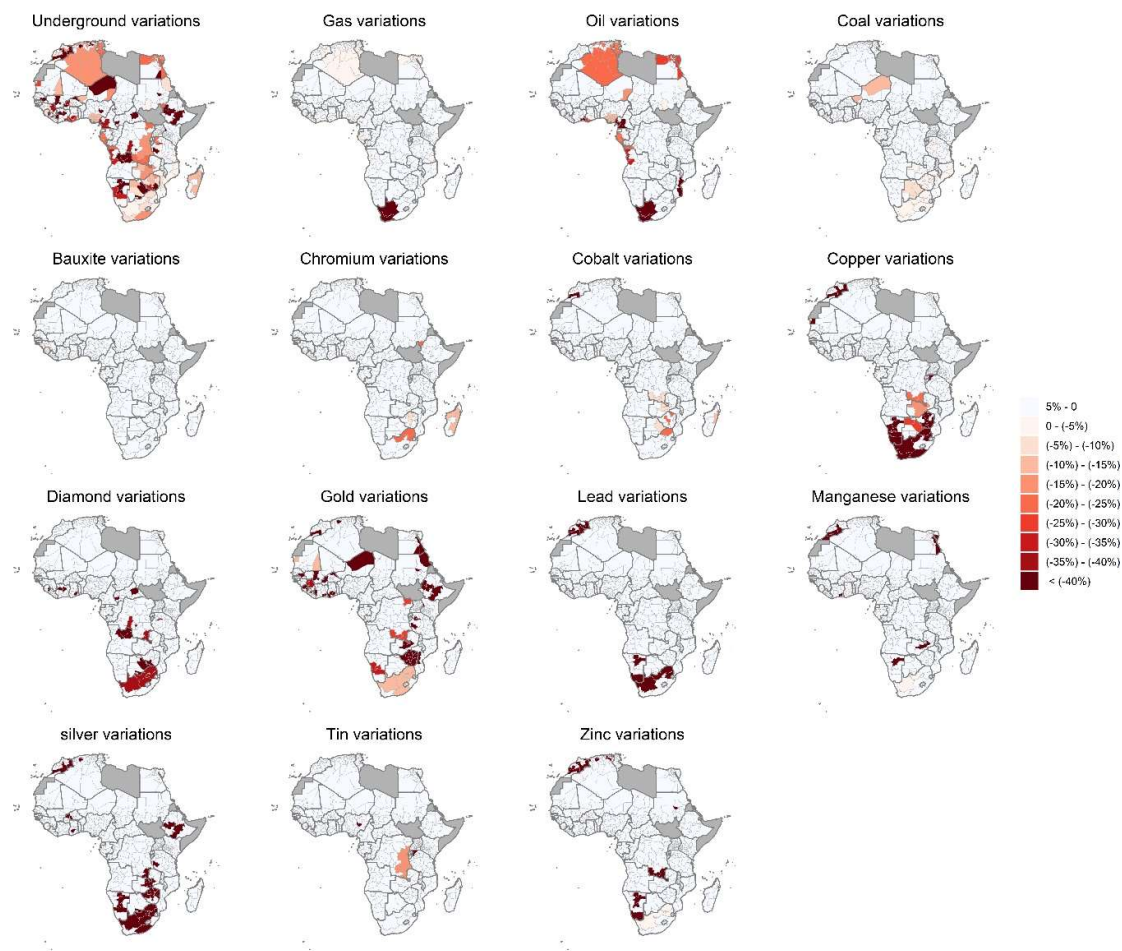


Table 3-8 Difference in income growth between key provinces and secondary provinces

| VARIABLES               | Variation of income per capita |
|-------------------------|--------------------------------|
| Capital dummy           | -0.0273<br>(0.0169)            |
| Constant                | 0.0133***<br>(0.00345)         |
| Observations            | 714                            |
| R-squared               | 0.641                          |
| Country Factor variable | Yes                            |

Robust standard errors in parentheses \*\*\* p<0.01, \*\* p<0.05, \* p<0.1. In this study, income act as a proxy of regional GDP. One of the three main definitions of GDP portray it as the sum of income from labor and from capital.

## **Chapter 4 Populations in slums are happier than rural populations: The case of Mumbai**

### **4.1. Introduction**

Contemporary economic growth promotes the development of more urban civilizations (Easterlin, Angelescu, and Zweig 2011; Glaeser 2011). However, several findings suggest that life in cities may be linked to lower perceptions of well-being, happiness, and life satisfaction, particularly in developed countries (Berry and Okulicz-Kozaryn 2012; Easterlin, Angelescu, and Zweig 2011; Shucksmith et al. 2009; Knight and Gunatilaka 2010; Okulicz-Kozaryn and R. Valente 2021). In developed countries, researchers attribute the lower happiness in large cities to the provisions of public services that reach populations nationally, and to the lifestyle in cities that are linked to factors such as pollution, population density and crime rate (Okulicz-Kozaryn and Mazelis 2018).

Nonetheless, these lower levels of happiness in urban areas relative to rural areas, a concept known as the urban-rural happiness gradient, often does not apply in developing countries. In the developing world, Easterlin, Angelescu, and Zweig (2011) and (Lenzi and Perucca 2021), for instance, show that city-dwellers feel happier than rural populations. These researchers explained the higher happiness in urban areas in developing countries by the absence of basic necessities (for instance, access to water, electricity) in rural areas (Okulicz-Kozaryn and R. Valente 2021), and/or the demographic differences between urban



and rural populations, with rural populations being younger) (Easterlin, Angelescu, and Zweig 2011).

However, by making inferences from a comparison of happiness between urban and rural areas, we miss the important spatial inequalities affecting urban areas in developing countries. These urban areas are largely divided between formal settlements, under the provision of public services, and informal settlements, commonly known as slums, with absence of services such as safe water and sanitation, similar to rural areas. Unlike in developed countries, informal settlements are common in developing countries. The World Bank estimates 60% of urban populations in low-income countries live in informal settlements, against less than 30% in middle income countries in 2018<sup>a</sup>. Considering the absence of provision of public services in slums and rural areas and the harsh living conditions in slums (e.g. population density, pollution), we can posit the hypothesis that the urban-rural gradient of happiness in developing countries exists between populations in informal settlements and rural areas. Yet, as described by (Marx, Stoker, and Suri 2013), the lack of representativeness of data in slums in research on this topic has limited further exploration.

Therefore, this study seeks to fill this knowledge gap by quantifying the difference in subjective wellbeing (SWB) between residents in slums and rural areas in the case study

---

<sup>a</sup> <https://data.worldbank.org/indicator/EN.POP.SLUM.UR.ZS>

of Greater Mumbai municipality. The case of Greater Mumbai provides a pertinent perspective on this topic because the city has a large slum population (the second largest in the world (Marx, Stoker, and Suri 2013)) and Mumbai's rural populations are pushed to move in informal urban areas by economic forces. In other words, owing to its rapidly growing population (Acharya and Nangia 2004) and its expansion over rural areas (Harari 2020), Greater Mumbai is the most populous agglomeration in India (12,691,836 inhabitants in 2021<sup>b</sup>). Since this urban growth outpaces the provision of services, it has led to the creation of informal settlements without access to safe water and sanitation, garbage collection, security of tenure and, poor infrastructure where forty-one percent of Mumbai's population live (Goudet et al. 2018; Gandhi 2012). On the other side, rural populations often move into in these informal settlements as they frequently sell their lands located in Mumbai's peripheral rural areas and build informal settlements as means to increase their income (Sreeja, Madhusoodhanan, and Eldho 2017; Harari 2020).

Hence, I compare the result of a survey about the happiness and satisfaction with life, taken as indices of well-being, between slum-dwellers and rural populations located in Greater Mumbai and compare their answers. This comparison suggests that populations in urban slums have on average a happiness higher than citizens of rural areas by 0.3 points out of 5 and a life satisfaction that is higher by 0.192 out of 5. Attempting to identify the factors

---

<sup>b</sup> <https://www.geonames.org/>

that determine the higher well-being in slums reveals that socio-economic profiles of respondents explain some but not all the difference in happiness and life satisfaction between residents in rural areas and slum-dwellers. However, while population density (Easterlin, Angelescu, and Zweig 2011) and economic development (Proto and Rustichini 2013) (as described by night light data) play an insignificant role in the difference in happiness between populations in slums and those in rural areas, and they explain an important share of the statistical difference in life satisfaction between populations in the two types of area.

Thus, this article relates to three bodies of literature. First, the results can add to the claims that subjective well-being in urban slums is higher than rural areas. Indeed, (Glaeser 2011) grounded this argument on the comparison of happiness between urban and nonurban citizens, and therefore could not ascertain these results for slum-dwellers uniquely.

Furthermore, studies such as (Turok and Borel-Saladin 2018; Chen, Kosec, and Mueller 2019; Perlman 2010) focus on the difference in well-being between slum and rural populations with wellbeing defined as an economic index. Since higher economic well-being does not necessarily lead to higher subjective well-being (Jebb et al. 2018), our results provide a perspective on this argument that reflects on the respondents' own definition of their well-being.

Furthermore, the results add to studies interested in migrations of rural populations to urban areas (Mitra 2010; Chen, Kosec, and Mueller 2019). Studies such as (Mitra 2010) in India and (Chen, Kosec, and Mueller 2019) in Pakistan, show that well-being of these migrants decreases shortly after they move to cities. Here, by estimating the structural difference in happiness between slum and rural populations, the results may suggest that migrants get habituated to temporary shocks that they feel when moving to slums.

Finally, the results relate to studies suggesting that policies against the formation and the expansion of slums may be ineffective (see (Marx, Stoker, and Suri 2013) for survey on this topic). Policies against slums may be relevant in instances like health and security issues. However, the results show that policies aiming to relocate slum-dwellers to rural areas as means to tackle the ills of informal settlements may be conducive to a lower happiness for these populations. Thus, policies aiming at improving life conditions within slums may be more appropriate on a happiness basis.

The remaining of the article is structured as follows. In Section 3, I present the data, and in Section 4, I describe the empirical strategy. Then, I describe the results in Section 5 and discuss the implication of these results in Section 6. In Section 7, I conclude the work.

## 4.2. Literature review

SWB referred to individuals' self-reported well-being. It is usually assessed as an index of happiness, life satisfaction or Cantril's ladder scale that portrays the perception oneself's life. As described by (Diener, Oishi, and Tay 2018), several studies aim to understand the concept of SWB and its characteristics because it has been linked to aspects such as higher work productivity and greater health. In addition, previous studies show that SWB index can provide the ground for policy decision on aspects such as tax rates (Blanchflower et al. 2014) or environmental issues (Coulibaly and Managi 2021). The pertinence of SWB in policy design led researchers to describe spatial differences in SWB. The goal of these efforts is to inform populations and governments of means to improve SWB. Most works carried out in this direction describe the difference in SWB between urban and rural residents.

For instance, (Glaeser 2011)'s seminal work uses several happiness surveys in twenty-five countries where GDP per capita is below \$10,000 to show that urban populations are happier than rural populations. With this result, he argues that populations prefer cities' lifestyle. (Easterlin, Angelescu, and Zweig 2011) provide nuances on this topic with Gallup World Poll's data coming from 80 countries. Although, urban populations have a life satisfaction higher than rural populations in developing countries, rural populations are as satisfied and often more satisfied with their lives than urban populations in developed

countries. (Easterlin, Angelescu, and Zweig 2011) argued that urban populations have a higher SWB in developing countries because of their significantly better occupational structure, education, income, and, access to public services.

Adam Okulicz-Kozaryn and his colleagues advance (Easterlin, Angelescu, and Zweig 2011)'s arguments by estimating a dissatisfaction with city life in developed countries. With the USA taken as a case study, they show that cities with the highest the population are the least happy and, thereby, put forward the concept of urban-rural gradient of happiness with data from General Social Survey (Berry and Okulicz-Kozaryn 2012), Behavioral Risk Factor Surveillance System (Okulicz-Kozaryn and Mazelis 2018), and, the General Social Surveys (Okulicz-Kozaryn 2017). They suggest that urban problems such as population density, pollution, stress, crime rate and lifestyle in cities are among the reasons explaining the lower SWB in cities. Nevertheless, (Okulicz-Kozaryn and R. Valente 2021) agree with (Easterlin, Angelescu, and Zweig 2011) that there is generally no urban-rural gradient of happiness in developing countries.

Works focusing on the difference in happiness between urban and rural areas, however, do not account for differences in living conditions between informal and formal settlements within cities in the developing world. Unlike in developed countries, informal settlements in and around cities, known commonly as “slums”, are widespread in developing countries (Marx, Stoker, and Suri 2013). These agglomerations of housing improper for

habitation mostly occur as the result of governments' inability to provide public services such as access to water, sanitation facilities as fast as cities grow. These slums are characterized by high population density but low levels of income, education, health conditions that is compound with lack of access to safe water and sanitation facilities or pollution (Pörtner and Su 2018; Marx, Stoker, and Suri 2013; Coulibaly et al. 2020). Thus, slums are areas where urban problems, described by (Okulicz-Kozaryn 2017), are concentrated and with population with socio-economic status that may be as low as that of rural populations (Easterlin, Angelescu, and Zweig 2011). Thus, slum dwellers have living conditions similar to that of rural populations coupled with an experience urban problems. Considering that most international surveys do not focus on slums in their sample collection, studies on the difference in happiness between urban and rural areas overlook the fact that life in cities may also be linked to lower SWB in developing countries.

To our knowledge, the only suggestions of difference in well-being between rural and slums populations come from works focusing on well-being of rural migrants and using survey more specific to their study. For instance, (Mitra 2010) suggests a higher well-being in slums over rural areas with data stemming from a survey administered in four Indian slums and with a definition of well-being based on economic, social, cultural and demographic variables. On the other side, (Chen, Kosec, and Mueller 2019) suggests the opposite in Pakistan where although rural-to-urban migrations lead to higher material well-

being it is linked to lower subjective well-being among migrants in their sample. These studies, however, do not inform on the structural difference in SWB between populations in slums and those in rural areas since they estimate the variation in well-being in time of migrants.

Therefore, the study aims to shed light on the difference in well-being between populations in slums and rural areas. For this purpose, I use the outcome of a survey that focus on the identification of slum dwellers and I focus on the case study of Greater Mumbai because more than 50% of its populations live in slums.

### **4.3. Data**

#### *4.3.1 Sample*

For this study, I use the results of a survey on well-being collected by Nikkei Research Inc. (Chapman, Fujii, and Managi 2019). Nikkei Research Inc. that provides the ability to differentiate between populations living in slums and those living in rural areas. This survey used a quota sampling method where after the researchers set goals to collect data for populations' age and gender categories, the respondents were randomly selected to fill these demographics. The survey was conducted between 25 September 2015 and 11 August 2015 in India. It investigated respondents' self-assessment of their well-being, and socio-economic status, the statistics for which are displayed in Table 4-1.



In addition to be a case study relevant for the current topic, I focus on Greater Mumbai because it is the only site where the survey conductors in (Chapman, Fujii, and Managi 2019), identify slums and rural areas. This data is the result of a survey administered as face-to-face interviews to 500 slum-dwellers and 385 rural individuals. The answers from slum dwellers come from 44 settlements presented in Appendix D (Chapman, Fujii, and Managi 2019)’s selection of these settlements in their survey was based on budget constraints and the list of informal settlements described as slums by the Indian government. On that matter, the Indian government defines slums as a housing that is “unfit for human habitation or detrimental to safety, health and morals of the inhabitants” (Nijman 2008).

#### *4.3.2 Definitions of the variables*

Our analysis aims to determine the difference in well-being between rural and slums’ populations. I use subjective well-being (SWB) as the index well-being for this purpose. The SWB can be defined as one’s perception of his/her quality of life and is a common measure of total utility in economics and psychology (Coulibaly and Managi 2021; Kahneman, Wakker, and Sarin 1997).

Following previous studies insights (Diener, Oishi, and Tay 2018; Zhang, Zhang, and Chen 2017), I use two measures of SWB in the analysis, the evaluative measure of SWB and the hedonic measures of SWB. As argued by (Zhang, Zhang, and Chen 2017), understanding

SWB through these concepts can be useful because they relate to factors with different significations. I focus on single indices of SWB for this analysis because they do not prejudge what people will consider a good life for themselves, but instead rely on the judgements based on whatever criteria respondents deem to be most important for their lives (Diener, Oishi, and Tay 2018; Haybron 2008).

The first concept is the hedonic happiness. It refers to the moment-to-moment experienced utility and directly links to immediate emotions and affection (Zhang, Zhang, and Chen 2017). Since the survey was carried out within a month across areas, this index can represent how different populations in slums and rural areas was perceiving their life during this period. I collected the measure of hedonic SWB as the answer to the following question: “Overall, how happy are you with your life?”. The answer of this question, known as the happiness measure, is recorded on a Likert type scale with 1 being the lowest level happiness and 5 the highest.

Then, I use an evaluative measure of SWB which reflects an overall assessment of respondents’ entire lives. This index is appropriate in describing a long-term difference in life assessment that slum-dwellers and rural populations have. This index is recorded as the answer to the question: “All things considered, how satisfied are you with your life?” This index is recorded on a Likert-type scale with 1 being the lowest level of satisfaction and 5 the highest.

Since our interest lies in determining the link between SWB and residency in either a slum or a rural area, I consider the respondents' factors that could influence their SWB in the regression analyses. By respondents' factors, I refer first to respondents' profiles. It is their self-assessed age, participations in the labor force, income, gender, health, education level, marital status, and wealth (as explained by ownership of at least a light bubble, a flushing toilet and a TV). The importance of these individual level factors is highlighted by (Diener, Oishi, and Tay 2018) who show that married and employed individuals are happier on average. Furthermore, they show that while happiness increases with higher income, wealth, health and education level, individuals on mid-age (30 to 40 years old) are less happy than those who are younger and older. Finally, the ownership of a light bubble and a TV enables the accountability of the life where there is an access to electricity and the ownership of a flushing toilet depicts a residency with access to water services.

Respondents' factors also include the total population and economic development of the residency of the respondents. Estimations from (Okulicz-Kozaryn and Mazelis 2018; Mitra 2010) show that population size and economic development contribute to the difference in SWB between urban and rural areas. I use population provided from WorldPop<sup>c</sup> as index of population and I use satellite night light data provided by VIIRS project of

---

<sup>c</sup> <https://www.worldpop.org/>

NASA<sup>d</sup> as a proxy of economic development (see the seminal work of (Henderson et al., 2012) for the relevance of night light satellite data in this procedure).

Summary statistics of the variables and t-test of the difference of these variables between slum dwellers and rural populations are presented in Table 4-1. These statistics show that slum dwellers have on average a higher happiness and life satisfaction than rural populations. However, since the profiles of the respondents of the survey in rural and urban areas exhibit both similarities and divergences, inferences based on simple differences in mean may not accurately inform our hypothesis. Regarding similarities of respondents across areas, 46% of the respondents are female, 70% are married, around 95% own a TV or an electric lamp, and the average age is around 36 and 37 years old. With regards to differences however, slums residents have an income range one class (on a scale of nineteen) higher than that of rural populations. Also, populations in slums are more educated than those in rural areas. 69% of slum dwellers have completed at least a secondary level of education against 50% of rural populations, while 18% of slum populations did not attend school against 30% of rural populations. Finally, respondents in rural areas indicate having a better access to flushing toilet than slum dwellers (14% in rural areas against 5% in informal settlements).

Differences in respondents' profile between urban and rural areas may be responsible of the higher SWB in slum areas. Thus, we need to perform regression analyses that compare

---

<sup>d</sup> <https://earthdata.nasa.gov/earth-observation-data/near-real-time/download-nrt-data/viirs-nrt>

respondents with similar profiles to better inform the differential SWB between respondents in slum and rural areas.

Table 4-1 Sample characteristics

|                                  | <b>Definition</b>                                                                   | <b>Rural</b> | <b>Slum</b> | <b>T-test</b> | <b>S.E.</b> | <b>Obs.</b> |
|----------------------------------|-------------------------------------------------------------------------------------|--------------|-------------|---------------|-------------|-------------|
| Happiness                        | Happiness with life. 1(Unhappy) - 5(Very happy)                                     | 4.426        | 4.683       | -0.257***     | (0.047)     | 830         |
| Life satisfaction                | Satisfaction with life.1 (Completely dissatisfied) - 5(Completely satisfied)        | 4.465        | 4.645       | -0.180***     | (0.046)     | 830         |
| Married                          | Equals 1 if the respondent is married and 0 otherwise                               | 0.701        | 0.701       | -0.000        | (0.032)     | 830         |
| income range                     | Household income range in Rupees. (see ranges in the Appendix)                      | 10.921       | 12.160      | -1.239***     | (0.129)     | 830         |
| Age                              | Respondent's age                                                                    | 36.640       | 37.421      | -0.780        | (0.982)     | 830         |
| Female                           | Equals 1 if the respondent is female and 0 otherwise                                | 0.459        | 0.459       | 0.000         | (0.035)     | 830         |
| Health poor                      | Equals 1 if subjective health is poor and 0 otherwise                               | 0.054        | 0.010       | 0.044***      | (0.013)     | 830         |
| Not attended school              | Equals 1 if the respondent has never attended a school and 0 otherwise              | 0.296        | 0.184       | 0.112***      | (0.031)     | 830         |
| Secondary education              | Equals 1 if the respondent attended at most the secondary education and 0 otherwise | 0.495        | 0.691       | -0.196***     | (0.034)     | 830         |
| Flushing owner                   | toilet Equals 1 if the respondent owns a flushing toilet and 0 otherwise            | 0.142        | 0.052       | 0.090***      | (0.022)     | 830         |
| TV owner                         | Equals 1 if the respondent owns a TV and 0 otherwise                                | 0.952        | 0.974       | -0.022        | (0.014)     | 830         |
| Lamp owner                       | Equals 1 if the respondent owns a lamp and 0 otherwise                              | 0.970        | 0.980       | -0.010        | (0.011)     | 830         |
| Population count<br>(,000 heads) | Average population at the residency of the respondent                               | 33.956       | 35.438      | -1.482***     | (0.443)     | 830         |
| Night emissions                  | light Average luminosity at the residency of the respondent                         | 32.637       | 33.395      | -0.758        | (0.523)     | 830         |

#### 4.4. Empirical approach

To test the hypothesis that the well-being of residents in rural areas is higher than that of residents in slums in Mumbai, I use a spatial error model (SEM). The choice of this model is owed to the ease of interpretation of coefficients of linear regressions and their ability to perform as well as the logistic regressions in well-being related analyses (Diener, Oishi, and Tay 2018). Thus, despite the categorical nature of the well-being indices, I opt for linear regressions of the SEM. Furthermore, I expect the respondents' answers to be influenced by their location because, in the large city of Mumbai, some individuals may live close to each other and thus be affected by the same geographical factors. Hence, the SEM is appropriate in controlling for the spatial auto-correlation of the respondents that live in the same neighbourhood.

The model we regress via a maximum likelihood estimation technique can be expressed as follows:

$$\begin{aligned} SWB_i = & \beta_0 + \beta_1 slum_i + \beta_2 Gender_i + \beta_3 Income_i + \beta_4 Age_i + \beta_5 LaborForce_i + \\ & \beta_6 Health_i + \beta_7 Maritalsatus_i + \beta_8 Education_i + \beta_9 Toilet_i + \beta_{10} TV_i + \beta_{11} Lamp_i + \\ & \beta_{12} Population_{gi} + \beta_{13} Light_{gi} + u_i \end{aligned} \quad (4.1)$$

$$u_i = \sigma W u + \epsilon_i \quad (4.2)$$

In Equation (4.1), where  $SWB_i$  is the subjective well-being of the respondent  $i$ . The term  $slum_i$  is the variable of interest which equals 1 if the respondent  $i$  resides in a slum and

0 if the residency is in a rural area.  $Maritalsatus_i$ ,  $Gender_i$ , and  $Laborforce_i$ , are bivariate dummy variables which equals 1 when the respondent  $i$  is married, a female, and employed respectively.  $Toilet_i$ ,  $TV_i$ , and  $Lamp_i$  are dummy variables which equals 1 when the respondent owns a flushing toilet, a TV, and an electric lamp respectively. The terms  $Education_i$ ,  $Health_i$  and,  $Income_i$  are measured as ordinal variables whose higher values indicate higher levels of education, subjective health and income range respectively.  $Population_{gi}$  and  $Light_{gi}$  are continuous variables indicating the population and light emissions at the grid  $g$  for the respondents  $i$  respectively. Finally,  $Age_i$  is the age of the respondent  $i$ , and the terms  $\beta$  are the coefficients to be estimated.

The sign of  $\beta_1$  that is the coefficient of  $slum_i$  enable to test the hypothesis that populations in rural areas have SWB higher than that of slum-dwellers. A statistically significant negative sign of  $\beta_1$  indicates that rural populations rate their well-being higher than slumdwellers. A statistically significant positive sign will show that rural populations rate their well-being lower than slum-dwellers. Finally, a results that is statistically insignificant suggest a similarity of well-being between populations in rural and slum areas.

Finally, Equation (4.2) displays the residual term  $u_i$  of Equation (4.1). Since Equation (4.1) describes a SEM,  $u_i$  is composed of a component that is idiosyncratic to respondent  $i$ , and a component  $\epsilon_i$  that is spatially correlated in space,  $\sigma W u$ . Regarding the spatially correlated component, I assume  $u$  to be independently and identically distributed across



individuals, and that the spatial correlation across individuals can be represented by a spatial weight matrix  $W$ . Elements of  $W$  are the inverse Euclidean distance between the respondents' residency in Mumbai.  $\sigma$  is the spatial auto-regressive parameter to be estimated.

## 4.5. Results

The results of the regression analyses are displayed in Table 4-2 with each columns representing a different model estimation. The coefficients in columns 1 to 3 refer to correlations of the independent variables with hedonic measure of SWB since the dependent variable is happiness. In the columns 4 to 6, I show the coefficients related to the evaluative measure of SWB with life satisfaction used as the dependent variable. The coefficient of interest, informing on the differential of SWB between slum dwellers and rural populations, is displayed in the first line of the coefficient estimates. Moving from columns 1 to 3 and from columns 4 to 6, this coefficient describes the slum-rural difference in SWB when increasing numbers of control variables are added to the model. In the columns 1 and 4, the regressions only control out the spatial auto-correlation across individuals. In the columns 2 and 5, I added to the control variables relative to the respondents' profile. Finally, I included population density and night light emissions at the grid level in columns 3 and 6. All model specifications take into account the spatial auto-correlation across observations and display the auto-regressive parameter of the SEM.

Starting with results related to the hedonic measure of SWB, I find that slum dwellers report on average a level of happiness that is 0.3 points (on a scale of 5 point) statistically higher than rural populations do ( $p < 0.01$ , Column 1 of Table 4-2). However, when controlling out the respondents' profile, this difference in happiness decreases to 0.164 points but still remains statistically significant ( $p = 0.02$ , Column 2 of Table 4-2). This result suggest that individual level characteristics explain some of the difference in happiness between individuals in slums and those in rural areas. Then, when controlling out respondent's residency factors (population and night light) in addition to their profiles, I find that slum dwellers report a SWB of 0.216 points higher than rural population ( $p < 0.01$ , Column 3 of Table 4-2).

On the other hand, the results of the analyses with the evaluative measure of SWB as the dependent variable generally fit the same report as that of results with the hedonic measure of SWB. On average, I find that slum dwellers expressed a level of life satisfaction that is

0.192 points (on a scale of 5 points) higher than that of rural populations ( $p = 0.01$ , Column 4 of Table 4-2). When I control out the characteristics relative to individual profiles, this difference decreases to 0.192 points but remain statically significant ( $p = 0.04$ , Column 5 of Table 4-2). However, there is no statistically significant difference in SWB between rural and urban after taking into account the levels of populations and night light emissions

of the residency of respondents. This last result suggests that differences in life satisfaction across areas is linked to differences in populations and economic activities. Therefore, on the basis of the statistical analysis I tend to reject the null hypothesis that rural populations are happier than slum dwellers.

Regarding the control variables, the regression analyses suggest that individuals with high income and good health are significantly happier than other individuals. This result fits previous studies' insights (Diener, Oishi, and Tay 2018). However, other control variables do not consistently display a statistically significant relationship with SWB, and therefore I refrain to make any inference from their coefficients. Furthermore, I find the presence of a statistically significant and positive spatial auto-correlation across observations. It means that the answers from respondents that live close to each other are similar.

Table 4-2 Differences in well-being across areas in Mumbai

| Dependent variable | Happiness           |                     |                     | Life satisfaction  |                     |                     |
|--------------------|---------------------|---------------------|---------------------|--------------------|---------------------|---------------------|
|                    | (1)                 | (2)                 | (3)                 | (4)                | (5)                 | (6)                 |
| Slum               | 0.300***<br>(0.071) | 0.164**<br>(0.075)  | 0.216***<br>(0.070) | 0.192**<br>(0.078) | 0.145**<br>(0.070)  | 0.092<br>(0.079)    |
| Married            |                     | 0.047<br>(0.056)    | 0.040<br>(0.059)    |                    | 0.038<br>(0.059)    | 0.032<br>(0.058)    |
| Income range       |                     | 0.047***<br>(0.013) | 0.050***<br>(0.013) |                    | 0.040***<br>(0.013) | 0.035***<br>(0.013) |
| Age                |                     | -0.017*<br>(0.010)  | -0.013<br>(0.010)   |                    | -0.011<br>(0.011)   | -0.011<br>(0.010)   |
| Age <sup>2</sup>   |                     | 0.000               | 0.000               |                    | 0.000               | 0.000               |

|                               |           |           |          |           |           |          |
|-------------------------------|-----------|-----------|----------|-----------|-----------|----------|
|                               | (0.000)   | (0.000)   |          | (0.000)   | (0.000)   |          |
| Female                        | 0.034     | 0.012     |          | 0.011     | 0.000     |          |
|                               | (0.044)   | (0.045)   |          | (0.045)   | (0.045)   |          |
| Health poor                   | -0.874*** | -0.867*** |          | -0.935*** | -0.926*** |          |
|                               | (0.124)   | (0.131)   |          | (0.132)   | (0.129)   |          |
| Not attended school           | -0.040    | -0.004    |          | -0.112    | -0.131*   |          |
|                               | (0.067)   | (0.070)   |          | (0.070)   | (0.069)   |          |
| Secondary education           | 0.074     | 0.083     |          | -0.097    | -0.106    |          |
|                               | (0.066)   | (0.068)   |          | (0.068)   | (0.068)   |          |
| Flushing toilet owner         | 0.141*    | 0.060     |          | 0.075     | 0.129     |          |
|                               | (0.077)   | (0.079)   |          | (0.080)   | (0.080)   |          |
| TV owner                      | -0.074    | -0.081    |          | -0.011    | -0.015    |          |
|                               | (0.109)   | (0.115)   |          | (0.116)   | (0.113)   |          |
| Lamp owner                    | 0.167     | 0.160     |          | 0.328**   | 0.297**   |          |
|                               | (0.131)   | (0.138)   |          | (0.138)   | (0.136)   |          |
| Population count (,000 heads) |           | 0.001     |          |           | 0.003     |          |
|                               |           | (0.005)   |          |           | (0.006)   |          |
| Night light emissions         |           | -0.012**  |          |           | -0.007    |          |
|                               |           | (0.005)   |          |           | (0.005)   |          |
| Constant                      | 4.403***  | 4.114***  | 4.320*** | 4.475***  | 3.990***  | 4.304*** |
|                               | (0.061)   | (0.287)   | (0.349)  | (0.067)   | (0.297)   | (0.370)  |
| W                             |           |           |          |           |           |          |
| $\sigma$ (Happiness)          | 0.956***  | 2.837***  | 0.956*** |           |           |          |
|                               | (0.042)   | (0.129)   | (0.042)  |           |           |          |
| $\sigma$ (Life satisfaction)  |           |           |          | 1.219***  | 0.940***  | 1.715*** |
|                               |           |           |          | (0.062)   | (0.056)   | (0.095)  |
| Observations                  | 830       | 830       | 830      | 830       | 830       | 830      |
| Pseudo $R^2$                  | 0.035     | 0.116     | 0.126    | 0.018     | 0.104     | 0.102    |

Standard errors in parentheses

\*  $p < 0.10$ , \*\*  $p < 0.05$ , \*\*\*  $p < 0.01$

## 4.6. Discussion

In summary, the results show that respondents in Mumbai's slum report being happier than respondents in rural areas, unconditionally to their socio-economic profiles and the total population and economic activities in their residency. Next, despite slum dwellers being more satisfied with their lives on average, their satisfactions are similar to that of individuals residing in rural areas where the level of population and light emissions are similar to slums.

With regards to the urban economics theory, these results are supported by previous studies addressing the absence of urban-rural gradient of happiness in developing countries (Lenzi and Perucca 2021; Easterlin, Angelescu, and Zweig 2011; Glaeser 2011; Turok and Borel-Saladin 2018; Chen, Kosec, and Mueller 2019; Perlman 2010). The absence of urban-rural gradient of happiness in developing countries refers to the higher levels of well-being linked to lifestyle in urban areas relative to that in rural areas (see (Diener, Oishi, and Tay 2018; Glaeser 2011; Okulicz-Kozaryn and Mazelis 2018)). Like in (Glaeser 2011)'s arguments, we can posit that residents in slums are happier than rural populations because of their proximity to formal urban areas. The proximity to wealthy formal urban areas enables slum dwellers to near higher socio-economic lifestyles, and higher opportunity of upward social mobility even when slum dwellers are not rich or do not have access to certain public services themselves. Consequently, these results suggest that the effects of relative income,

leading populations to report low SWB when they live close to individuals wealthier than they are (Luttmer 2005), has a minor importance in the differential of well-being between slum dwellers and rural populations. Indeed, compared to rural populations, slum dwellers report higher happiness and life satisfaction despite the fact that they live close to wealthy populations in formal urban areas.

Our results can provide other insights through the lens of general economic theories. The interpretation of SWB as a proxy of total utility informs that the results suggest that populations in rural areas will migrate to slum areas in the long-term (Kahneman, Wakker, and Sarin 1997). In economics, one assumes that individuals aim to maximize their total utility through their decisions like the choice of their residency. Thus, a higher happiness in slums means that rural populations will be willing to migrate toward slums. Although previous studies show that in the short-term, migrations from rural areas to slums are linked to a decrease in reported happiness (Chen, Kosec, and Mueller 2019), I believe that the structural difference in well-being between slum and rural areas will continue to push rural populations towards slums. This explanation originates from the hedonic adaptation hypothesis. This hypothesis states that there is a psychological process that attenuates the long-term emotional impact of a favourable or unfavourable change in circumstances (Galiani, Gertler, and Undurraga 2018). Despite this perspective, we owe to note that local economic development in rural areas may slow down migrations from rural to slum areas.

Individuals in rural areas that live in areas with total population and light emissions similar to that in slums report a satisfaction with their lives similar to that of slum dwellers. Thus, they will be less tempted to migrate to slums to obtain a higher satisfaction with their lives.

Finally, with regards to Mumbai's urban policy, the results show that policy that seek to displace slum dwellers toward rural areas are not conducive on well-being terms. Mumbai has experienced a long history of policy dealing slums that are summarized in (Bardhan et al. 2015)'s work. Most of these governmental efforts were unsuccessful in eradicating or substantially improving lives of slum dwellers. According to (Bardhan et al. 2015), this failure reflects the lack of bottom-up approaches in the creation process of policy where slum dwellers opinions are considered. In this context, the result can first explain the failure of the Slum Area Improvement and Clearance Act, 1956. During the enforcement of this act, the clearance of informal settlements led most slum dwellers to relocate in other slums in Mumbai. Our results, exhibiting a higher SWB in slums relative to rural areas, explain why slum-dwellers refrained from relocating in rural areas and instead remained in slums within Mumbai. Second, the result can aid in policy creation as it suggests that to incorporate a bottom-up approach in the creation of policies dealing with slums, improving economic activities in rural areas may have some impact the migration of slum-dwellers towards rural areas as life satisfaction between the two areas will be reduced.

## **4.7. Conclusions**

Based on evidence of harsh living conditions in slums, the study investigated the difference in well-being between samples of resident in slums and rural areas in Greater Mumbai. I found that respondents in slums report being happier and more satisfied with their lives than respondents in rural areas, irrespective to their socio-economic' profiles. Furthermore, slum dwellers report being happier even when they are compared with rural individuals who live in areas where same populations and economic development resemble that of slums.

The issues raised in this study, and the limitations of our approaches, give rise to many research opportunities. First, the study focuses on the difference in SWB between residents in informal settlements and residents in rural areas. It disregards happiness of residents in formal urban areas. To provide a more general information to the knowledge in urban economics' literature, there is a need to understand the difference in SWB between formal urban areas, informal urban and rural areas.

Second, the scope of the study is defined around the estimation of the higher SWB of slum dwellers, and thereby it does not account slum dweller's opinions in the changes needed in slum's rehabilitation. Although slums may be appropriate for some of its residents, in the general urban settings, they cause issues linked to health, sanitation, and crimes for



government. Thus, determining with slums the conducive approach for them and for the government to rehabilitate their living standards will be necessary.

## References

- Acharya, Arun Kumar, and Parveen Nangia. 2004. "Population Growth and Changing Land-Use Pattern in Mumbai Metropolitan Region of India." *Caminhos de Geografia* 5 (11).
- Bardhan, Ronita, Sayantani Sarkar, Arnab Jana, and Nagendra R. Velaga. 2015. "Mumbai Slums since Independence: Evaluating the Policy Outcomes." *Habitat International* 50 (December): 1–11. <https://doi.org/10.1016/j.habitatint.2015.07.009>.
- Berry, Brian J.L., and Adam Okulicz-Kozaryn. 2012. "The City Size Distribution Debate: Resolution for US Urban Regions and Megalopolitan Areas." *Cities* 29 (March): S17–23. <https://doi.org/10.1016/j.cities.2011.11.007>.
- Blanchflower, David G., David N.F. Bell, Alberto Montagnoli, and Mirko Moro. 2014. "The Happiness Trade-Off between Unemployment and Inflation." *Journal of Money, Credit and Banking* 46 (S2): 117–41. <https://doi.org/10.1111/jmcb.12154>.
- Chapman, Andrew, Hidemichi Fujii, and Shunsuke Managi. 2019. "Multinational Life Satisfaction, Perceived Inequality and Energy Affordability." *Nature Sustainability* 2 (6): 506–14. <https://doi.org/10.1038/s41893-019-0303-5>.
- Chen, Joyce, Katrina Kosec, and Valerie Mueller. 2019. "Moving to Despair? Migration and Well-Being in Pakistan." *World Development* 113 (January): 186–203. <https://doi.org/10.1016/j.worlddev.2018.09.007>.
- Coulibaly, Thierry Yerema, and Shunsuke Managi. 2021. "The Multinational and Heterogeneous Burden of Air Pollution on Well-Being." *Journal of Cleaner Production* 318 (October): 128530. <https://doi.org/10.1016/j.jclepro.2021.128530>.
- Coulibaly, Thierry Yerema, Mihoko Wakamatsu, Moinul Islam, Fukai Hiroki, Shunsuke Managi, and Bingqi Zhang. 2020. "Differences in Water Policy Efficacy across South African Water Management Areas." *Ecological Economics* 175: 106707. <https://doi.org/10.1016/j.ecolecon.2020.106707>.
- Diener, Ed, Shigehiro Oishi, and Louis Tay. 2018. "Advances in Subjective Well-Being Research." *Nature Human Behaviour* 2 (4): 253–60. <https://doi.org/10.1038/s41562-018-0307-6>.
- Easterlin, Richard A., Laura Angelescu, and Jacqueline S. Zweig. 2011. "The Impact of Modern Economic Growth on Urban–Rural Differences in Subjective Well-Being." *World Development* 39 (12): 2187–98. <https://doi.org/10.1016/j.worlddev.2011.04.015>.
- Galiani, Sebastian, Paul J Gertler, and Raimundo Undurraga. 2018. "The Half-Life of Happiness: Hedonic Adaptation in the Subjective Well-Being of Poor Slum Dwellers to

- the Satisfaction of Basic Housing Needs.” *Journal of the European Economic Association* 16 (4): 1189–1233. <https://doi.org/10.1093/jeea/jvx042>.
- Gandhi, Sahil. 2012. “Economics of Affordable Housing in Indian Cities.” *Environment and Urbanization ASIA* 3 (1): 221–35. <https://doi.org/10.1177/097542531200300112>.
- Glaeser, Edward. 2011. *Triumph of the City: How Urban Spaces Make Us Human*. Pan Macmillan.
- Goudet, S., A. Jayaraman, S. Chanani, D. Osrin, B. Devleesschauwer, B. Bogin, N. Madise, and P. Griffiths. 2018. “Cost Effectiveness of a Community Based Prevention and Treatment of Acute Malnutrition Programme in Mumbai Slums, India.” Edited by Vijayaprasad Gopichandran. *PLOS ONE* 13 (11): e0205688. <https://doi.org/10.1371/journal.pone.0205688>.
- Harari, Mariaflavia. 2020. “Cities in Bad Shape: Urban Geometry in India.” *American Economic Review* 110 (8): 2377–2421. <https://doi.org/10.1257/aer.20171673>.
- Haybron, Daniel M. 2008. “Philosophy and the Science of Subjective Well-Being.” In *The Science of Subjective Well-Being*, edited by Eid Michael and Larsen J. Randy, 17–43. New York: the Guilford Press.
- Jebb, Andrew T, Louis Tay, Ed Diener, and Shigehiro Oishi. 2018. “Happiness , Income Satiation and Turning Points around the World.” *Nature Human Behaviour* 2 (January). <https://doi.org/10.1038/s41562-017-0277-0>.
- Kahneman, Daniel, Peter P. Wakker, and Rakesh Sarin. 1997. “Back to Bentham? Explorations of Experienced Utility.” *Quarterly Journal of Economics* 112 (2): 375–405. <https://doi.org/10.1162/003355397555235>.
- Knight, John, and Ramani Gunatilaka. 2010. “The Rural-Urban Divide in China: Income but Not Happiness?” *Journal of Development Studies* 46 (3): 506–34. <https://doi.org/10.1080/00220380903012763>.
- Lenzi, Camilla, and Giovanni Perucca. 2021. “Not Too Close, Not Too Far: Urbanisation and Life Satisfaction along the Urban Hierarchy.” *Urban Studies* 58 (13): 2742–57. <https://doi.org/10.1177/0042098020962397>.
- Luttmer, E. F. P. 2005. “Neighbors as Negatives: Relative Earnings and Well-Being.” *The Quarterly Journal of Economics* 120 (3): 963–1002. <https://doi.org/10.1093/qje/120.3.963>.
- Marx, Benjamin, Thomas Stoker, and Tavneet Suri. 2013. “The Economics of Slums in the Developing World.” *Journal of Economic Perspectives* 27 (4): 187–210.

<https://doi.org/10.1257/jep.27.4.187>.

- Mitra, Arup. 2010. "Migration, Livelihood and Well-Being: Evidence from Indian City Slums." *Urban Studies* 47 (7): 1371–90. <https://doi.org/10.1177/0042098009353621>.
- Nijman, Jan. 2008. "Against the Odds: Slum Rehabilitation in Neoliberal Mumbai." *Cities* 25 (2): 73–85. <https://doi.org/10.1016/j.cities.2008.01.003>.
- Okulicz-Kozaryn, Adam. 2017. "Unhappy Metropolis (When American City Is Too Big)." *Cities* 61 (January): 144–55. <https://doi.org/10.1016/j.cities.2016.04.011>.
- Okulicz-Kozaryn, Adam, and Joan Maya Mazelis. 2018. "Urbanism and Happiness: A Test of Wirth's Theory of Urban Life." *Urban Studies* 55 (2): 349–64. <https://doi.org/10.1177/0042098016645470>.
- Okulicz-Kozaryn, Adam, and Rubia R. Valente. 2021. "Urban Unhappiness Is Common." *Cities*, August, 103368. <https://doi.org/10.1016/j.cities.2021.103368>.
- Perlman, Janice E. 2010. *Parsing the Urban Poverty Puzzle: A Multi-Generational Panel Study in Rio de Janeiro's Favelas, 1968-2008*. WIDER Working Paper.
- Pörtner, Claus C., and Yu hsuan Su. 2018. "Differences in Child Health Across Rural, Urban, and Slum Areas: Evidence From India." *Demography* 55 (1): 223–47. <https://doi.org/10.1007/s13524-017-0634-7>.
- Proto, Eugenio, and Aldo Rustichini. 2013. "A Reassessment of the Relationship between GDP and Life Satisfaction." Edited by Mariano Sigman. *PLoS ONE* 8 (11): e79358. <https://doi.org/10.1371/journal.pone.0079358>.
- Shucksmith, Mark, Stuart Cameron, Tanya Merridew, and Florian Pichler. 2009. "Urban–Rural Differences in Quality of Life across the European Union." *Regional Studies* 43 (10): 1275–89. <https://doi.org/10.1080/00343400802378750>.
- Sreeja, K.G., C.G. Madhusoodhanan, and T.I. Eldho. 2017. "Processes of Peri-Urban Resource – Livelihood Transitions: Glimpses from the Periphery of Greater Mumbai City, India." *Land Use Policy* 69 (December): 49–55. <https://doi.org/10.1016/j.landusepol.2017.09.008>.
- Turok, Ivan, and Jackie Borel-Saladin. 2018. "The Theory and Reality of Urban Slums: Pathways-out-of-Poverty or Cul-de-Sacs?" *Urban Studies* 55 (4): 767–89. <https://doi.org/10.1177/0042098016671109>.
- Zhang, Xin, Xiaobo Zhang, and Xi Chen. 2017. "Happiness in the Air: How Does a Dirty Sky Affect Mental Health and Subjective Well-Being?" *Journal of Environmental Economics and Management* 85 (September): 81–94.

<https://doi.org/10.1016/j.jeem.2017.04.001>.

## Appendix 4

| <b>List of slums</b>                 |
|--------------------------------------|
| Ghatla Village Chembur               |
| Lebur Camp Matunga                   |
| Pratiksha Nagar Sion Koliwada        |
| Sardar Nagar GTB Nagar               |
| Sardar Nagar 2 Sion Chunabhatti Sion |
| Sardar Nagar 4 Sion                  |
| Shivaji Nagar Govandi                |
| Nityanand Nagar Ghatkoper(w)         |
| Tanaji Chowk Kurla (w)               |
| Jaiydeep Colony Nahur                |
| Ambedkar Nagar Vidhyavihar           |
| Sai Nagar Bhandup(e)                 |
| Kokan Nagar Bhandup(w)               |
| Karve Nagar Kanjur Marg (e)          |
| Maharashtra Nagar Kanjur Marg(e)     |
| BDD Chawl Worli                      |
| BDD Chawl Naigaon Dadar(e)           |
| BIT Chawl Lower Parel                |
| BIT Chawl Mumbai Central             |
| Shastri Nagar Mahalaxmi              |
| Umarkhadi Sandhurst Road             |
| BIT Chawl Byculla(w)                 |
| BIT Chawl Love Lane Byculla(e)       |
| MHB Colony Dahisar                   |
| Quarter Road No-3 Boriwali (e)       |
| Malad Kurar Village(e)               |
| Rani Sati Marg Malad(e)              |
| Jawahar Nagar Goregaon(w)            |
| Bandrekar Wadi Jogeshwari(e)         |
| Pratap Nagar Jogeshwari(e)           |
| Motilal Nagar Goregaon(w)            |
| Shradhanand road Ville Parle         |
| Neharu Nagar Ville Parle(w)          |

|                              |
|------------------------------|
| Shahu Nagar Mahim(e)         |
| Telli Galli Andheri(e)       |
| Koldongri Andheri(e)         |
| Gundavali Gavthan Andheri(e) |
| Beharam pada Bandra (e)      |
| Government Colony Bandra(e)  |

## **Chapter 5 Using satellite images to assess the economic impact of the World Cup at the city level**

### **5.1. Introduction**

Economic indices averaged at the national level cannot accurately portray the impact of mega-events (Baade & Matheson, 2004; Feddersen & Maennig, 2012). In South Africa, despite the massive investment, neither the national GDP nor even their provincial GDP seemed to have changed as a result of the World Cup 2010 (see GDP trends in Figure 5.5 and 5.5 in the Appendix). Notwithstanding, South Africa and several developing countries (i.e. Brazil, Russia and Qatar) perceive this mega-event to be an opportunity to boost their economy and thus increasingly bid to host the event (Pop, Kanovici, Ghic, & Andrei, 2016). Hence, relying on the arguments developed by Baade and Matheson (2004) and Feddersen and Maennig (2012), it is necessary to measure the impact of the World Cup at such a level to perceive its economic effects.

However, unlike developed countries, South Africa has neither periodic nor reliable city-level economic indices to assess the World Cup's economic impacts (Cavalcanti, Limont, Dziedzic, & Fernandes, 2017). This lack of data contributes to differences of perception on the economic effects of the World Cup in South African host cities. Indeed, host cities benefited from new infrastructure intended to improve their economy and to alleviate poverty by the implementation of more efficient transportation systems, sports activities, boost in tourism



industry in host cities. Certain investigations such as Peeters et al. (2014) with international tourism inflow data and Pfeifer et al. (2018) with night light data suggest that the goals linked to the boost of the tourism industry and the improvement of the local economy was achieved respectively. However, studies such as Easmonn and Trish (2015) who conducted post-event interviews with event officials and Steinbrink et al (2011) who compiled a collection of studies on the impact of the World Cup claimed that investment for the event were inefficient. These divergences in the assessments of the impact of the World Cup 2010 contrast with the commonly recognized evaluation among scholars of editions of the World Cup held in developed countries. This body of research reveals an absence of any economic advantage in hosting the World Cup at the city level in developed countries, with findings supported by the use of more fine-scale, commonplace and reliable economic indices (Baade & Matheson, 2004; Coates & Humphreys, 2002; Feddersen & Maennig, 2012; Siegfried & Zimbalist, 2010).

Therefore, this study addresses the issue of data unavailability at the city level in the estimation of the economic effect linked to hosting the World Cup 2010 in host cities. In South Africa, the total income at the city level is the most disaggregated economic data. Since it is available only for 2007, it lacks periodicity to investigate the effect of the World Cup. Thus, I rely on a body of literature in economics arguing that satellite night light data are relevant proxies of economic activities and use it to predict income growth at the city level in South Africa from 1992 to 2013 (Henderson et al., 2012; Michalopoulos & Papaioannou, 2018). The

reason underlying the relationship between night light and economic development is the need for luminosity for most human nocturnal activities. However, the quantity of light emitted at the macro-level depends on the level of development of citizens in the areas and the capacity of government to provide light (i.e., electricity). The pertinence of luminosity to predict income is salient in (Pokhriyal & Jacques, 2017)'s analyses. They highlight that out of 120 factors including mobile phone records, satellite imageries, or characteristics of local markets, all of which are recognized in economics literature to impact income, lack of night lights was the most important feature in predicting poverty (defined as the lack of income) in Senegal at the commune level.

Hence, to predict income from 1992 to 2013, I use the coefficients of a geographically weighted regression (GWR) between light and income in 2007. Then, I create a yearly proxy of income per municipality by multiplying the GWR coefficients with the night light data for each year where the data are available. I use the GWR in the investigation to address two measurement errors that night light data may display in studies like ours. Hunter et al. (2017) highlight that places with large agglomerations of bright lights may experience top-coding that censors part of their increase in light. Meanwhile, Michalopoulos et al. (2018) suggest that the nature of light in every spatial unit is often specific to this spatial unit. Thus, assuming that a unit in variation in light represents a proportional unit in economic growth regardless the location where the increase of light occurs may understate the lucrativeness of economic

activities in affluent cities and overstate it elsewhere. Hence, the GWR enables us to leverage the interdependence of cities' economies across space, so-called agglomeration economies (Antonelli, Patrucco, & Quatraro, 2011) so that income growth linked to light growth for each city matches the geographical cluster of wealth and light where this city is located. Hence, the construction of this proxy enables robust assessment of the effect of the World Cup by comparing income growth between host cities and non-host cities.

To carry out this comparison, I test the hypothesis that income growth in host cities was not significantly different than that of other South African cities. Our results show that despite the investment they received, host cities had an income growth similar to that of other South African cities. However, host cities that invested the most in the construction and the renovations of their stadiums had an income growth that 9% lower than that of other host cities after the event.

Overall, while this study focus on the World Cup 2010, the method can provide insights in research corpus seeking to evaluate policies at the sub-national level with either no or not reliable data (Barclay, 2009; Wray, Todes, Gotz, & Harrisson, 2014). In their survey describing the use of night light in economics, Michalopoulos et al. (2018) highlights that night light data have some shortcomings linked to their hihgly correlated spatial nature. Here, with a spatial econometrics' approach, I show that variations in night light can better explain economic

variations in cross-sectional and longitudinal analyses when the variations in night light incorporate concepts of spatial economic agglomeration.

Furthermore, Pfeifer et al. (2018) already use night light data for the economic valuation of the World Cup 2010 and find an income growth in South Africa host cities from 2004 to 2008. This analysis extends this previous attempt by accounting for changes in land cover and population satellite maps and by doing so contradict the previous interpretations. Thus, the results show the importance of omitted variables when using night light satellite data for economic analyses (Michalopoulos & Papaioannou, 2018).

Finally, the findings can provide insights in the organization of mega-events in developing countries. In case of South Africa, there is a need for investment that spur both the formal and the informal sectors and to divert away from the construction of costly stadiums to benefit from the event.

This study proceeds as follows. Section 2 reviews the body of literature investigating the impact of the World Cup in South African host cities. Section 3 describes the dataset, Section 4 describes the methodology, and Section 5 describes the results. Finally, Sections 6 and 7 present the discussion and conclusions.

## **5.2. The World Cup and the economy of South African host cities**

Since the end of the Apartheid in 1995, the South African government has strived to reduce income inequalities while maintaining economic growth (Celik, 2011; Fleischer, Fuhrmann, Haferburg, & Krüger, 2013; Steinbrink et al., 2011). Thus, the government exploited the nation's general enthusiasm for sports as a source of economic and political leverage and hosted the Rugby World Cup in 1995, the Cricket World Cup in 2003, and the FIFA World Cup in 2010 (Steinbrink et al., 2011). However, unlike the previous mega-events it hosted, the colossal investments required for the FIFA World Cup incited South Africa, the least wealthy country to ever host this event, to introduce strict policies to benefit from it (Müller, 2015).

To maximize the benefits during the World Cup, local governments must rely on revenues from the sale of match tickets and the sales of services and goods to tourists and soccer fans, since FIFA collects most broadcasting revenues and a large share of sales income from goods with their label. Therefore, South African local governments invested in improving general security, and the constructions and renovations of airports, stadiums, fan parks, and shopping malls in their municipalities to attract international investors and visitors as well as to stimulate local economic activities (Celik, 2011; Kaplanidou et al., 2013; Peeters et al., 2014). Public transportation systems were also expanded via the introduction of the first South African high-speed train and new bus lines in host cities to improve South African lives. Overall, in

addition to important private investment (Pfeifer et al., 2018), public investment for the event amounted to USD 3.9 billion, with the construction and renovation of stadiums worth USD 1.3 billion (Easmonn & Trish, 2015; Steinbrink et al., 2011).

Owing to these investment, Pfeifer et al. (2018) with statistical comparison of night light data between host cities and counterfactual cities suggest an economic growth in host cities of a net additional GDP of USD 3.5 billion (1% national GDP) per year between 2005 and 2008. Based on the comparison of national monthly tourists' inflow in South Africa with that of other countries worldwide, (Peeters et al., 2014) suggest that the World Cup participate in attracting 294,804 tourists additional arrivals in 2010 (a 12% increase) and thereby additional revenue for this industry.

Nevertheless, the hosting of the World Cup was linked to a slow-down of certain economic activities in host cities. Large fiscal spending were used to finance the required tournament's infrastructure, at the expense of other projects was thus linked to some opportunity costs (Pillay & Bass, 2008). In this regard, Steinbrink et al. (2011) argue that there were some cuts in subsidized housing projects despite 10% of South African population living without appropriate shelters (2006 statistics) and some displacements of low-income populations for renewal and gentrification of conveniently located areas (e.g. Elis Park stadium in Johannesburg and major highways in Cape Town).

In addition to these redirections of funds, Celik (2011) shows that certain policy for the event aimed at reducing economic activities in the informal sector despite the fact that it provides 60% of jobs (estimates as of 2000). The reason for this policy was security matters, the mismatch that this sector constitute for FIFA image requirements, the protection of FIFA trademark rights regarding counterfeit goods (Steinbrink et al., 2011) and the world-class city images that local governments attached to host cities (Fleischer et al., 2013). Steinbrink et al. (2011) revealed the large-scale policy against informal housing, vendors, and vehicles, which contribute to the disruption of the local economy's business cycle.

Finally, certain evidence suggests that investment during the World Cup were over-optimistic and inefficient. Stemming from the answers of interviews with city and stadium managers, Easmonn and Trish (2015) exhibited the short-sighted investment strategy for the World Cup stadiums as newly built stadiums do not generate enough revenue to cover their annual costs. Other evidence suggest that the event was linked to an oversupply of private investment such as hostels (Easmonn & Trish, 2015; S. L. Ferreira & Boshoff, 2014) and houses (National Treasury, 2011) in the year of the event as demands lower than expected.

Therefore, these conflicting perceptions raise interest in assessing the overall economic effects linked to the organization of the World Cup in host cities. Since the policy for the event were mostly implemented in host cites, the interest lies in their ability to spur an overall

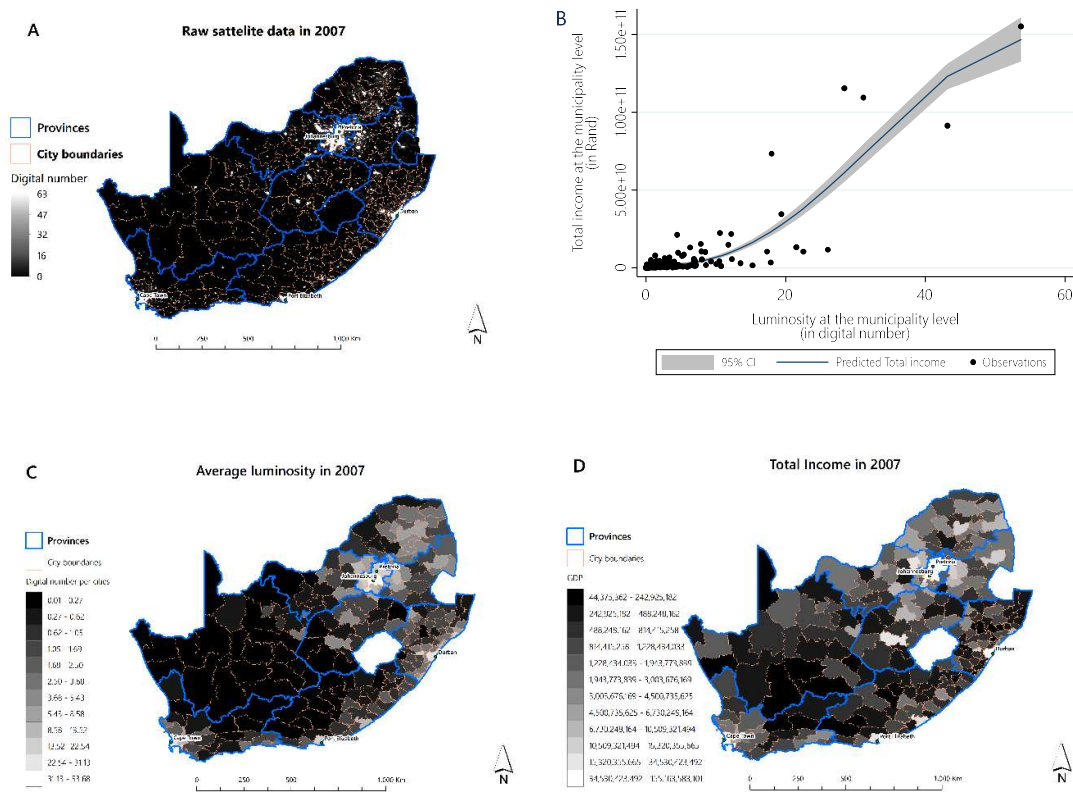
economic growth in host cities, larger than that of other South African cities. I do so by generating a proxy of city-level total income from satellite monitored dataset.

### **5.3. Data and summary statistics**

This study uses data from 234 South African municipalities (cities) of which 9 are host cities. Delimitations of cities are acquired from the Database of Global Administrative Areas (GADM). Then, I collected data on lights at night from the Defense Meteorological Program Operational Linescan System (DMSP-OLS) as indices of yearly economic activities. These data are created by monitoring daily satellite images at 20:30 to 22:00 local time describing persistent human-made light in cities, towns, and other sites while excluding temporary events, such as fires or nonhuman-related activities (*Figure 5-1A*). The unit of night light, the digital number, ranges from 0 to 63. NASA provides annual average night light intensity maps at a resolution of 2.3 km<sup>2</sup> of which I use the average night light intensity by South African municipality for the period 1992–2013.



Figure 5-1 Night light and South African total income in 2007



Then, I predict variation in total income based on variation in light to provide an economic interpretation. For this purpose, I collect cross-sectional cities' total income in 2007 from (Pfeifer et al., 2018), which is to the best of our knowledge, the only dataset portraying the total income distribution in South Africa at the municipal level (see *Figure 5-1C*). Total income is the sum of income from labor (e.g., compensation and bonuses) and income from capital (e.g., rent and dividends). In this study, data on income may be linked to remuneration to labor force and capital for the construction and the use of stadiums, houses, airports, or buildings that emit light. Whereas a reduction in the number of shops, the relocation of housing or fewer cars on the road, which is correlated with less light are linked to lesser income. Also, I collected provincial income as estimated by the South African government from 1995 to 2013 (<http://www.statssa.gov.za/?p=12056>). Provincial income is used to test the ability of variations of income predicted from light to be a reliable index.

As control variables, I use satellite data describing economic activities, affected by the World Cup and linked to night light variations (Ebener, Murray, Tandon, & Elvidge, 2005; Keola, Andersson, & Hall, 2015; B. Zhang, Nozawa, & Managi, 2020). I use cropland and urban area satellite maps from the Terra and Aqua Combined Moderate Resolution Imaging Spectroradiometer (MODIS) (see *Figure 5.7* and (Sulla-Menashe & Friedl, 2018)). The relevance of these variables in the study relates to the different use of land in host cities that may have occurred because of the World Cup (e.g. displacement of informal settlers). In

addition, I use population data derived from LandScan (<https://web.ornl.gov/sci/landscan/>).

This data is produced using demographic census, geographic data, and remote sensing imagery techniques in a multivariate modeling framework (see *Figure 5.8*). The population variable enables us to account for population migration that occurred before, during, and after the event as a result of the World Cup. These control variables were used for the period 2001 – 2013 given their availability. The summary statistics of the data are presented in Table 5-1.

Table 5-1 Summary statistics

| Variables                                          | Unit            | Observations   | Mean                   | Standard Deviation       | Minimum            | Maximum                  |
|----------------------------------------------------|-----------------|----------------|------------------------|--------------------------|--------------------|--------------------------|
| $H_i$ (equals 1 if city is a host and 0 otherwise) |                 | 4,446          | 0.04                   | 0.19                     | 0                  | 1                        |
| Stadiums per city                                  | unit            | 4,446          | 0.04                   | 0.22                     | 0                  | 2                        |
| Light                                              | Digital number  | 4,446<br>(171) | 3<br>(19.5)            | 6<br>(15.3)              | 0<br>(3.4)         | 58<br>(58)               |
| Urban areas                                        | km <sup>2</sup> | 3,042<br>(117) | 33.4<br>(33.7)         | 96.8<br>(262)            | 0<br>(20)          | 779<br>(779)             |
| Cropland areas                                     | km <sup>2</sup> | 3,042<br>(117) | 545<br>(568)           | 791<br>(354)             | 0<br>(42.1)        | 6,780<br>(1,290)         |
| Population                                         | heads           | 3,276<br>(126) | 195,479<br>(1,588,899) | 412,490.8<br>(1,185,719) | 5,474<br>(176,018) | 2,172,733<br>(4,172,733) |

Note: The descriptive statistics in parentheses corresponds to host cities' statistics

## 5.4. Methods

Our objective is to quantify the differential total income growth in host cities compared to other cities. In the first part of the method, I present the estimation of the relationship between light and income per city in 2007 and how this estimation enables the prediction of variations in total income from 1992 to 2013. In the second part, I present the econometrics model used to estimate differential of income growth between host cities and other South African cities.

### 5.4.1 *The spatially varying relationship between total income and night light data*

The usual model to interpret variations of night light as an economic index is the following:

$$\log(eco_i) = \delta_0 + \delta_1 \log(light_i) + \varepsilon_{1i} \quad (5.1)$$

where  $light_i$  represents the average light digital number,  $eco_i$  represents the economic index (such as total income or GDP) of city  $i$ , and the terms  $\delta_0$  and  $\delta_1$  describe the average national relationship between the two previous variables. Building on the relationship in Figure 5-1B, the parameter  $\delta_1$  in (5.1) can be estimated via ordinary least squares (OLS) as  $\hat{\delta}_1 = (\mathbf{x}'\mathbf{x})^{-1}\mathbf{x}'\mathbf{y}$  ( $\mathbf{x}$  and  $\mathbf{y}$  being the vector of independent and dependent variable respectively).

However, there are two concerns with this model.

First, in night light maps, several areas within wealthy cities display the highest unit of measure of light recordable by satellites (Hunter, Maxim, & Xavier, 2017). Thus, in time-series

analysis, the most developed cities may have a light variation that is less elastic than that of other cities. However, since this inelasticity of light is not evidence of income inelasticity, it may cause bias. Second, the total income correlated to a unit growth of light (such as that due to the construction of infrastructure) is larger in developed cities than that in the least wealthy cities (Michalopoulos & Papaioannou, 2018). Considering, the higher salaries and a higher return on investment in a few South African developed cities, this issue is crucial (van Niekerk et al., 2016).

To address these concerns, I carry out a GWR instead of linear model (5.1) (Mellander, Lobo, Stolarick, & Matheson, 2015). This model relaxes the assumption of a stationary relationship between night light and total income. Each city is imputed a unique relationship between light and income that depends on both its own observation and those of nearest cities in terms of distance. We can express the GWR model as follows:

$$\log(eco_i) = \gamma_0(i) + \gamma_1(i) \log(light_i) + \varepsilon_{2i} \quad (5.2)$$

With the coefficient  $\gamma_1(i)$ , estimated as follows:

$$\widehat{\gamma_1(i)} = (\mathbf{x}'\mathbf{W}(i)\mathbf{x})^{-1}\mathbf{x}'\mathbf{W}(i)\mathbf{y} \quad (5.3)$$

where  $y_i$  is the dependent variable,  $x_i$  is the independent variable and  $\varepsilon_{2i}$  is the error term of the city  $i$ . The terms  $\gamma_0(i)$ , and  $\gamma_1(i)$  are the coefficients specific to the city  $i$  that characterize the relationship between  $\log(light_i)$  and  $\log(eco_i)$ .

The estimate  $\widehat{\gamma_1(i)}$  is different from the OLS estimate  $(\widehat{\delta_1})$  by the presence of the square  $\mathbf{W}(i)$ . The calibration of the GWR consists in estimating a linear regression that is inherent to the city  $i$  owing to the presence of  $\mathbf{W}(i)$ . This matrix defines the influence, based on the distance, that other cities have on the calibration of the relationship between light and income at city  $i$ . Each element of this matrix is a weight at city  $i$  expressing its relationship with city  $j$  as follows:

$$w_{ij} = \begin{cases} \exp\left(-\frac{1}{2} \left(\frac{d_{ij}}{b_i}\right)^2\right) & \text{if } d_{ij} \leq b_i, \\ 0 & \text{elsewhere} \end{cases} \quad (5.4)$$

where  $w_{ij}$  is the weight of city  $j$  at city  $i$ ,  $d_{ij}$  is the Euclidean distance (i.e., the shortest distance between the geographical center of two cities) between  $i$  and  $j$ , and  $b_i$  is the bandwidth at observation  $i$ . I use the adaptive bandwidth limit method (built-in function in ArcGIS) to determine the bandwidth that best fit the distribution of the data. Optimal bandwidths were estimated by the minimization of the corrected Akaike information criterion (AICc) technique.

The GWR's central assumption, which stipulates that geographically close observations have a greater influence on one another fits this study because discrepancies in the spatial distribution of light levels and income levels appeared to be geographically clustered (see Figure 5-1C and 5-1D). Furthermore, this assumption leverages agglomeration economies

(Antonelli et al., 2011) between South African cities, embodied by the divergence in the sectors of activity across South African provinces<sup>c</sup> that generate different levels of income.

Thus, I perform a GWR of the total income on night light per municipality in 2007. Then, I assume that there is a similar relationship between light images and income across satellites and years, and predict the income per municipality for 1992–2013 as follows:

$$\log(\widehat{income})_{it} = \widehat{\gamma_0(i)} + \widehat{\gamma_1(i)}\log(light)_{it} \quad (5.5)$$

Where  $\log(\widehat{income})_{it}$  is the estimated logarithm of income in city  $i$  at year  $t$ .

#### 5.4.2 Estimation of differential income growth in host cities with a panel data

Similar to (Baade & Matheson, 2004; Du Plessis & Maennig, 2012; Siegfried & Zimbalist, 2010), I apply a difference-in-difference fixed-effects estimation technique (DID) to estimate the differential economic growth in host cities relative to cities that did not host any soccer game:

$$\log(income)_{it} = \alpha_i + \eta_t + \sum_{k=1}^m \beta_k X_{ikt} + \beta_{post} H_i * post + \vartheta_{it1} \quad (5.6)$$

where  $\log(income)_{it}$  is the logarithm of total income in city  $i$  at year  $t$ . The term  $\alpha_i$  is the city  $i$  fixed effects measuring city specific characteristics steady in time. Since, light is reported by

---

<sup>c</sup> <https://southafrica-info.com/infographics/animation-economic-sectors-of-south-africas-provinces/>

5 different satellites with calibration performance degrading over time, the predicted income will be affected by some measurement errors. Thus, I use a set of year factor variables  $\eta_t$  to capture differences in satellite measures and national shocks in the economy across years (Xi Chen & Nordhaus, 2010; Hunter et al., 2017).

Terms included in  $X_{ikt}$  refers to the total population, and the total areas of urban and agricultural land in the city  $i$ . The coefficient  $\beta_k$  of these variables relates to the effects of a yearly unit increase of the corresponding  $X_{ikt}$  on percentage of income. The inclusions of  $X_{ikt}$  enable the difference-in-difference estimate to measure differential in income when these variables are held constant.

The variable of interest in [5] is the interaction term  $H_i * post$ . The first term of this interaction,  $H_i$ , is a dummy variable indicating that city  $i$  is a host city. The second term,  $post$ , is a dummy variable which equal one for year between 2010 and 2013. The term  $\beta_{post}$  measures the difference in percentage of income of host cities from 2010 to 2013 when income of host cities is compared to (1) the average income in host cities prior to 2010 and (2) the average income in other cities over the full panel data. Hence, I consider a positive value of  $\beta_{post}$  as evidence of higher income growth in host cities relative to other cities. In contrast, a negative value of  $\beta_{post}$  is interpreted as lower growth of income in host cities than in other cities. I assume that  $\beta_{post}$  displays the economic shocks in host cities because the Apartheid history still limits economic growth spillovers across areas (Naude & Krugell, 2003; Yerema



et al., 2020) and the economic impact of the World Cup does not affect non-host cities' income (Baade & Matheson, 2004; Feddersen & Maennig, 2012). As serial correlation is an issue in DID estimations, I use the technique developed by Bertrand, Duflo, & Mullainathan (2004) and performed a block-bootstrap standard error with 1000 repetitions in the regressions.

The identification of the term  $\beta_{post}$  suggests that in the absence of this event, variations of income in host cities and that of other cities would have been similar. This identification strategy lies on the common trend assumption between the host cities and the other cities and emphasizes the need for these “other cities” to be relevant counterfactuals. At face value, non-host South African cities cannot be regarded as relevant contractual since the selection of host cities by the national government was guided by the peculiarity of their large population and income. However, the percentage growth of non-host South African cities may act as meaningful counterfactuals despite host cities being more economically developed. Based on to evidence from Henderson et al. (2012), Duranton (2008), and Gabaix (2009) the percentage growth rate in night light, in income and in population respectively should be similar across cities, independently of their size (e.g. Figure 5-7 shows that administrative data of provincial income growth is similar across provinces despite their dissimilar economic development). Thus, in the study “other cities” refers to South African cities that (1) did not host any World Cup match and (2) have a population larger than 100,000 inhabitants in 2001. I limit my sample to cities with at least 100,000 inhabitants (the median values of population South African

municipalities in 2001) to disregard local municipalities so small that their economies do not behave like cities.

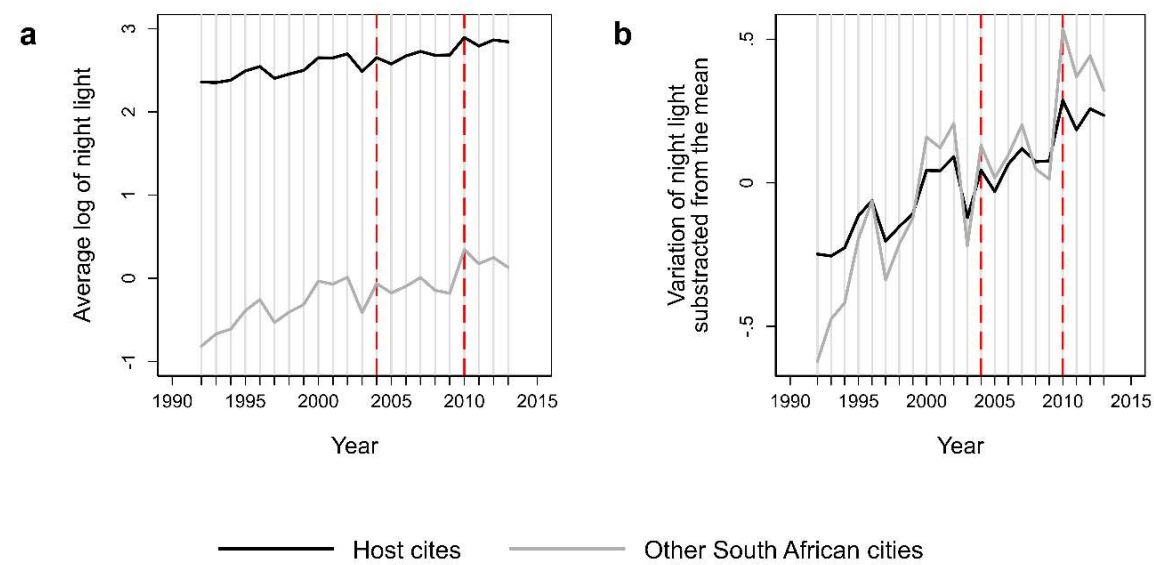
The model (5.6) leverages the aforementioned evidence since  $\beta_{post}$  indicates a percentage growth. The common assumption trend seems to apply at the city-level since plotting the variations of logarithm of light between host and non-host cities subtracted by their respective mean, show a broad similarity between the two types of cities after 1995 (Figure 5-2; In Figure 5-9, I show that this pattern applies to the predicted income). I recognize that this descriptive interpretation of the common assumption trend may be insufficient because of measurement errors of satellite data causing erratic variations in light across years and investment made in anticipation of the event. Thus, building on (Feddersen & Maennig, 2012), I further examine the common assumption trend with a statistical model where I estimate the difference of income growth between host cities and other cities for every year taken individually, starting from the year of the announcement of the event in 2004 with the model:

$$\log(income)_{it} = \alpha_i + \eta_t + \sum_{k=1}^m \beta_k X_{ikt} + \sum_{y=200}^{2013} \beta_y H_i * t_y + \vartheta_{it2} \quad (5.7)$$

Nevertheless, I refrain from claiming any causal estimate because summary statistics in Table 5-1 show a structural difference between these two types of host cities and non-host cities. Also, I avoid using methods like a Synthetic Control Method (SCM) as a main estimate

because I believe that the assumption that there were no population migration and economic exchange across host and non-host cities from 2003 to 2013 is too strong.

Figure 5-2 Trends in light growth between host cities and other cities



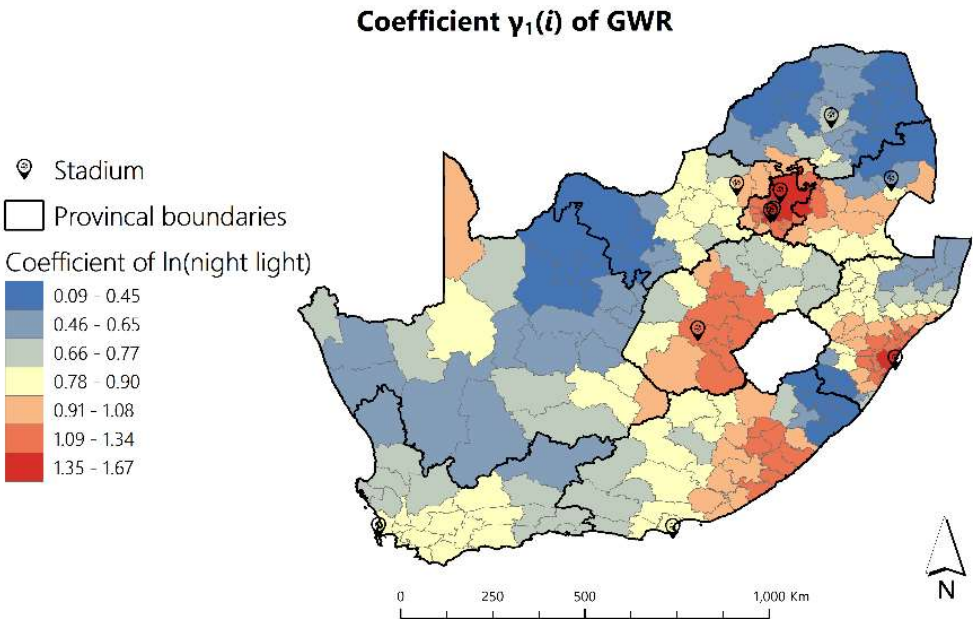
Note: The figure-a plots the yearly change in the log night light in host cities and other cities. The Figure b refers to the yearly change in the night light intensity subtracted from their mean during the period 1992 – 2013 in host cities and other South African cities

## **5.5. Results and discussion**

### *5.5.1 The relationship between economic development and night light*

The GWR results, which portray the spatial relationship between night light and total income, are shown in Figure 5-3. Depending on locations, there are wide differences in the elasticity between total income and night light, as a 1% increase in night light relates to an increase in total income ranging from 0.09% to 1.67% (Figure 5-3). This association is strong in host cities, as their night light coefficients range between 0.78% and 1.67%.

Figure 5-3 GWR estimates of the relationship between total income and night light in 2007



To observe the robustness of the GWR results, I compare its performance with OLS estimates of the average national relationship between light and income using Equation (5.1). OLS estimates show that an increase in night light of 1% is related to an increase of 0.66% in total income (see Table 5-2). However, this national average relationship appears to be less robust than the city-level relationships for two reasons. First, the GWR results is consistent with previous studies' findings (Henderson et al., 2012; Hodler & Raschky, 2014). Indeed, the range of variation in the estimates varies around the estimates in previous studies that associated a 1% increase in light to a 0.3% increase in the GDP internationally.

On the other hand, since both his AICc is lower and his R-squared is higher than that of the OLS when training the models, the goodness of fit of the GWR outperforms that of the OLS (see Figure 5.-1). In addition, to test of prediction performances, I use GWR and OLS models to predict city-level income between 1995-2013, then I average these predictions at the provincial level. I compare the respective ability of these estimates in explaining yearly variations of the most disaggregated income dataset provided by the South African government between 1995-2013. It results that GWR predictions are more accurate than that of the OLS since it predicts 70% of yearly provincial GDP variations while OLS predicts 55% (see Table 5-3).

In accordance with Mellander et al. (2015), I assume that the difference in results between the OLS and GWR can be attributed to the underestimation of income growth in wealthy cities and its overestimation in smaller towns by national averages relationship between income and light. Nevertheless, in the following estimations of the difference in income growth on income in host cities, I conduct a secondary analysis with the commonly used OLS (Chen and Nordhaus, 2010; Henderson, Storeygard and Weil, 2012; Pfeifer, Wahl and Marczak, 2018) to appreciate this bias.

Table 5-2 The OLS estimates of the relationship between total income and night light in 2007

| VARIABLES                   | Log of total income   |
|-----------------------------|-----------------------|
| Log of light ( $\delta_1$ ) | 0.659***<br>(0.0467)  |
| Constant ( $\delta_0$ )     | 20.713***<br>(0.0583) |
| Observations                | 234                   |
| R-squared                   | 0.549                 |

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



Table 5-3 The comparative ability of GWR and OLS estimates to predict provincial income

|                                   | (1)                                     | (2)                       |
|-----------------------------------|-----------------------------------------|---------------------------|
| Dependent variable                | South Africa official provincial income |                           |
| Model used for income predictions | OLS                                     | GWR                       |
| OLS income predictions            | 8.34e-05***<br>(2.58e-05)               |                           |
| GWR income predications           |                                         | 4.33e-05***<br>(4.66e-06) |
| Constant                          | 20,723<br>(55,953)                      | 42,250**<br>(20,634)      |
| Observations                      | 171                                     | 171                       |
| R-squared                         | 0.545                                   | 0.696                     |
| Year fixed effects                | YES                                     | YES                       |
| Number of Provinces               | 9                                       | 9                         |
| Period of study                   | 1995-2013                               | 1995-2013                 |

Block bootstrapped standard errors with 1000 repetitions in parentheses \*\*\* p<0.01, \*\*

p<0.05, \* p<0.1

### *5.5.2 Differential income growth in host cities*

#### *5.5.2.1 Yearly difference in income growth between host cities and other cities*

In Figure 5-4, I display the estimates of the differential income growth and light growth between host cities and other cities for every year taken individually from 2004 to 2013, while controlling for changes population, as well as in urban and agricultural areas. In these graphs, I highlight the point estimates of the difference in growth and the 95% confidence interval of each estimate (exact estimates are presented in Table 5-6).

The yearly estimates of the differential income growth between host cities and other South African cities show a similarity in growth rate between host cities and other cities is perceived. Thus, for most of the year, there is no significant superior economic activities in host cities relative to other cities since the difference in income growth (and light growth) between these two types of cities is not statistically different from zero. Hence, these results first display a broad common trend assumption in income and night light growth between host and non-host cities. In Figure 5-11 in the Appendix, I confirm the common assumption trend between 1992 and 2013 while excluding control variables since they are unavailable prior to 2001. Despite being mostly insignificant, the estimates display three trends in the differential income growth between host cities and other cities. Two patterns from 2004 to 2009 and from

2010 and 2013 highlight a gradual increase in economic activities in host cities relative to that of other cities and a pattern from 2009 to 2010 pointing a sharp decline in economic activities.

Among the estimates of differential income growth between host cities and other cities only those in 2009 and 2010 display some statistical significance at the 95% confidence level when using night light data as the dependent variable. It appears that light growth in host cities were 13.4% ( $p=0.05$ ) larger than that of other cities in 2009, while in 2010 it was 14.2% lower<sup>f</sup> ( $p=0.04$ ). If we were to assume that night light is linearly linked an income variation, these estimates would suggest an income growth of 8.8% in 2009 and -9.3% in 2010 in host cities relative to other cities<sup>g</sup>. However, when accounting for the spatially varying relationship between night light and income with the GWR, I do not find any significant difference in income growth between host cities and other cities in 2010 and in 2009. Hence, despite the lower light growth than in host cities in certain years, estimates from GWR suggest that they do not reflect a difference in income growth between host cities and other South African cities. Therefore, we fail to reject the hypothesis no difference in income growth between host cities and other cities.

From this analysis, we cannot confirm a substantial impact linked to the event. Hence, the results differ from those of Pfeifer et al. (2018), who claimed a positive impact of the World

---

<sup>f</sup> Because of the log-linear relationship in our models, the percentage growth is estimated as  $(\exp(\beta) - 1) * 100\%$ .

<sup>g</sup> Coefficient of differential light growth \* OLS relationship between light and income

Cup in South African host cities between 2004 and 2008 with variations in night light data. The relevance of the estimates compared to that of Pfeifer et al. (2018) resides in the data and method used. Regarding the data, the use of population, urban land and cropland variations as control variables in the current study provides more pertinent economic interpretations of the effect of the World Cup than the data used in the previous study. In an effort to make a clear comparison between the data and those of Pfeifer et al. (2018), I carry out an SCM and show that the additional income growth in host cities between 2004 and 2008 is not statistically significant as suggested by the previous study (see Figure 5-14). Regarding the method, I believe that the panel data estimation technique better fits the analysis since it enables the account of satellite measurement errors across years and World Cup-related migrations occurring after 2004. Finally, building on Easmonn and Trish (2015) and Peeters et al. (2014), who argue that economic effects linked to the World Cup picked around 2010, and not before, I believe that the yearly estimate of the differential income growth between host cities and other cities fit previous findings.

Figure 5-4 The difference in predicted income growth between host cities and other cities from 2002 to 2013



Note: This graph plots to the coefficient estimate of the difference in total income growth between host cities and other cities from 2004 to 2013. The markers in white circle represent the estimate results and the error-bar around these estimates represent the confidence interval at a 95-confidence level.

#### *5.5.2.2 Persistent income growth between host cities and other cities*

Table 5-4 presents the results of the differential growth in total income and night light between host cities and other South African cities. The columns 1 and 2 use GWR-predicted income as dependent variables while the columns 3 and 4 use light as dependent variable. Also, the columns 1 and 3 control for changes in population, urban and cropland areas with a time period restricted between 2001 and 2013 because of the availability of the control variables. Meanwhile the columns 2 and 4 include data from 1992 to 2013 but ignore any confounding factor.

In the preferred model specifications using predicted income as dependent variable, there is no statistically significant difference in income growth between host cities and other South African cities (Columns 1 and 2). However, results using light as dependent variable show that income growth in host cities is lower than that of other South African cities between 2010 and 2013 when estimated jointly ( $p < 0.01$ ). If we were to assume a linear relationship between income and light, this analysis would suggest a lower income growth between 7% and 14.3% in host cities (in Column 3 and 4 respectively). However, building on the evidence of heterogeneous relationship between income and light at the city level (see Methods and GWR results), we posit that estimates that use GWR-predicted income are more reliable. Thus, we fail to reject the hypothesis against the similarity of income growth between host and other cities.

Overall, the absence of difference in income growth between host cities and other South African cities may be explained by the brevity of the event (lasting a month from June 11 to July 11), which was likely too short for the associated tourist activities, hotel industry usage, use of stadiums, and other opportunities to result in the desired changes in the economy (Du Plessis & Maennig, 2012; S. L. Ferreira & Boshoff, 2014; Kaplanidou et al., 2013).

Table 5-4 The differential income growth in host cities

| Dependent variables                            | Log(Total income)   |                     | Log(Light)            |                       |
|------------------------------------------------|---------------------|---------------------|-----------------------|-----------------------|
|                                                | (1)                 | (2)                 | (3)                   | (4)                   |
| Host cities between 2010 and 2013              | -0.0525<br>(0.0359) | -0.0655<br>(0.0593) | -0.127***<br>(0.0395) | -0.245***<br>(0.0617) |
| Cropland (10 <sup>9</sup> km <sup>2</sup> )    | -0.07**<br>(0)      |                     | -0.07**<br>(0)        |                       |
| Urban areas (10 <sup>9</sup> km <sup>2</sup> ) | 2.21<br>(30.5)      |                     | 11.6<br>(33.4)        |                       |
| Population (10 <sup>9</sup> heads)             | 74.5<br>(60.7)      |                     | -23.6<br>(90.3)       |                       |
| Constant                                       | 21.15***<br>(1.731) | 20.76***<br>(0.135) | 0.0851<br>(1.867)     | -0.0515<br>(0.165)    |
| Observations                                   | 1,521               | 2,574               | 1,521                 | 2,574                 |
| R-squared                                      | 0.558               | 0.544               | 0.568                 | 0.558                 |
| Number of cities                               | 117                 | 117                 | 117                   | 117                   |
| Period                                         | 2001-2013           | 1992-2013           | 2001-2013             | 1992-2013             |
| Cities FE                                      | YES                 | YES                 | YES                   | YES                   |
| Year FE                                        | YES                 | YES                 | YES                   | YES                   |

Block bootstrapped standard errors with 1000 repetitions in parentheses \*\*\* p<0.01, \*\*

p<0.05, \* p<0.1



### *5.5.2.3 Heterogeneous income growth among host cities between host cities and other cities*

Despite the results failing to exhibit a significant difference in income growth between host and non-host cities, findings from (Easmonn & Trish, 2015) may imply some difference in income growth across host cities. This difference could be linked to local governments allocating too large investment on stadium infrastructure to the detriment of inclusive investment. In (Easmonn & Trish, 2015), several government officials admitted that the construction of new stadiums were not essential to the holding of the event and some officials allude to cases of embezzlement and over expenditure for these sport facilities.

I perform two additional analyses in this regard. First, I compare the income growth of host cities that invest in new stadiums (Cape Town, Durban, Nelspruit, Port Elizabeth and Polokwane) with that of cities that renovate an existing stadium. In the second analysis, I compare the income growth between host cities that allocated the most resources for the construction or the renovation of their stadiums to the income growth of other host. For this purpose, I split host cities in two groups of equal number according to expenses for their stadiums, with Cape Town, Johannesburg, Durban, Port Elizabeth, and Polokwane being the cities with the costliest stadiums (see cost of stadiums in Table 5-7).

The results of this analysis are displayed in Table 5-5. It appears that cities that construct a new stadium had an income growth lower than host cities that did not engage in such an investment. However, inference from these results are not appropriate since this estimate is not

statistically significant ( $p=0.11$ ). However, cities that allocate large financial resources on their stadiums experience a statistically significantly lower income growth than other host cities ( $p<0.01$ ). Our estimates suggest that cities with the costliest stadiums had an income growth that was 9.7% lower than that of other host cities from 2010 to 2013 on average. I consider these results as a piece of evidence inefficient investment and an opportunity cost for these local governments.

Table 5-5 The differential income growth across host cities

| Dependent variables                                  | Log(Total income)     |                       |
|------------------------------------------------------|-----------------------|-----------------------|
|                                                      | (1)                   | (2)                   |
| Host cities between 2010 and 2013                    | 0.00112<br>(0.0614)   | 0.0136<br>(0.0413)    |
| Cities with a new stadium between 2010 and 2013      | -0.0911<br>(0.0683)   |                       |
| Cities with costliest stadiums between 2010 and 2013 |                       | -0.102***<br>(0.0386) |
| Cropland (10 <sup>9</sup> km <sup>2</sup> )          | -0.0720**<br>(0.0294) | -0.0722**<br>(0.0290) |
| Urban areas (10 <sup>9</sup> km <sup>2</sup> )       | 1.878<br>(30.60)      | 2.492<br>(30.42)      |
| Population (10 <sup>9</sup> heads)                   | 63.74<br>(59.86)      | 80.80<br>(56.89)      |
| Constant                                             | 21.17***<br>(1.751)   | 21.13***<br>(1.734)   |
| Observations                                         | 1,521                 | 1,521                 |
| R-squared                                            | 0.559                 | 0.560                 |
| Number of cities                                     | 117                   | 117                   |
| Period                                               | 2001-2013             | 2001-2013             |
| Cities FE                                            | YES                   | YES                   |
| Year FE                                              | YES                   | YES                   |

Block bootstrapped standard errors with 1000 repetitions in parentheses \*\*\* p<0.01, \*\*

p<0.05, \* p<0.1

## 5.6. Conclusions

This study attempted to estimate the additional income growth in South African cities that host the World Cup 2010 using satellite data as an economic proxy. In a first part of the chapter, I improve the economic significance of night light variations by predicting subnational total income between 1992 and 2013 from night light maps. This prediction is based on the spatially varying relationship between light and income with GWR to correct the spatially specific correlation between light and income. Hence, this study is the first to show that variations in night light can better explain economic variations in cross-sectional and longitudinal analyses when the variations in night light incorporate the spatial economic agglomeration.

Then, in the second part of the study, I show that host cities have considerably lower light growth than other South African cities, but this does not reflect a difference in overall income growth between host cities and other South African cities. This finding is in line with most of the scholarly ex-post analyses on mega-events at the city-level. Subsequently, I found that host cities whose local government spent the most on stadium renovation and construction had a lower income growth than other host cities.

Overall these findings revealed two types of opportunity costs linked to the event. Despite the large investment they acquired from national government, host cities had an economic growth similar to that of other South African cities that did not invest any public

capital for the event. It shows the limited ability of this World Cup to stimulate local economic growth. Results indicating a lower income growth in cities with the costliest stadiums imply that these sorts of investments might have been better spent elsewhere.

Despite the results, nations and cities may still have incentives to host the World Cup as it is linked to trade openness, the improvement of the image of the country and changes to citizens' perceptions of politics (Fourie & Santana-Gallego, 2011; Kaplanidou et al., 2013; Peeters et al., 2014; Rose & Spiegel, 2011). In such circumstances, I believe that the organization of events of smaller magnitude and the organization of events in collaboration with other countries could reduce the amount of investment in sport-related infrastructure and redirect investment into areas that could generate inclusive economic growth.

Finally, I need to acknowledge that I cannot ascertain that host cities did not benefit from hosting the World Cup in the longer term due to data limitations. Moreover, despite the inclusion of satellite dummy variables in the fixed-effects, it should be noted that the results of the GWR rely on the assumption that the relationship between night light maps and income is the same across satellites and years. Thus, there is a need for further research to extent the periodicity of night light maps and their robustness across years and satellite. These issues remain for future research to further explore.

## References

- Antonelli, Cristiano, Pier Paolo Patrucco, and Francesco Quatraro. 2011. "Productivity Growth and Pecuniary Knowledge Externalities: An Empirical Analysis of Agglomeration Economies in European Regions." *Economic Geography* 87 (1): 23–50. <https://doi.org/10.1111/j.1944-8287.2010.01104.x>.
- Baade, Robert A., and Victor A. Matheson. 2004. "The Quest for the Cup: Assessing the Economic Impact of the World Cup." *Regional Studies* 38 (4): 343–54. <https://doi.org/10.1080/03434002000213888>.
- Barclay, Jonhatan. 2009. "Predicting the Costs and Benefits of Mega-sporting Events: Misjudgement of Olympic Proportions?" *Economic Affairs* 29 (2): 62–66.
- Bertrand, Marianne, Esther Duflo, and Sendhil Mullainathan. 2004. "How Much Should We Trust Differences-in-Differences Estimates?" *Quarterly Journal of Economics*. <https://doi.org/10.1162/003355304772839588>.
- Cavalcanti, Clarissa de Oliveira, Marcelo Limont, Maurício Dziedzic, and Valdir Fernandes. 2017. "Sustainability of Urban Mobility Projects in the Curitiba Metropolitan Region." *Land Use Policy* 60: 395–402. <https://doi.org/10.1016/j.landusepol.2016.11.015>.
- Celik, Ercument. 2011. "The Exclusion of Street Traders from the Benefits of the FIFA 2010 World Cup in South Africa." *African Journal of Business Economic Research* 6 (1).
- Chen, Xi, and William D. Nordhaus. 2010. "The Value of Luminosity Data as a Proxy for Economic Statistics." *Proceedings of the National Academy of Sciences* 108 (21): 8589–8594. <https://doi.org/10.2139/ssrn.1666164>.
- Coates, Dennis, and Brad R. Humphreys. 2002. "The Economic Impact of Postseason Play in Professional Sports." *Journal of Sports Economics* 3 (3): 291–99. <https://doi.org/10.1177/1527002502003003005>.
- Diener, Ed, Daniel Kahneman, William Tov, and Raksha Arora. 2010. "Income's Association with Judgements of Life versus Life Feelings." In *International Differences in Well-Being*, edited by Ed Diener, John F. Helliwell, and Daniel Kahneman, 2nd ed., 3–33. New York: Oxford University Press.
- Duranton, Gilles. 2008. "Cities : Engines of Growth and Prosperity for Developing Countries?" 12. Washington DC.
- Easmonn, Molloy, and Chetty Trish. 2015. "The Rocky Road to Legacy: Lessons from the 2010 FIFA World Cup South Africa Stadium Program." *Project Management Journal* 46 (3): 88–107. <https://doi.org/10.1002/pmj>.

- Ebener, Steeve, Christopher Murray, Ajay Tandon, and Christopher C. Elvidge. 2005. "From Wealth to Health: Modelling the Distribution of Income per Capita at the Sub-National Level Using Night-Time Light Imagery." *International Journal of Health Geographics*. <https://doi.org/10.1186/1476-072X-4-5>.
- Feddersen, Arne, and Wolfgang Maennig. 2012. "Sectoral Labour Market Effects of the 2006 FIFA World Cup." *Labour Economics* 19 (6): 860–69. <https://doi.org/10.1016/j.labeco.2012.07.006>.
- Ferreira, Sanette Lacea, and Adriaan Boshoff. 2014. "Post-2010 FIFA Soccer World Cup: Oversupply and Location of Luxury Hotel Rooms in Cape Town." *Current Issues in Tourism*. <https://doi.org/10.1080/13683500.2013.776524>.
- Fleischer, Matthias, Maximilian Fuhrmann, Christoph Haferburg, and Fred Krüger. 2013. "'Festivalisation' of Urban Governance in South African Cities: Framing the Urban Social Sustainability of Mega-Event Driven Development from Below." *Sustainability* 5 (12): 5225–48. <https://doi.org/10.3390/su5125225>.
- Fourie, Johan, and María Santana-Gallego. 2011. "The Impact of Mega-Sport Events on Tourist Arrivals." *Tourism Management* 32 (6): 1364–70. <https://doi.org/10.1016/j.tourman.2011.01.011>.
- Gabaix, Xavier. 2009. "Power Laws in Economics and Finance." *Annual Review of Economics* 1 (1): 255–94. <https://doi.org/10.1146/annurev.economics.050708.142940>.
- Haybron, Daniel M. 2008. "Philosophy and the Science of Subjective Well-Being." In *The Science of Subjective Well-Being*, edited by Eid Michael and Larsen J. Randy, 17–43. New York: the Guilford Press.
- Henderson, J Vernon, Adam Storeygard, and David N Weil. 2012. "Measuring Economic Growth from Outer Space." *American Economic Review* 102 (2): 994–1028. <https://doi.org/10.1257/aer.102.2.994>.
- Hodler, Roland, and Paul A Raschky. 2014. "Regional Favoritism." *Quarterly Journal of Economics* 129 (2): 995–1033. <https://doi.org/10.1093/qje/qju004>. Advance.
- Hunter, Clark, Pinkovskiy Maxim, and Sala-i-Martin Xavier. 2017. *China's GDP Growth May Be Understated*. National Bureau of Economic Research. Vol. No. w23323.
- Kahneman, Daniel, Schakade A. David, Fisher Claude, Krueger B. Alan, and Krilla Amy. 2010. "The Structure of Well-Being in Two Cities." In *International Differences in Well-Being*, edited by Ed Diener, John F. Helliwell, and Daniel Kahneman, 2nd ed., 16–33. New York: Oxford University Press.

- Kaplanidou, Kyriaki, Kostas Karadakis, Heather Gibson, Brijesh Thapa, Matthew Walker, Sue Geldenhuys, and Willie Coetzee. 2013. "Quality of Life, Event Impacts, and Mega-Event Support among South African Residents before and after the 2010 FIFA World Cup." *Journal of Travel Research* 52 (5): 631–45.  
<https://doi.org/10.1177/0047287513478501>.
- Keola, Souknilanh, Magnus Andersson, and Ola Hall. 2015. "Monitoring Economic Development from Space: Using Nighttime Light and Land Cover Data to Measure Economic Growth." *World Development* 66: 322–34.  
<https://doi.org/10.1016/j.worlddev.2014.08.017>.
- Mellander, Charlotta, José Lobo, Kevin Stolarick, and Zara Matheson. 2015. "Night-Time Light Data: A Good Proxy Measure for Economic Activity?" *PLoS ONE*.  
<https://doi.org/10.1371/journal.pone.0139779>.
- Michalopoulos, Stelios, and Elias Papaioannou. 2018. "Spatial Patterns of Development: A Meso Approach." *Annual Review of Economics* 10 (1): 383–410.  
<https://doi.org/10.1146/annurev-economics-080217-053355>.
- Müller, Martin. 2015. "What Makes an Event a Mega-Event ? Definitions and Sizes." *Leisure Studies* 34 (6): 627–42. <https://doi.org/10.1080/02614367.2014.993333>.
- National Treasury. 2011. "Budget Review 2011." Pretoria, South Africa.
- Naude, W. A., and W. F. Krugell. 2003. "Are South Africa's Cities Too Small?" *Cities*.  
[https://doi.org/10.1016/S0264-2751\(03\)00011-8](https://doi.org/10.1016/S0264-2751(03)00011-8).
- Niekerk, Adriaan van, Danie du Plessis, Ilze Boonzaaier, Manfred Spocter, Sanette Ferreira, Lieb Loots, and Ronnie Donaldson. 2016. "Development of a Multi-Criteria Spatial Planning Support System for Growth Potential Modelling in the Western Cape, South Africa." *Land Use Policy* 50: 179–93. <https://doi.org/10.1016/j.landusepol.2015.09.014>.
- Oishi, Shigehiro. 2010. "Culture and Well-Being: Conceptual and Methodological Issues." In *International Differences in Well-Being*, edited by Ed Diener, F. Helliwell John, and Kahneman Daniel, 34–69. New York, USA: Oxford University Press.
- Peeters, Thomas, Victor Matheson, and Stefan Szymanski. 2014. "Tourism and the 2010 World Cup: Lessons for Developing Countries." *Journal of African Economies*.  
<https://doi.org/10.1093/jae/ejt031>.
- Pfeifer, Gregor, Fabian Wahl, and Martyna Marczak. 2018. "Illuminating the World Cup Effect: Night Lights Evidence from South Africa." *Journal of Regional Science* 58 (5): 887–920. <https://doi.org/10.1111/jors.12410>.



- Pillay, Udesb, and Orli Bass. 2008. "Mega-Events as a Response to Poverty Reduction: The 2010 FIFA World Cup and Its Urban Development Implications." *Urban Forum* 19 (3): 329. <https://doi.org/10.1007/s12132-008-9034-9>.
- Plessis, Stan Du, and Wolfgang Maennig. 2012. "The 2010 FIFA World Cup High-Frequency Data Economics: Effects on International Tourism and Awareness for South Africa." In *International Handbook on the Economics of Mega Sporting Events*, 28:482–503. Cheltenham, UK: Edward Elgar Publishing. <https://doi.org/10.1080/0376835X.2011.595994>.
- Pokhriyal, Neeti, and Damien Christophe Jacques. 2017. "Combining Disparate Data Sources for Improved Poverty Prediction and Mapping." *Proceedings of the National Academy of Sciences* 114 (46): E9783–92. <https://doi.org/10.1073/pnas.1700319114>.
- Pop, Iuliana, Adrian Kanovici, Grațela Ghic, and Mădălina Andrei. 2016. "The Economic Effects of the Mega Sport Events on Tourism in the BRICS Countries Case." *Amfiteatru Economic Journal* 18: 960–75.
- Rose, Andrew K., and Mark M. Spiegel. 2011. "The Olympic Effect." *Economic Journal* 121 (553): 652–677. <https://doi.org/10.1111/j.1468-0297.2010.02407.x>.
- Siegfried, John, and Andrew Zimbalist. 2010. "The Economics of Sports Facilities and Their Communities." *Journal of Economic Perspectives* 14 (3): 95–114. <https://doi.org/10.1257/jep.14.3.95>.
- Steinbrink, Malte, Christoph Haferburg, and Astrid Ley. 2011. "Festivalisation and Urban Renewal in the Global South: Socio-Spatial Consequences of the 2010 FIFA World Cup." *South African Geographical Journal* 93 (1): 15–28. <https://doi.org/10.1080/03736245.2011.567827>.
- Sulla-Menashe, Damien, and Mark A Friedl. 2018. "User Guide to Collection 6 MODIS Land Cover (MCD12Q1 and MCD12C1) Product." Reston, VA, USA. <https://doi.org/10.5067/MODIS/MCD12Q1>.
- Wray, Chris, Alisson Todes, Graeme Gotz, and Philip Harrisson. 2014. *Changing Space, Changing City: Johannesburg after Apartheid*. Wits University Press. <https://doi.org/10.18772/22014107656>.
- Yerema, Coulibaly Thierry C.T., Mihoko Wakamatsu, Moinul Islam, Fukai Hiroki, Shunsuke Managi, and Bingqi Zhang. 2020. "Differences in Water Policy Efficacy across South African Water Management Areas." *Ecological Economics* 175: 106707. <https://doi.org/10.1016/j.ecolecon.2020.106707>.
- Zhang, Bingqi, Wataru Nozawa, and Shunsuke Managi. 2020. "Sustainability Measurements

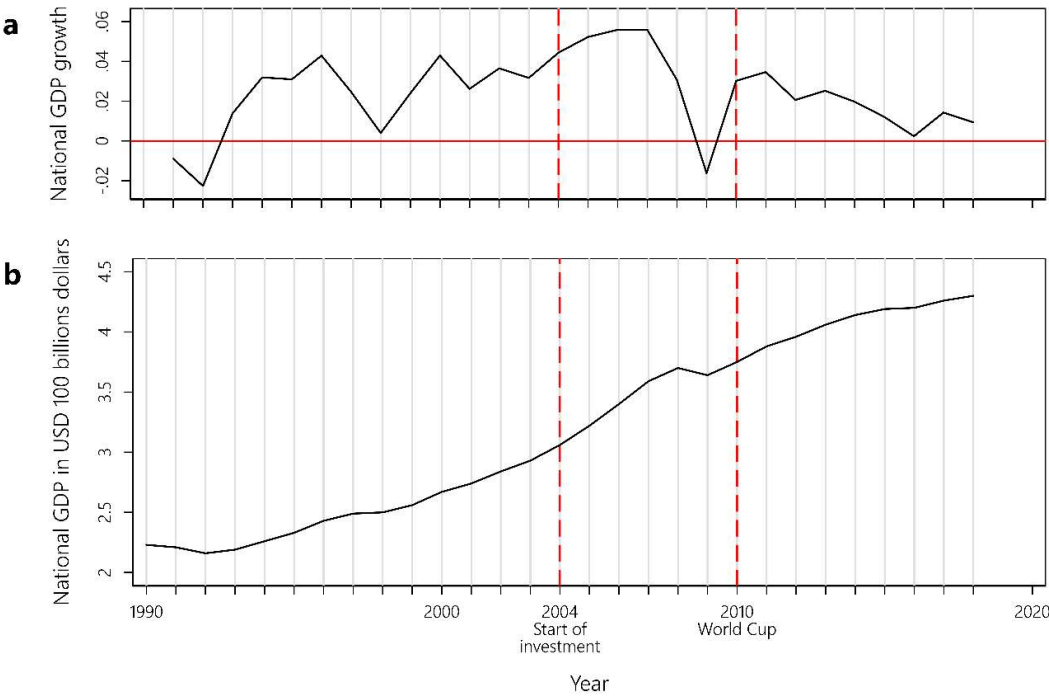
in China and Japan: An Application of the Inclusive Wealth Concept from a Geographical Perspective.” *Regional Environmental Change* 20 (2). <https://doi.org/10.1007/s10113-020-01658-x>.

Zhang, Xin, Xiaobo Zhang, and Xi Chen. 2017. “Happiness in the Air: How Does a Dirty Sky Affect Mental Health and Subjective Well-Being?” *Journal of Environmental Economics and Management* 85 (September): 81–94. <https://doi.org/10.1016/j.jeem.2017.04.001>.

# Appendix 5

## Appendix 5 - Section A: Supplementary tables and figures

Figure 5-5 National South African GDP from 1990 to 2018



Source: The World Bank Group

Figure 5-6 Provincial South African GDP from 1990 to 2018

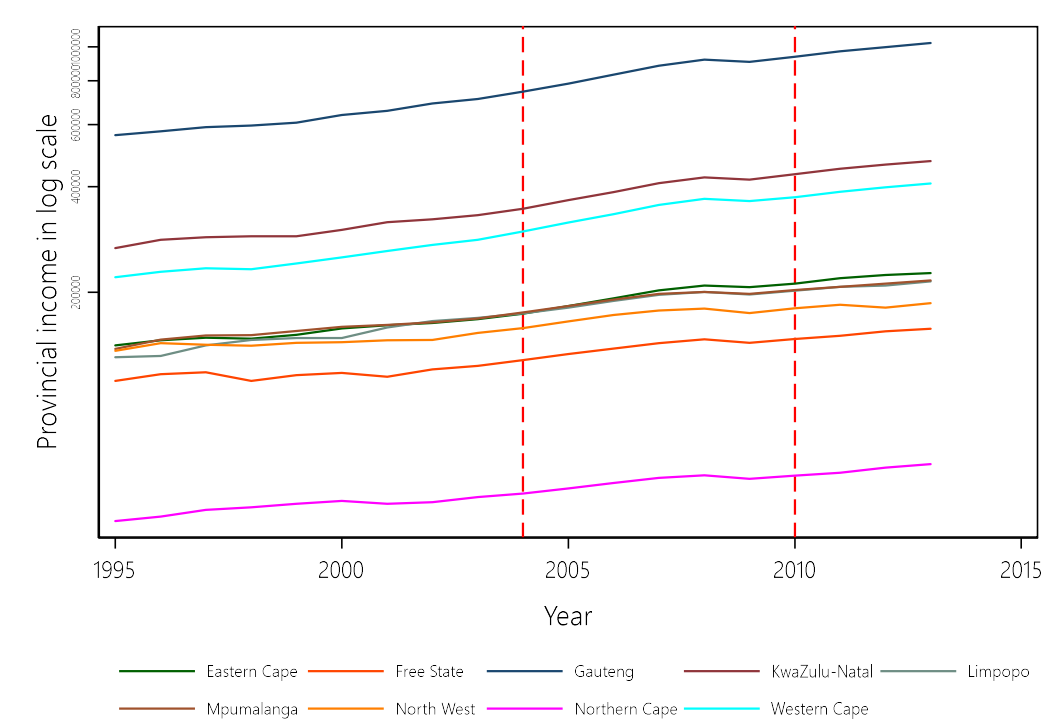


Figure 5-7 The distribution of the land cover in 2010

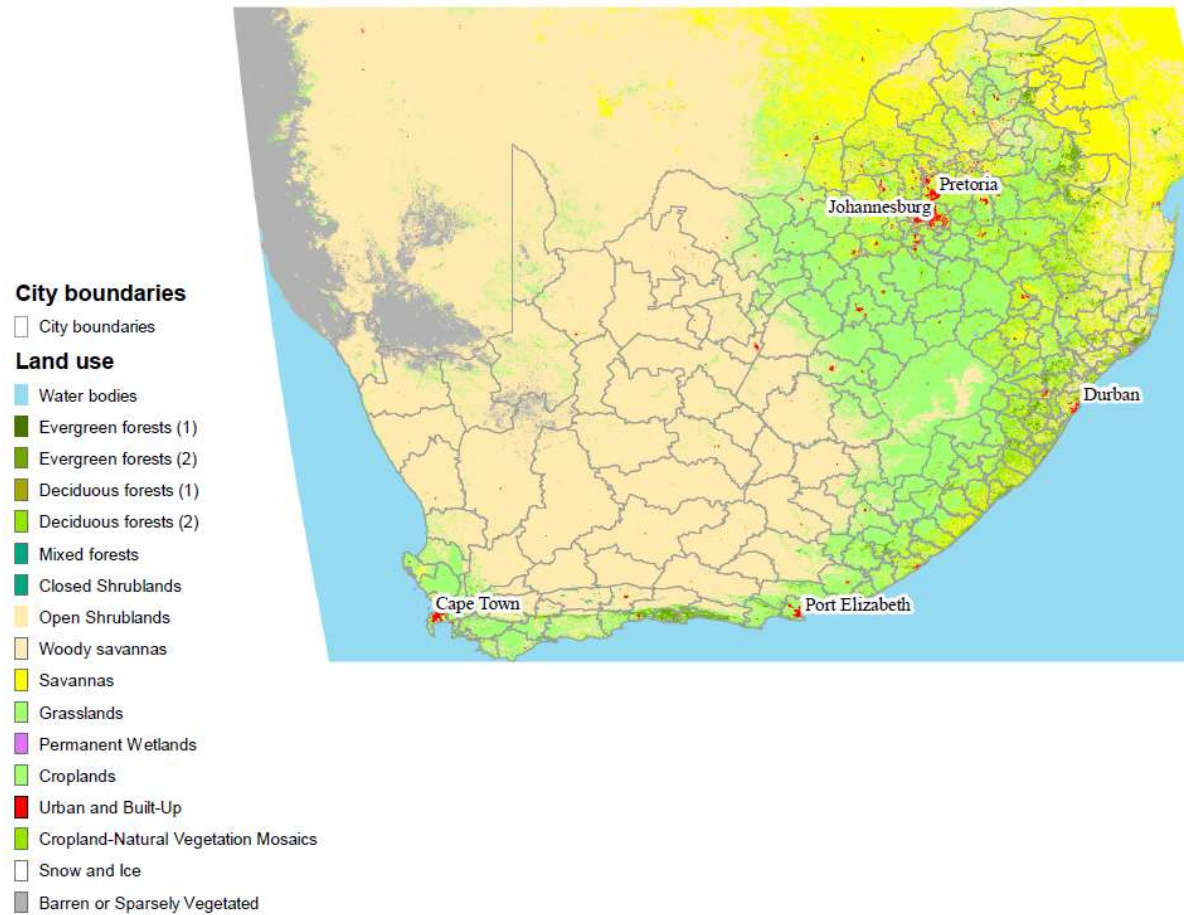


Figure 5-8 The distribution of the population in 2010

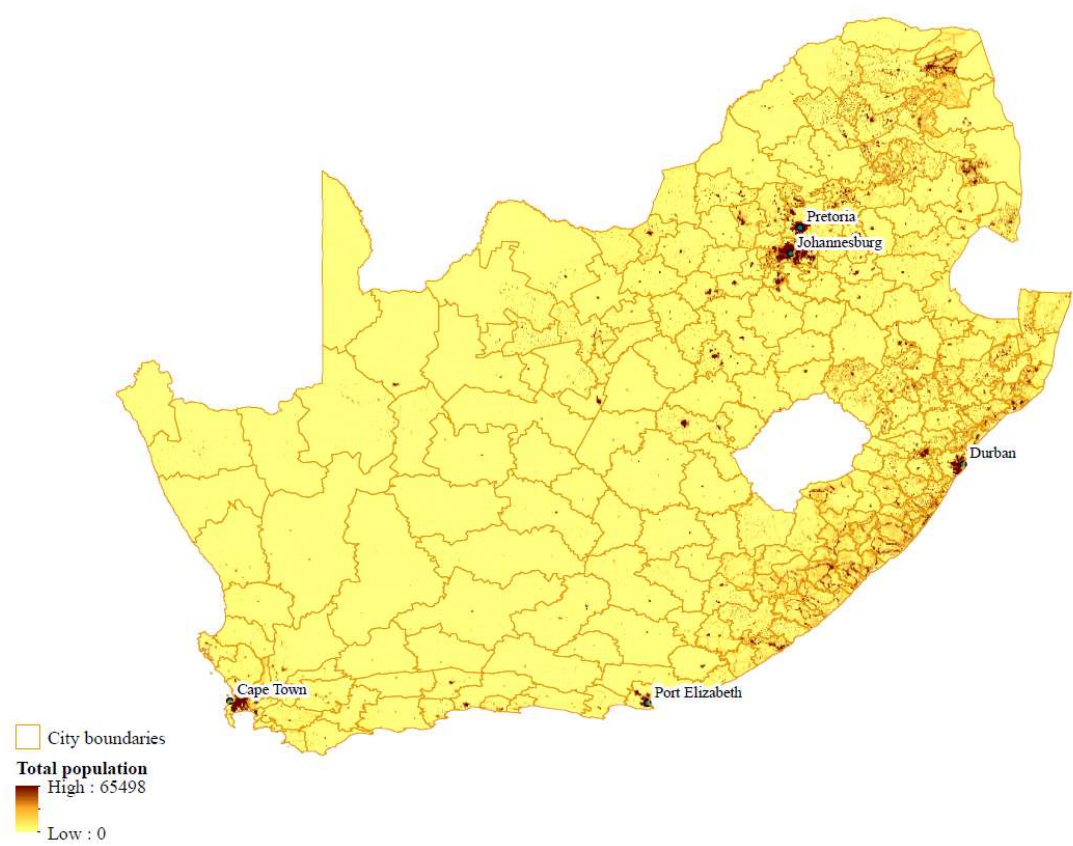


Figure 5-9 Trends of the growth of night light between host cities and other cities

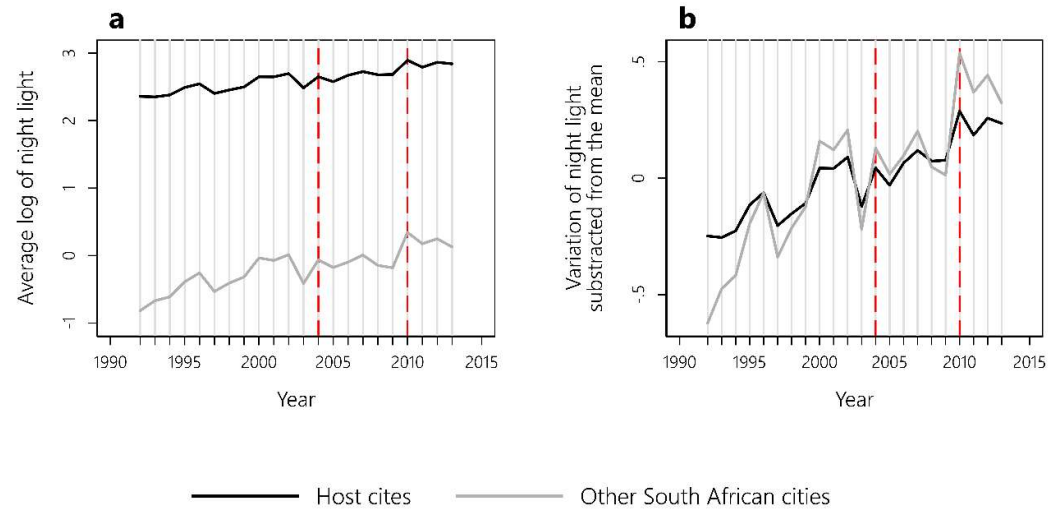
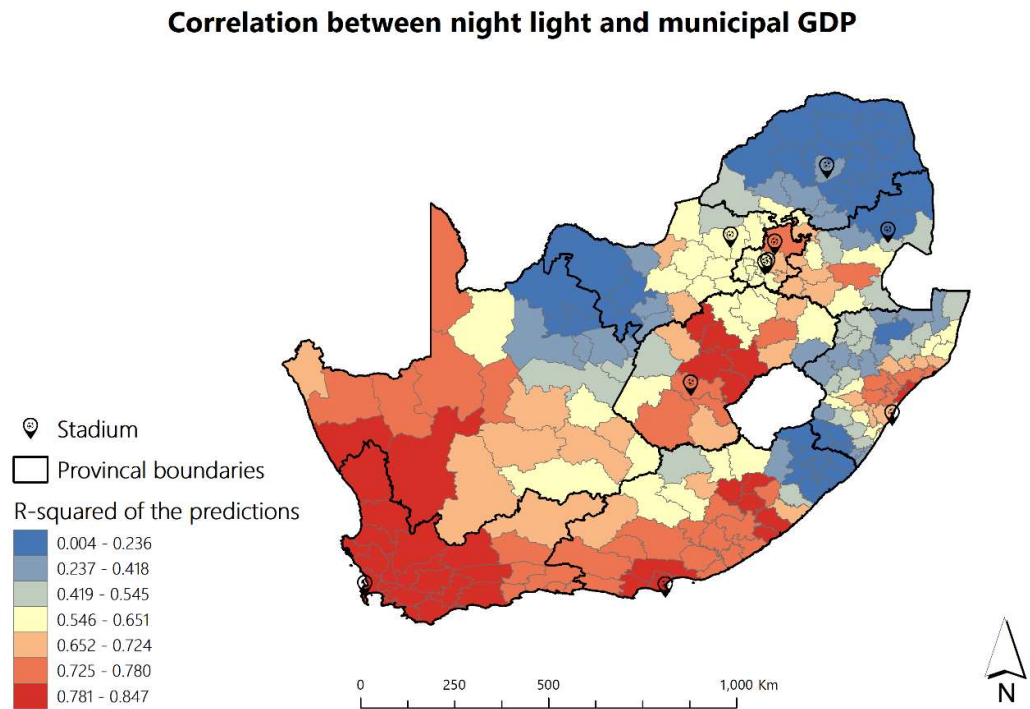


Figure 5-10 Prediction accuracy of the GWR model



GWR and OLS prediction performance

|                                 | OLS        | GWR          |
|---------------------------------|------------|--------------|
| Bandwidth (number of neighbors) | -          | 26           |
| R-squared                       | 0.5497     | 0.71462      |
| Sum of squared residuals        | 197.627515 | 95.541497388 |
| AICc                            | 628.532    | 562.1360     |



Table 5-6 Estimation of the difference of economic activities between host and non-host cities  
for every taken individually

|                     | (1)                                                                | (2)                                                                      | (3)                                                                      | (4)                                                                   | (5)                                                                      | (6)                                                                      |
|---------------------|--------------------------------------------------------------------|--------------------------------------------------------------------------|--------------------------------------------------------------------------|-----------------------------------------------------------------------|--------------------------------------------------------------------------|--------------------------------------------------------------------------|
| Dependent variables | <b>Total income</b>                                                |                                                                          |                                                                          | <b>Light</b>                                                          |                                                                          |                                                                          |
| in logarithm form   |                                                                    |                                                                          |                                                                          |                                                                       |                                                                          |                                                                          |
| Model description   | Estimation<br>between 2004<br>and 2013 with<br>control<br>variable | Estimation<br>between 2004<br>and 2013<br>without<br>control<br>variable | Estimation<br>between 1992<br>and 2013<br>without<br>control<br>variable | Estimation<br>between<br>2004 and<br>2013 with<br>control<br>variable | Estimation<br>between 2004<br>and 2013<br>without<br>control<br>variable | Estimation<br>between 1992<br>and 2013<br>without<br>control<br>variable |
| Host cities in 1992 |                                                                    |                                                                          | 0.200**<br>(0.0908)                                                      |                                                                       |                                                                          | 0.548***<br>(0.122)                                                      |
| Host cities in 1993 |                                                                    |                                                                          | 0.0934<br>(0.0920)                                                       |                                                                       |                                                                          | 0.415***<br>(0.118)                                                      |
| Host cities in 1994 |                                                                    |                                                                          | 0.0722<br>(0.0919)                                                       |                                                                       |                                                                          | 0.369***<br>(0.112)                                                      |
| Host cities in 1995 |                                                                    |                                                                          | 0.0324<br>(0.0843)                                                       |                                                                       |                                                                          | 0.265***<br>(0.0952)                                                     |
| Host cities in 1996 |                                                                    |                                                                          | -0.0160<br>(0.0716)                                                      |                                                                       |                                                                          | 0.161**<br>(0.0821)                                                      |
| Host cities in 1997 |                                                                    |                                                                          | 0.0203<br>(0.0908)                                                       |                                                                       |                                                                          | 0.263***<br>(0.0947)                                                     |
| Host cities in 1998 |                                                                    |                                                                          | -0.000348<br>(0.0816)                                                    |                                                                       |                                                                          | 0.189**<br>(0.0838)                                                      |
| Host cities in 1999 |                                                                    |                                                                          | -0.0380<br>(0.0763)                                                      |                                                                       |                                                                          | 0.122<br>(0.0763)                                                        |
| Host cities in 2000 |                                                                    |                                                                          | -0.0700<br>(0.0545)                                                      |                                                                       |                                                                          | 0.0212<br>(0.0567)                                                       |
| Host cities in 2001 |                                                                    |                                                                          | -0.0437<br>(0.0475)                                                      |                                                                       |                                                                          | 0.0402<br>(0.0537)                                                       |

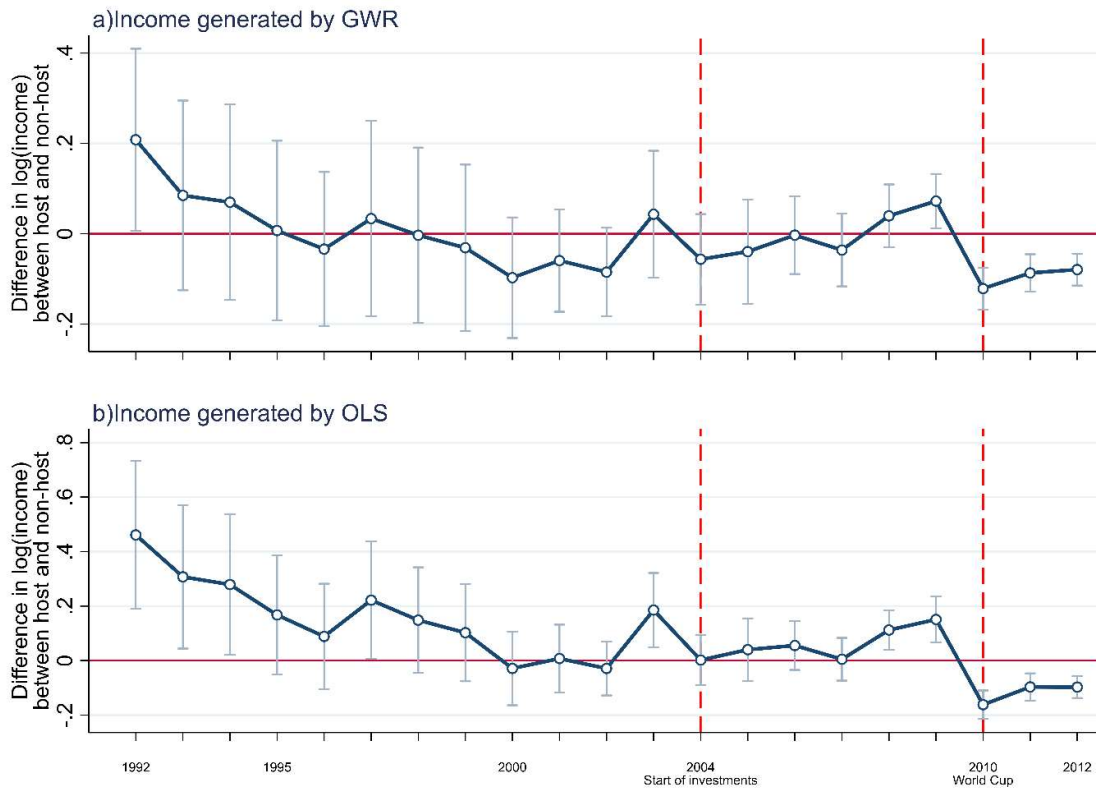
|                     |           |          |            |          |           |            |
|---------------------|-----------|----------|------------|----------|-----------|------------|
| Host cities in 2002 |           |          | -0.0689*   |          |           | 0.00297    |
|                     |           |          | (0.0401)   |          |           | (0.0430)   |
| Host cities in 2003 |           |          | 0.0295     |          |           | 0.194***   |
|                     |           |          | (0.0575)   |          |           | (0.0580)   |
| Host cities in 2004 | -0.0361   | -0.0744  | -0.0728*   | -0.0341  | -0.174*** | -0.0151    |
|                     | (0.0350)  | (0.0500) | (0.0403)   | (0.0455) | (0.0510)  | (0.0401)   |
| Host cities in 2005 | -0.0348   | -0.0574  | -0.0711    | -0.0195  | -0.135*** | -0.000471  |
|                     | (0.0358)  | (0.0386) | (0.0457)   | (0.0470) | (0.0479)  | (0.0484)   |
| Host cities in 2006 | 0.0159    | -0.0209  | -0.0204    | 0.00231  | -0.120*   | 0.0200     |
|                     | (0.0390)  | (0.0512) | (0.0329)   | (0.0516) | (0.0620)  | (0.0386)   |
| Host cities in 2007 | -0.0127   | -0.0539  | -0.0467    | -0.0432  | -0.171**  | -0.0197    |
|                     | (0.0390)  | (0.0523) | (0.0309)   | (0.0510) | (0.0683)  | (0.0326)   |
| Host cities in 2008 | 0.0496    | 0.0218   | 0.0213     | 0.0836   | -0.0633   | 0.0947***  |
|                     | (0.0528)  | (0.0550) | (0.0294)   | (0.0691) | (0.0647)  | (0.0335)   |
| Host cities in 2009 | 0.0819    | 0.0545   | 0.0566**   | 0.126*   | -0.0246   | 0.141***   |
|                     | (0.0510)  | (0.0538) | (0.0245)   | (0.0648) | (0.0563)  | (0.0356)   |
| Host cities in 2010 | -0.0599   | -0.139*  | -0.0905*** | -0.153** | -0.337*** | -0.143***  |
|                     | (0.0650)  | (0.0742) | (0.0181)   | (0.0744) | (0.0805)  | (0.0217)   |
| Host cities in 2011 | -0.0562   | -0.105   | -0.0859*** | -0.126*  | -0.272*** | -0.115***  |
|                     | (0.0547)  | (0.0654) | (0.0161)   | (0.0683) | (0.0743)  | (0.0239)   |
| Host cities in 2012 | -0.0444   | -0.0972  | -0.0686*** | -0.0960  | -0.273*** | -0.0964*** |
|                     | (0.0668)  | (0.0730) | (0.0126)   | (0.0859) | (0.0787)  | (0.0143)   |
| Host cities in 2013 | 0.0226    | -0.0179  | -          | 0.00170  | -0.176**  | -          |
|                     | (0.0639)  | (0.0688) | -          | (0.0835) | (0.0753)  | -          |
| Cropland areas      | -0.0717** |          |            | -0.0719* |           |            |
|                     | (0.0294)  |          |            | (0.0369) |           |            |
| Urban areas         | 6.153     |          |            | 18.08    |           |            |
|                     | (32.65)   |          |            | (34.17)  |           |            |
| Total population    | 55.77     |          |            | -49.37   |           |            |
|                     | (73.35)   |          |            | (113.5)  |           |            |
| Constant            | 20.92***  | 20.23*** | 20.74***   | -0.293   | -0.694*** | -0.0937    |
|                     | (1.839)   | (0.0888) | (0.132)    | (1.828)  | (0.114)   | (0.161)    |

|                     |           |           |           |           |           |           |
|---------------------|-----------|-----------|-----------|-----------|-----------|-----------|
| Observations        | 1,521     | 1,521     | 2,574     | 1,521     | 1,521     | 2,574     |
| R-squared           | 0.562     | 0.537     | 0.548     | 0.573     | 0.549     | 0.567     |
| Number of _ID       | 117       | 117       | 117       | 117       | 117       | 117       |
| Control variables   | Y         | Y         | Y         | Y         | N         | Y         |
| Observations period | 1992-2013 | 1992-2013 | 1992-2013 | 2001-2013 | 1992-2013 | 1992-2013 |
| Cities FE           | Y         | Y         | Y         | Y         | Y         | Y         |

---

Block bootstrapped standard errors with 1000 repetitions in parentheses \*\*\* p<0.01, \*\* p<0.05, \* p<0.1

Figure 5-11 The difference in income growth between host cities and other cities from 1992 to 2013



**Note:** This graph shows the estimate of the difference in logarithm of income (a) and night light (b) between host cities and other cities from 1992 to 2012 carried out through Equation [9] while however excluding the control variables because they are available from 2001. The 2013 dummy variable is set to zero. The standard error is block bootstrapped with 1000 repetitions and confidence interval around the estimates represents a 95 statistical level.

Table 5-7 Stadiums and venues

| FIFA World Cup<br>host cities | Stadiums                            | Cost (USD)* | New<br>Stadium | Costliest<br>stadium |
|-------------------------------|-------------------------------------|-------------|----------------|----------------------|
| Johannesburg                  | FNB stadium / Ellis Park<br>Stadium | 267,372,010 |                | Yes                  |
| Pretoria                      | Loftus Versfeld Stadium             | 7,776,373   |                |                      |
| Polokwane                     | Peter Mokaba Stadium                | 87,906,832  | Yes            | Yes                  |
| Nelspruit                     | Mbombela Stadium                    | 72,472,828  | Yes            |                      |
| Bloemfontein                  | Free State Stadium                  | 25,019,636  |                |                      |
| Rustenburg                    | Royal Bafokeng Stadium              | 20,286,192  |                |                      |
| Port Elizabeth                | Nelson Mandela Bay Stadium          | 139,554,072 | Yes            | Yes                  |
| Durban                        | Moses Mabhida Stadium               | 20,9623,984 | Yes            | Yes                  |
| Cape Town                     | Cape Town Stadium                   | 304,292,880 | Yes            | Yes                  |

**Source:** Easmonn, M. and Trish, C. (2015) 'The rocky road to legacy: Lessons from the 2010 FIFA World Cup South Africa stadium program', *Project Management Journal*, 46(3), pp. 88–107. doi: 10.1002/pmj.

## *Appendix 5 - Section B: SCM results*

Our study focuses on the results of the difference-in-difference over the SCM for two reasons. First, it provides the ability to account for migrations that occurred across cities as a result of the World Cup by controlling for population variations and thereby avoiding so-called substitution bias (Coates and Humphreys, 2002; Pillay and Bass, 2008). Second, the panel data allows us to compare income variations after 2004 across specific human activities accounted for as land uses.

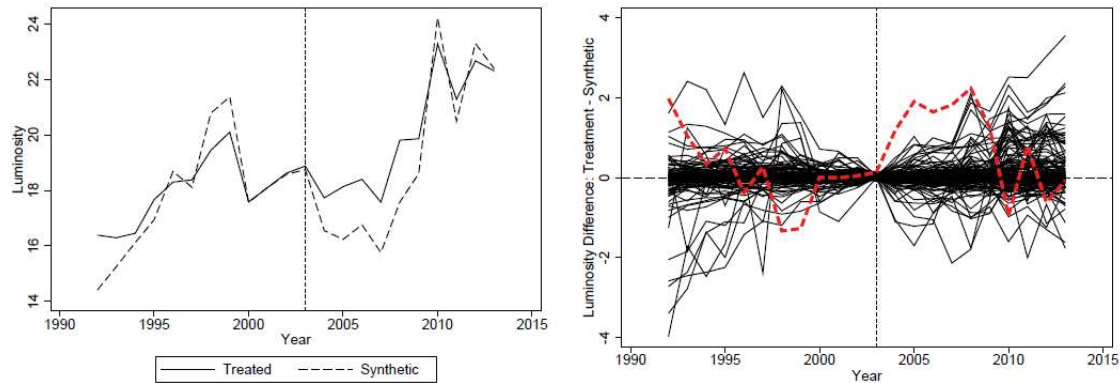
Nonetheless, here, I estimate the SCM with night light data as the dependent variable to understand how the account of urban areas and population, participate in the difference between the results and that of (Pfeifer et al., 2018). First, I attempt to replicate the results of (Pfeifer et al., 2018) with the same variables as theirs (Figure 5-12). I use data on municipalities area, elevation, the proportion of the white population (instead of the share of the indigenous population), the proportion of tertiary education, index of soil quality, and the unemployment rate for this purpose. Similar to (Pfeifer et al., 2018), I exclude the cities neighboring host cities to account for spatial correlation in the donor pool. The result of the replication is presented in Figure 5-13. Unfortunately, I am not able to exactly produce the same results as the one of the previous study since Figure 5-13 displays a growth in light in host cities that is lower than what was described by (Pfeifer et al., 2018). I believe this is due to the difference in some data sources, the use of the share of the white population instead of the share of indigenous

populations as a predictor. Although it is less significant than in the results of the previous study, I can notice a higher night light in host cities relative to that of the counterfactual cities between 2004 and 2008.

Then, I add population as well as urban, and cropland areas as factors necessary to construct the synthetic control cities in Figure 5-14. The results show that adding these three variables, improve the construction of the synthetic control unit in the period anterior to 2003. Furthermore, adding these variables in the construction of the synthetic control diminish the previously estimated gap in night light between host cities and their counterfactual units between 2004 to 2010.

Therefore, I take the result of this exercise as evidence of the importance of omitted variables when using night light satellite data for economic analyses. I believe that the identification of the counterfactuals of cities can be improved by the account of their levels of population, as well as urban and rural areas. Overall, the estimates from SCM fail to reject the hypothesis of no difference in light between host cities and their counterfactuals.

Figure 5-12 SCM results from (Pfeifer et al., 2018)

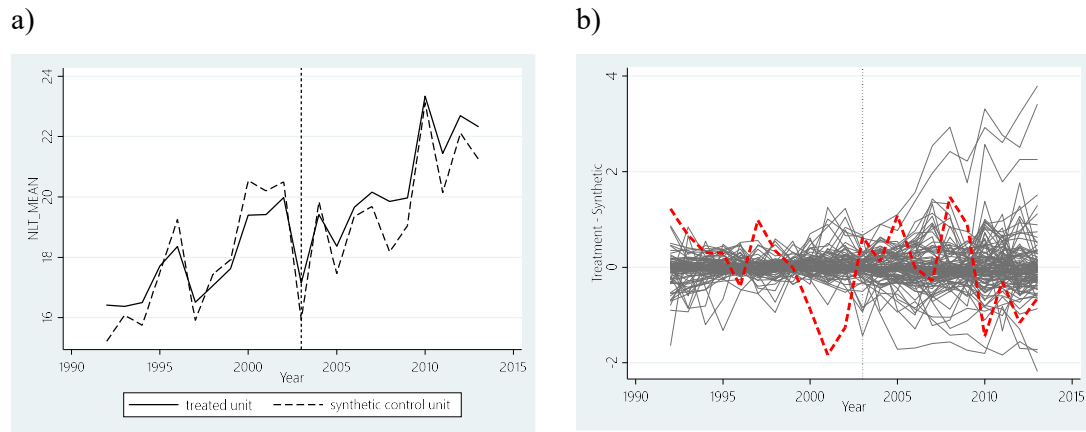


**Source:** Pfeifer, G., Wahl, F., Marczak, M., 2018. Illuminating the World Cup effect: Night lights evidence from South Africa. *Journal of Regional Science* 58, 887–920.

| Predicator                                     | Treated  | Synthetic |
|------------------------------------------------|----------|-----------|
| Average night light (1992(1)2003)              | 18.8757  | 18.74631  |
| Area                                           | 3722.812 | 1111.97   |
| Distance to Railway                            | 0        | 0         |
| Elevation                                      | 907.6775 | 332.8766  |
| Proportion white of people                     | 0.033037 | 0.033079  |
| Proportion of population with higher education | 0.066015 | 0.057515  |
| Index of soil quality                          | 39.97758 | 49.94023  |
| Unemployment rate                              | 37.65556 | 40.74773  |



Figure 5-13 Replication of SCM results from (Pfeifer et al., 2018) with the data



In Figure 5-13b, placebo test for night light, one line for each municipality in the donor pool.

Host cities in red

The weights of cities in the synthetic control unit of host cities

| City          | Weight |
|---------------|--------|
| The Msunduzi  | 0.573  |
| uMhlathuze    | 0.339  |
| Steve Tshwete | 0.089  |

| Predictor                                      | Treated   | Synthetic |
|------------------------------------------------|-----------|-----------|
| Average night light                            | 17.70693  | 17.72075  |
| Area                                           | 21.39318  | 33.54259  |
| Distance to Railway                            | 0         | 0         |
| Elevation                                      | 885.3843  | 696.4918  |
| Proportion white of people                     | 0.1313636 | 0.078529  |
| Proportion of population with higher education | 0.1659091 | 0.137373  |
| Index of soil quality                          | 39.69958  | 39.00272  |
| Unemployment rate                              | 19.26     | 28.15341  |

Figure 5-14: Account of population, urban and cropland areas in SCM analysis



In Figure 5-14b, placebo test for night light, one line for each municipality in the donor pool.

Host cities in red

City weights in the synthetic

| City         | Weight |
|--------------|--------|
| Buffalo City | .134   |
| uMhlathuze   | .866   |

| Predicator                                     | Treated  | Synthetic |
|------------------------------------------------|----------|-----------|
| Average night light                            | 17.70693 | 17.71663  |
| Area                                           | 153.2428 | 39.72011  |
| Elevation                                      | 885.1419 | 843.8148  |
| Proportion white of people                     | 0.135546 | 0.140138  |
| Proportion of population with higher education | 0.173636 | 0.151924  |
| Index of soil quality                          | 40.21996 | 42.55794  |
| Unemployment rate                              | 25.635   | 20.32562  |
| Urban built-up areas                           | 9.82E+07 | 1.01E+08  |
| Cropland areas                                 | 4.27E+08 | 4.68E+08  |

## **Chapter 6 Differences in water policy efficacy across South African water management areas**

### **6.1. Introduction**

During the colonial and apartheid eras, water policies in South Africa mainly targeted white-dominated areas. Since the end of the apartheid, the South African government has striven to balance this historical inequality with several political reforms (Constitution of the Republic of South Africa, 1996; Funke et al., 2007). However, at a provincial level, Gauteng and Western Cape, the largest former white-dominated areas and the wealthiest provinces in the country, appear to have significantly better water quality and services than the remaining provinces (Funke et al., 2007; Hopolang & Dennis, 2018).

Nonetheless, despite this positive relationship between water quality and economic development at a provincial level, eutrophication due to the phosphorus industry and issues related to acid mine drainage (AMD) raise concern about the spatial distribution of water pollution within each province. Furthermore, studies from a provincial approach may mask the differentiated reality of some townships that are more economically developed than rural areas. For instance, in the case of Khayelitsha and Soweto, near Cape Town and Johannesburg respectively, living standards and water quality are low due to several unauthorized and informal settlements (Goldblatt, 1999; Makaudze & Gelles, 2015). Additionally, province-level research cannot be used to tackle the economic disparity and/or the legacies of racial injustices because within each province, there is considerable diversity in economic prosperity across

areas related to former apartheid and colonization. It is therefore vital to conduct fine-scale level research to bring to light hidden inequalities in water policies. However, despite the existence of some region-specific studies in South Africa, difficulty in attaining fine-scale data makes this type of study scarce at a national level.

This study attempts to reveal the relationship between economic development and surface water quality in South Africa at a water management area level nationwide (Figure 6-3 in the Appendix). The water management area is defined as the unit area in which the water policies are implemented. Considering that some previous policy attempts to adjust for the apartheid and colonization legacies of injustice and economic injustice have not been successful (DWS, 2018), I believe that such an investigation will allow policymakers in South Africa to evaluate the effectiveness of their policies at the smallest level of implementation. Based on these legacies of apartheid at a local level and the concentration of wealth in former white-dominated areas, I assume that better water policies and their enforcement are positively related to economic development (this assumption is discussed in the next section). Thus, I formulate the hypotheses that (1) on average, economic development in South Africa improves surface water quality and (2) the impact of economic development on water quality is not related to former apartheid segregation.

In this study, I focus on surface water, because it accounts for 77% of the water available for human uses, government and nongovernmental organizations have issued water quality

policies in this area (Feris & Kotzé, 2014; Schreiner & Van Koppen, 2003; Turpie, Marais, & Blignaut, 2008). Then, since there is limited access to data to represent the economic development variable at a fine-scale level, I use the average night-time light and its growth as a proxy of the average economic level and the proxy of economic growth. Additionally, I assume that the land uses per area give an index of the nature of water polluting-activities in each region, allowing us to interpret the night-time light data variables as the magnitude of the economic development not related to the human activities of a certain region. A geographically weighted regression (GWR) was applied to graphically detect the water management areas wherein there might be a positive association between water pollution and either the average economic level or economic growth. Subsequently, I use a spatial Durbin model (SDM) to assess the statistical significance of the GWR results.

The results of the estimations imply a dual impact of economic development on water pollution: a negative impact in the eastern provinces of the country and a positive impact in the western provinces, highlighting a weaker efficacy of water policies in the latter provinces. Moreover, the results show that former white-dominated areas during apartheid have better water policies, suggesting a legacy of the apartheid unfair water policies.

This study relates to economic literature claiming a dual causal effect of economic development on water quality, as it can either improve or degrade water quality. It mentions the environmental Kuznets curve (EKC) theory where countries considerably pollute their

environment at the beginning of their development, but, after a certain level of per capita income, they improve this environment through significant technological advances. To our knowledge, despite evidence that this dual impact may apply at a fine scale level (Anderson, Romani, Phillips, Wentzel, & Tlabela, 2007; Goldblatt, 1999; Xu et al., 2019), no economic studies are carried at such level nationwide, because of the lack of economic variable at such level.

Therefore, I first contribute to the economic literature by using satellite night light data as a proxy of economic development at a fine scale for the assessment of its impact on water quality. Second, unlike province level study, the fine scale level of my study enables us to disentangle the effect of land use from economic development in estimating the relationship between economic development and water quality.

Furthermore, I combine geographical and environmental researchers' techniques that estimate the impact of human activities on water quality at a fine-scale level (Giri & Qiu, 2016; Mainali, Chang, & Chun, 2019; Tu & Xia, 2008) to the economic literature. It consists in the estimation of the spatial distribution of the relationship between economic development and water quality. Indeed, within countries, evidence suggest that although economic activities might cause water pollution (Xu et al., 2019), citizens and municipalities in economically developed areas are more concerned about the quality of life and thus demand improvement in water management (Anderson et al., 2007; Goldblatt, 1999). It also consists in the account for

upstream river pollution in the estimation of the local relationship between economic development and water quality in each quaternary area. Previous researches highlighted that not including this spatial dependence were source of bias estimates in this relationship (Chakir & Le Gallo, 2013; Q. Li et al., 2014; E. Wang et al., 2018).

The remainder of the chapter is structured as follows: the second section describes the background of South African water policy and the relationship between economic development and effective water quality; the third section details the methodology; the fourth section provides the results of the estimations of the GWR and SDM; and the fifth section presents the political implications and conclusion.

## **6.2. South African water policies and local economic development**

During the apartheid, economic development and its benefits (including water policies) were divided between two populations separated in two areas. On one side, there were the economically developed white dominated areas, that enjoyed water resources and services due to effective government actions. On the other side, there were the areas where lived most black, and partially the colored population, characterized by the absence of economic development, and by the absence of government actions, and water and sanitation services (Beck, Rodina, Lucker, & Harris, 2016).

Since their first democratic election in 1994, South Africans put an end to this political system and strive to eradicate racial inequalities. An important measure taken for this aim is Part B of Schedule 4 of the *Constitution*, calling for local governments to provide water and sanitation services, including the potable water supply systems and domestic wastewater and sewage disposal systems (Feris & Kotzé, 2014). For the enforcement of this article, (local and district) municipal managers, provincial and national managers and water drainage officials cooperate to provide these services, while being supported by equitable shares of the national water sanitation budget across provinces (Beck et al., 2016; Hopolang & Dennis, 2018). In this arrangement, the overall policies are taken at a national level, while specific roles are allocated among different line functions (Beck et al., 2016; Feris & Kotzé, 2014). The goal of the government is to enable all areas of the country, independently of its level of development to enjoy sufficient water services.

Nevertheless, despite these nationwide water policies, the wealthy Gauteng and Western Cape provinces (hosting the largest white population during apartheid) still have better water supply and sanitation services, and have better wastewater treatment plants than other provinces (DWS, 2013; Hopolang & Dennis, 2018). The main factor maintaining this inequality at a provincial level seems to be related to the inability of the Government to provide cost effective infrastructure in less developed areas (Adewumi, Ilemobade, & Van Zyl, 2010; DWA, 2008). According to the South African ministry of sanitation, most of capital investments



in current pipes, dams and other water-related infrastructures were made before the 1990's (Beck et al., 2016; DWA, 2008). Therefore, the development of new infrastructure in areas with no water infrastructure during the Apartheid is costly. Also, citizens in less developed areas may not sustain an eventual construction of performant water infrastructure as these citizens cannot provide enough resources (taxes), enough electricity and qualified personal for its maintenance (Ahjum et al., 2018; DWA, 2008). Finally, these inequalities across provinces seem to perpetuate within the wealthiest provinces due to cost-recovery constraints, as the townships that are less wealthy than the major cities, have poor water policies (Beck et al., 2016; Smith & Hanson, 2003).

Along with this persistent relationship between local economic development and effective water policies, the agricultural, manufacturing and mining industries act as sources of development that degrade water quality. In this regard, despite government policies and regulations, which hold polluting industries responsible to clean up after their waste, some industries escape these policies because of the government weak enforcement. For instance, in 2012, a legal case opposed the Federation for Sustainable Environment against the Minister of Water Affairs<sup>h</sup>, where residents in (economically small) towns of Carolina, Caropark and Silobela suffered from an acid mine contamination in their surface water. The water was not appropriately treated, either by the mining industry causing the AMD or by the municipality

---

<sup>h</sup> <http://www1.saflii.org/za/cases/ZAGPPHC/2012/128.html> on 06/17/2019

responsible for the water treatment, leading to supply to improper water for consumption in these towns (Feris & Kotzé, 2014). These ineffective regulations could be the consequences of the lack of co-ordination between the national, provincial, and municipal powers in resolving water issues (Adler, Claassen, Godfrey, & Turton, 2007; Feris & Kotzé, 2014). In this context, some private industries generate economic development (or growth) that pollute surface water, leading to significant risk in the municipalities where they occur, and in the nearby politically (and economically) weak municipalities.

In sum, two facets prevail in the assessment of the efficacy of water policy in South Africa. On one hand, local economic development is an important predicator for the implementation of effective water policies because, eventually of the Apartheid legacies, and of the cost-constraints related to the settlement of performant water amenities in non-economically developed areas. On the other hand, dynamic economic activities are sources of pollution because the Government fail to enforce an effective polluter-pays principle. I take advantage of these links between economic development and water policies in the pose the assumptions, stating that economic development can either be related to water quality through efficient water policies or it can relate to water pollution through water polluting activities. I attempt to disentangle these two effects in the model specification and in the definition of the proxy variables.

## 6.3. Methodology

### 6.3.1 Geographically weighted regression

I use a GWR as an exploratory model to account for spatial heterogeneity and the spatial dependency in the relationship between economic development and water quality across water management areas. GWR is a nonparametric regression technique that extends the scope of the average relationship techniques by allowing for variation of the parameter in space (Brunsdon, Fotheringham, & Charlton, 2010; LeSage, 2013). I use this model because 1) several studies highlight its strength over a global analytic approach (approaches estimating the aggregate relationship in space) in capturing the variation in water quality caused by human activities (Giri & Qiu, 2016; Tu & Xia, 2008); and because 2) the nature of economic activities across provinces in South Africa differs over space.

A GWR is calibrated by performing a local (relative to each observation of the sample) ordinary least squares (OLS) regression weighted by the neighboring observations to assess the direct relationship between the dependent variables and the independent variable in an area. In a matrix form, it can be expressed as follows:

$$\mathbf{Y} = (\boldsymbol{\beta} \otimes \mathbf{X})\mathbf{1} + \boldsymbol{\varepsilon} \quad (6.1)$$

where  $\mathbf{Y}$  is the matrix of dependent variables (water pollutants), and  $\otimes$  is the logical multiplication operator, in which each element of  $\boldsymbol{\beta}$  (representing the estimate) is multiplied by the corresponding element  $\mathbf{X}$ , the matrix of independent variables (the level of development

and economic growth). Both  $\beta$  and  $X$  are  $n \times (k+1)$  matrices, and  $\mathbf{1}$  is a  $(k+1) \times 1$  vector of 1 s.

The matrix  $\beta$  consists of  $n$  sets of local parameters that correspond to the value of the estimates for each independent variable (at the columns of the matrix) for each observation (at the row of the matrix). The parameters at each row (each observation) of the matrix  $\beta$  are estimated by the equation:

$$\hat{\beta}_i = [X'W_iX]^{-1}X'W_iY \quad (6.2)$$

where  $W_i$  is the weight matrix at the point  $i$ . For each unit  $i$  of the sample, the parameter is calibrated by weighting the regression results of different observations from this point  $i$  using the weight matrix  $W_i$ . The weight matrix  $W_i$  follows a Gaussian curve constructed by an inverse Euclidean distance technique, allowing the observations closer to the regression point to weigh more than those that are farther away.

The calculation of the weight matrix includes a limit adaptive bandwidth for each point of estimation, allowing the observations with relatively fewer neighboring observations to have a wider bandwidth and observations with a high density of neighbors to have a lower bandwidth. I use an adaptive bandwidth as it better suit data in general (Brunsdon et al., 2010). The adaptive bandwidths for each observation were estimated by minimizing the corrected Akaike information criterion<sup>i</sup>. The weight matrix at observation  $i$ , expressing its relationship with observation  $j$ , is described as follows:

---

<sup>i</sup> As robustness estimates, we also used a cross-validation technique to determine the bandwidths of the weight matrix, because the definition of the weight of the GWR plays an important role in the results of the

$$w_{ij} = \begin{cases} \exp\left(-\frac{1}{2} \left(\frac{d_{ij}}{b_i}\right)^2\right) & \text{if } d_{ij} \leq b_i, \\ 0 & \text{elsewhere} \end{cases} \quad (6.3)$$

where  $w_{ij}$  is the weight of observation  $j$  at observation  $i$ ,  $d_{ij}$  is the Euclidean distance between  $i$  and  $j$ , and  $b_i$  is the bandwidth at observation  $i$ .

The coefficient  $\hat{\beta}_i$  represents the estimate of the relationship between the economic variable and water pollution indices at the observation  $i$ . Since water quaternary management areas are the smallest units of policy implementation, I interpret a negative relationship as an effective water quality regulation policies and/or legacies of the effective water policies implemented during the apartheid era, while a positive association may indicate poor or ineffective policy (Herselman, Steyn, & Snyman, 2006; Nleya, 2008; Torras & Boyce, 1998).

### 6.3.2 *The spatial Durbin model*

Consequently, I use a spatial econometric approach to account for the effect of the spatial relationship between regions because economic and pollution spillover effects of one area on its neighboring areas are very plausible. The spatial economics literature embodies several models; of which I apply the SDM for two reasons.

First, I believe that water pollution in an area may be a factor of water pollution in neighboring areas through streams (Lintern et al., 2018). To estimate the local relationship

---

of GWR. Using the cross-validation method provides similar spatial distribution of the relationship between economic development and water quality (see Figure 6-9 and 6-10 in the Appendix).

between economic development and water pollution in certain areas, we need to control for this neighboring effect. Thus, I account for this relationship by including a spatially autocorrelated term of the dependent variable (water pollutants). To portray this relationship, I assume that the surface water pollution from one area to another is related to their immediate proximities and is related to the average river gradient in each area. Accordingly, I use a contiguity weight matrix ( $W_1$ ) because it considers areas that have a common border as neighboring observations. Along with the spatial matrix, I include the average river order code in each water area as a proxy of the river gradients; this code portrays the relationship between a river and its affluent as it is an integer whose value increases when the river flows downstream (Mettetal, 2019).

Second, I believe that economic development in one area may be a factor of water improvement in neighboring areas (Amin, 2016). Quaternary water management areas are the smallest areas of policy implementation, but water policy decisions are made at the local and district municipal levels as well as at the hydrological, provincial, and national levels (see the preceding section and (Beck et al., 2016; Feris & Kotzé, 2014)). Hence, I assume that economic development in a water management area influences the decision-making at different administrative and hydrological levels. Therefore, I use the spatial lag of the economic variables to estimate this relationship. I use the Euclidean distance weight matrix ( $W_2$ ) for the economic development variables to portray a relationship in which one water management area

may affect the water policies of distant water areas when the areas are, for example, in the same district, the same province or the same basin.

The SDM in the analysis is expressed as follows:

$$\mathbf{y} = \rho \mathbf{W}_1 \mathbf{y} + \alpha \mathbf{I}_n + \beta \mathbf{X} + \mathbf{W}_2 \mathbf{Z} \theta + \boldsymbol{\varepsilon}, \quad (6.4)$$

where  $\mathbf{y}$  is an  $n \times 1$  vector representing the dependent variable (expressed as the logarithm of the water pollutants),  $\rho$  is the endogenous spatial coefficient of the dependent variable, and  $\mathbf{W}_1 \mathbf{y}$  is an  $n \times 1$  vector of spatial lag of the dependent variable.  $\mathbf{X}$  is an  $n \times (k+1)$  matrix describing the independent variables (including the level of economic development and the control variables), and  $\theta$  illustrates the exogenous spatial interaction coefficient associated with the lagged independent variables  $\mathbf{W}_2 \mathbf{Z}$ . In the current case,  $\mathbf{Z}$  concerns only the economic development variables. The term  $\boldsymbol{\varepsilon}$  is an  $n \times 1$  vector representing the error term. Finally,  $\mathbf{W}_1$  and  $\mathbf{W}_2$  are  $n \times n$  spatial weight matrices<sup>j</sup>.

To overcome the identification issue inherent to estimating the spatial lag of the dependent variable ( $\rho$ ), I calibrate the SDM using the two stages least squares technique (Gibbons & Overman, 2012). The average river order, the average area and the average runoff per quaternary areas are used as instrumental variables to estimate the spillover pollution ( $\rho$ ) from one area to another. I believe that these variables are exogenous because 1) they are not directly related to pollution in neighboring areas and because 2) they could only affect the

---

<sup>j</sup> See Pre-estimation test on Figure 6-6 and 6-7 in the Appendix

neighboring areas level of pollution through their effects on the local area's level of pollution. Additionally, I consider them to be relevant instruments because previous studies show their impact on the estimation of the local water quality (Abildtrup, Garcia, & Stenger, 2013; Duflo & Pande, 2007; Mainali et al., 2019; Mettetal, 2019).

Because the economic variables have two impacts ( $\beta_k$  and  $\theta$ ) on water quality and because there is a water pollution effect from neighboring areas, the coefficients of the economic variables cannot be interpreted as a marginal effect (which I refer as “impact” throughout the analysis). Hence, the direct impact of the economic variables is expressed by the following equation:

$$\frac{\partial y_i}{\partial X_{ik}} = S_k(W)_{ii} \quad (6.5)$$

The indirect impact of the economic variables can be expressed by the following equation:

$$\frac{\partial y_i}{\partial X_{jk}} = S_k(W)_{ij} \quad (6.6)$$

With  $S_k(W)_{ii}$  and  $S_k(W)_{ij}$ , respectively, the  $ii$  and the  $ij$  elements of the square matrix of parameters  $S_k(W)$  that is estimated as follows:

$$S_k(W) = (I_n - \rho W_1)^{-1}(I_n \beta_k + W_2 \theta) \quad (6.7)$$

Finally, the readings of SDM results require two tables because, as mentioned above, the table coefficient of the regression does not inform on the impact of the variables. In this study, I report the results of the different impacts of the economic variables in the main text and provide the table of the other coefficients in the supplementary materials.



## 6.4. Data

I focus on surface water because the pollution of surface water is a direct risk to human health as it is a source for drinking (directly or after additional treatment), recreation, and agriculture (Ashton, Love, Mahachi, & Dirks, 2001; Dikio, 2010; Hopolang & Dennis, 2018). I use Total Dissolved Solids (TDS), nitrogen, phosphorus and sulfate concentrations as water pollution indices. These are meaningful water indices as the levels of TDS, nitrogen, and phosphorus can be interpreted to estimate the degree of eutrophication (Best, 2019), and the TDS and phosphorus concentrations indicate the level of acid mine damage (Akcil & Koldas, 2006). Finally, concentrations of TDS and sulfate can pose a serious risk for epidemic disease (Chien, Robertson, & Gerrard, 1968; World Health Organization, 2017).

I collect the water pollutant indices dataset provided by the Water Research Commission (WRC - <http://waterresourceswr2012.co.za/>) and constructed by the Department of Water and Sanitation of South Africa (DWS). Broadly, South Africa is divided into 22 basins, which themselves are subdivided into 1845 quaternary areas for the management of water resources. However, I restricted my sample to 305 quaternary areas as I dropped quaternary areas with no water quality stations during the sample period (2005 to 2010) and I dropped areas with several gaps in the monthly data collection (Figure 6-3 displays the sample location)<sup>k</sup>.

---

<sup>k</sup> The sample provided by the WRC included the level of Ammonium, Fluoride, and pH. We did not

The sample in each quaternary area consists of the quality of surface water in rivers, dams (generally over 1 million m<sup>3</sup> but also including farm dams), and lakes. However, the WRC only provides the median value of the water pollution index per quaternary areas to the public. Nevertheless, in spite of the reduction of explaining power when using a single aggregate value instead of a time-trend water indices, median values of water indices still perform well in explaining the relationship between water quality and anthropogenic activities (Chang, 2008; Kang et al., 2010; Peterson, Sheldon, Darnell, Bunn, & Harch, 2011; Tu & Xia, 2008).

Regarding the economic development variables, I collected night-time light raster images of the National Geophysical Data Center (NOAA) - Defense Meteorological Satellite Program (DMPS/OLS). This dataset is a yearly average of daily satellite images of human-made night-time light at 2.3×2.3 km resolution. The data provided by NOAA are corrected for disturbances such as scanning errors, errors in mapping and natural (nonhuman made) sources of light. The unit of the night-time light are the digital number, varying from 0 to 63, the average of which I calculated in each quaternary zone.

Several studies have highlighted the strength of using night light as a proxy of the economic differences within countries (Henderson et al., 2012). This relationship is based on the fact that places with large economic activities are more illuminated in the night so that there

---

include the level of Ammonium, Fluoride in our sample because they had significantly less data than the indices used in the study. Concerning the pH, the interpretation of its value would be less clear than the index chosen in this study, because a too low value and a too high value of pH can be both indices of human pollution.

is a positive relationship between GDP and night light digital number, which can be expressed as follows:

$$\log(NLT_i) = \gamma \log(x_i) + \mu_i, \quad (6.8)$$

where  $x_i$  represents economic development,  $NLT_i$  is the average night-time light digital number of a water management area  $i$ ;  $\mu$  is the stochastic error term, and  $\gamma$  is a positive real number representing the parameter of this relationship. However, estimating how the inverse of the parameter  $\gamma$  can be used to estimate a monetary economic development is beyond the scope of the study. Hence, I estimate the relationship between water quality and the natural logarithm of night light digital number in each quaternary area.

In addition, building on the EKC literature, I believe that economic growth has an importance in the relationship between the average economic status in 2010 and the median value of water pollution in the period 2005 to 2010. Therefore, I include the natural logarithm of the ratio of night-time light in 2010 to night-time light in 2005 (light growth) as a proxy of economic growth during the water sample collection period.<sup>1</sup>

Regarding the land uses data, I utilize the agriculture, mine, industry, and urban settlements as anthropogenic sources of pollution. Indeed, evidence show that pollutants in surface water are primarily found as a result of mining, industry (particularly the phosphorus

---

<sup>1</sup> We use the characteristic of the log function  $\log(a/b) = \log(a) - \log(b)$ . Therefore, the coefficient of night light growth will refer to the difference between logarithm of night in 2010 minus this in 2005. For this estimation, we assume a straightforward comparison of the average night-time light digital number from year to year.

and gold industries in South Africa), agriculture, and urban waste (Best, 2019). Regarding their relationships with the water pollutant indices I used, mining is known to be linked to total dissolved solids (TDS) and phosphorus (Ashton et al., 2001). While, agriculture is linked to TDS, nitrogen, and phosphorus (Hess et al., 2016; Verheyen et al., 2015). Industry is linked to TDS, sulfate, and phosphorus. Urban spaces are linked to TDS and sulfates. Although, each of these pollutants may be found across land uses depending on the specific activities present.

I collect the land use data as raster format data at a 30×30 m resolution that were provided by the WRC for the year 2011. The impact of land use on surface water quality is estimated by inverse distance-weighted metrics of the land use areas relative to the nearest water bodies in each water management area. Hence, land uses closer to water bodies have a stronger impact on the potential pollution in these water bodies. The expression of the distance-weighted land surface in each area is as follows:

$$\text{Weighted Land Use} = \sum_{i=1}^n I_k W_i, \quad (6.9)$$

where  $I_k$  is the raster cell of a given land use  $k$ , and  $W_i$  is the weight estimated via a Euclidean distance to water bodies with  $(d + 1)^{-1}$ . Land use estimation with this weight yields a higher linear relationship with water pollutant indices than the average area of land use in a given area (Peterson et al., 2011). Additionally, this metric allows for proportionate estimations of the impact of each land use across each water quaternary on water pollution. Finally, this metric simplifies the SDM model by assuming that pollution caused by a land use is carried to

neighboring areas' rivers or dams only through their pollution in local rivers or dams. Further, I include the population density, as a final anthropogenic index, collected via the WRC for the year 2011 as areas with large population tend to be more polluted (Tu & Xia, 2008).

Finally, I acknowledge that land use and population density are collected for the year 2011, while economic variables and water pollution indices are collected for the years 2005 and 2010. However, due to the lack of data, I use these variables as proxies of the location of human activities and population density in 2010. The summary statistics of the data are presented in the Table 6-5 in the Appendix.

## **6.5. Results and discussion**

### *6.5.1 The geographical distribution of the relationship between economic development and water quality*

The prediction of the GWR based on the sample data is displayed in Figures 6-1 and 6-2. Figure 6-1 displays four maps describing the relationship between the average level of development in 2010 and water pollutants, and Figure 6-2 shows the relationship between the economic growth from 2005 to 2010 and water pollutants. In each figure, the pollutants described are sulfate, nitrogen, TDS, and phosphorus concentrations<sup>m</sup>.

---

<sup>m</sup> The accuracy of fit of the GWR is presented in Table 6-6 and 6-7 in the Appendix.

I first describe the relation between the level of development in 2010 and water pollutants. The relationships between sulfate and nitrogen and the level of development seem to be the lowest in the middle of the Northern Cape province and the highest in the provinces of Free State, Mpumalanga, and Limpopo (The provinces name are displayed in Figure 6-3 and 6-4). There is a strong positive relationship in the northeast of the country, including cities such as Johannesburg and Pretoria. The spatial relationship between phosphorus and level of development is similar to the relationship of the above-described indices, except that the negative association is highest at the boundary between the Northern Cape and Eastern Cape provinces. A larger area of the country suffers a positive relationship between pollution and level of development, still concentrated in the northeast.

Finally, the negative association between TDS and economic development is highest in Western Cape province around Cape Town, and the highest positive values are in the eastern part of the country, covering the whole of Limpopo, Mpumalanga, Gauteng, Kwazulu Natal and North West provinces, as well as in large parts of Free State province. However, compared to the other water indices, the range of variation of the relationship between average economic development and TDS is smaller across the country, and this relationship is almost homogeneous in space.

Moving to the impact of economic growth (2005 – 2010) on water quality, I note a large range of variation and several fluctuations in space of the coefficients across and within

provinces. The association between economic growth and sulfate, phosphate, and TDS concentrations are similar to each other in the sense that they are strongly negative in Western Cape and North Cape provinces, including the city of Cape Town and part of Port Elizabeth. However, these relationships are highly positive in the eastern and southern parts of South Africa, particularly in the provinces of Eastern Cape, Limpopo, and Mpumalanga. These positive relationships appear around the cities of Johannesburg, Pretoria, and Durban without including them.

The impact of economic growth on nitrogen concentration is negative in Western Cape province around the city of Cape Town as well as in major cities in other provinces. In contrast with the other water pollutant concentrations, the association with nitrogen is at its highest in the middle of the country, mostly in the provinces of Northern Cape and Free State.

The dissimilarities in the spatial relationships between the different water pollutant indices and these two economic variables might be explained by the divergence in land use and in water pollution policies across provinces. For instance, provinces such as Free State and Gauteng have a large mining industry (Figure 6-5 in the Appendix) such that the economic development of these regions may worsen the water quality, on average. Such an impact would mainly cause the TDS and phosphorus concentrations to be high in the eastern part of the country.

Acknowledging these divergent impacts of economic development on the water pollutant indices, we can still draw some analogous results across each map. First, it appears that the fine-scale data highlight a difference in the relationship between economic variables and water pollution within and across provincial boundaries (represented in gray on the maps). This relationship varies widely within the provinces of Free State, Northern Cape, Limpopo, Eastern Cape, and Mpumalanga. Therefore, it is essential to go beyond provincial policies and identify significant stakeholders causing these differences for policy development purposes.

Second, allowing the relationship between the level of development and water quality to be nonstationary highlights a difference between the western and the eastern parts of South Africa. Generally, the level of economic development appears to be a factor that decreases the water quality in the provinces of North West, Free State, Gauteng, Mpumalanga and Limpopo, while it is positively related to higher water quality in Western Cape and most of Eastern Cape and North Cape provinces. This finding may follow the observations of previous works claiming that a large number of dams in eastern South Africa were significantly vulnerable to eutrophication (Matthews & Bernard, 2015). Moreover, the impact of economic growth on pollutant concentration is negative within major cities but positive around these cities, particularly cities in the east (Johannesburg, Pretoria, and Durban). This relationship could be evidence of the enforcement of effective water policies in large cities but not in their poorer suburbs. Therefore, we tend to reject null hypothesis (1) that economic development is



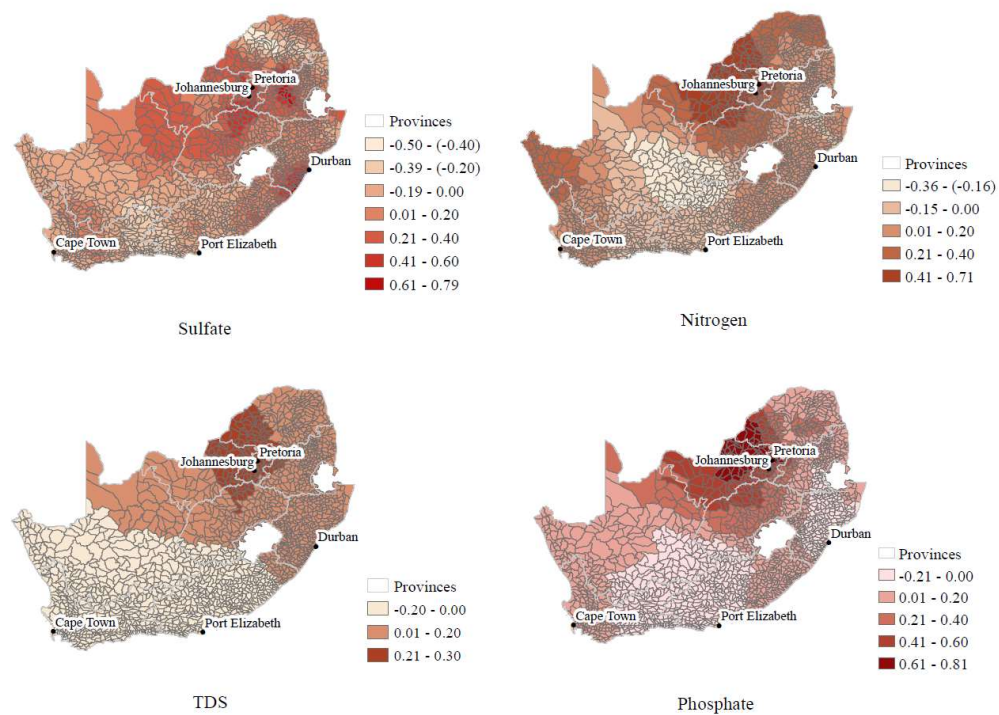
generally a factor of water quality improvement because in some regions, it appears to be a cause of water quality degradation.

According to the literature, we postulate two explanations (that will be tested in the following section of the analysis) for the positive relationship between the level of development and water pollution in the eastern part of the country. First, it could be the result of the large mining, industrial, and agricultural activities in the East (see land use concentration in the East in Figure 6-5 and (McCarthy, 2011; Wood, 1990)). Indeed, previous localized studies show that this human-driven land use in South Africa might be the cause of severe surface water pollution. For instance, the positive relationship between economic development and phosphorus concentration in the regions of Limpopo could be driven by the large phosphorus mining that is located there (Roux et al., 1989).

Second, there might be a link between water pollution and background as a former homeland or black-dominated area during apartheid, as these areas were mainly concentrated in the eastern part of the country (Makaudze & Gelles, 2015; Schreiner & Van Koppen, 2003; Smith & Hanson, 2003; Wood, 1990). In particular, when assessing the relationship between economic growth and water quality, negative relationships appear within major cities (which were the largest white-dominated areas) and not around them. This provides evidence against the null hypothesis (2) that the impact of economic development is not related to former apartheid segregation.

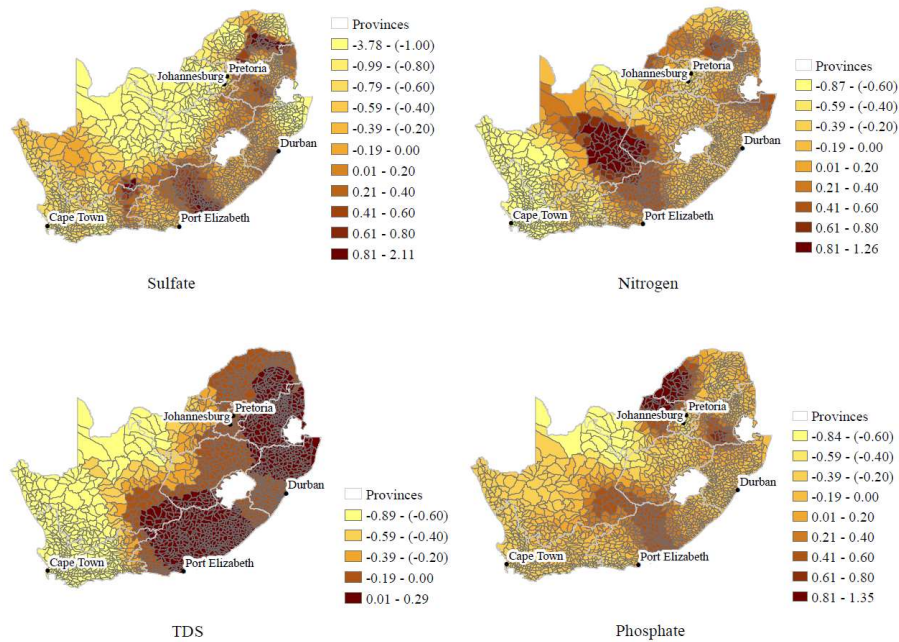
Following these interpretations, I test the significance of the GWR results while focusing on the eastern regions and the former white-dominated areas to reveal the cause of the positive relationship between water pollution and economic development.

Figure 6-1 The impact of the level of development on water quality indices



**Note:** The provincial data boundaries are collected from <https://gadm.org/about.html>

Figure 6-2 The impact of economic growth on water quality indices



**Note:** The provincial data boundaries are collected from <https://gadm.org/about.html>

### *6.5.2 The impact of economic development and its statistical significance*

The investigation of the statistical significance of the relationship between economic development and water pollutant indices is presented in Table 6-1. Regarding, the direct relationship, the estimation results confirm that there is a positive impact of the average economic level on water pollution at the national level. This may suggest that pollution in the eastern part of the country impacts the national average to such a degree that it obscures the negative relationship found in the west. Interestingly, there is a negative spillover effect or indirect impact of the economic level on water pollution. Water management areas close to economically developed areas tend to witness a low pollution in water.

This uncommon relationship, implying that economic development is not related to a reduction of the level of pollution where there is high economic activity but is significantly related to a reduction of water pollution in neighboring areas, has already been observed in some studies estimating the relationship between economic development and CO<sub>2</sub> emissions in China (Meng & Huang, 2018) and the relationship between economic development and the ecological footprint internationally (Y. Wang, Kang, Wu, & Xiao, 2013).

Here, I believe this relationship is a consequence of the dual effect of increased pollution and improved policy owing to economic development, noticeable only within the specific water management area. Water quality regulations serving to combat the augmentation of pollution (influenced by economic level) do not stop at the water management boundaries

but also spread to neighboring areas. Indeed, the water management areas where policy is actively implemented do not correspond to the administrative boundaries (which include several water management areas) where policy is planned (Beck et al., 2016; Feris & Kotzé, 2014). Therefore, the average level of development in a water area, by increasing the economic level in different administrative spheres, may allow several water management areas to benefit from better water policy. In that configuration, although the water quality of suburbs near large cities is low, these suburbs may benefit from the water quality policies associated with the neighboring economic level. The sum of the direct and indirect impacts of the average economic level on water pollutants is negative. Hence, despite some pollution that might occur due to human activities, average economic level is, on average, a determinant of better water quality.

However, economic growth does not have a significant direct impact on water pollutants but has a positive and significant effect on neighboring areas of water pollutants. This implies that the greater the economic growth in a water management area was, the higher the water pollutant concentrations in neighboring water management areas. This result, which contrasts that of the average level of development, may manifest informal settlements close to areas with high economic growth. These informal settlements appear due to the creation of new jobs in neighboring cities, but they do not generally have proper water policies because of their informal nature (Smith & Hanson, 2003). Hence, they are a manifestation of the “dual economy”

in South Africa since apartheid (Acemoglu & Robinson, 2012), where economically dynamic cities hire poorer laborers from neighboring areas to generate economic growth while maintaining certain segregation with these neighboring areas in terms of infrastructure, wages, and the environment. In sum, the total impact of economic growth is positive to water pollutant concentration, although this relationship is significant only for sulfate and phosphate concentrations. This finding highlights the potential unsustainability of economic growth in South Africa.

Subsequently, I include the land use variables to control the effect of the nature of activities per water management area (Table 6-2). Controlling for land use, however, considers only the surface area and the nature of the economic activity in each area. These variables, therefore, neglect the intensity and level of development of the given land use, although this information is included in the economic variables.

The direct impact of night-time light intensity is no longer significant. The drop in significance in these results highlights that the pollution due to the economic level is driven by the land use variables. Thus, these results strengthen the assumption that the overall water pollution caused by the economic level is shaped by the nature of activities. Regarding the impact of the average economic level in neighboring areas, the coefficients are similar to those of the model without any land uses variables and are significant at the 1% level. This bolsters

the previous interpretation linking the average economic level to water policies beyond the boundaries of water management areas.

Moving to economic growth, the inclusion of land use variables appears to have reduced the significance of its relationship with neighboring areas water pollutants. Thus, the nature of economic growth only partially influences the positive effect of growth on neighboring areas water pollution. This result can still owe to the South African dual economy.



Table 6-1 The average impact of economic development on water quality

|                                                                       |                                       | <b>Log TDS</b> | <b>Log SO4</b> | <b>Log PO4</b> | <b>Log NO3-NO2-N</b> |
|-----------------------------------------------------------------------|---------------------------------------|----------------|----------------|----------------|----------------------|
| <b>Direct<br/>impact</b>                                              | <b>Log (light) 2010</b>               | 0.032          | 0.076**        | 0.105***       | 0.105***             |
|                                                                       | <b>Log (light growth) 2005 - 2010</b> | -0.111         | -0.250         | -0.019         | -0.119               |
| <b>Indirect<br/>impact</b>                                            | <b>Log (light) 2010</b>               | -0.502***      | -0.652***      | -0.508***      | -0.368***            |
|                                                                       | <b>Log (light growth) 2005 - 2010</b> | 1.942*         | 3.532***       | 1.471***       | 0.958                |
| <b>Total<br/>impact</b>                                               | <b>Log (light) 2010</b>               | -0.470***      | -0.576***      | -0.402***      | -0.262**             |
|                                                                       | <b>Log (light growth) 2005 - 2010</b> | 1.830          | 3.282***       | 1.452**        | 0.838                |
| Robust standard errors in parentheses, *** p<0.01, ** p<0.05, * p<0.1 |                                       |                |                |                |                      |

Table 6-2 The average impact of economic development on water quality after controlling for the impact of land use

|                                                                       |                                       | <b>Log TDS</b> | <b>Log SO4</b> | <b>Log PO4</b> | <b>Log NO3-NO2-N</b> |
|-----------------------------------------------------------------------|---------------------------------------|----------------|----------------|----------------|----------------------|
| <b>Direct impact</b>                                                  | <b>Log (light) 2010</b>               | -0.017         | 0.022          | -0.002         | 0.0333               |
|                                                                       | <b>Log (light growth) 2005 - 2010</b> | -0.04          | -0.217         | 0.127          | 0.001                |
| <b>Indirect impact</b>                                                | <b>Log (light) 2010</b>               | -0.543***      | -0.793***      | -0.586***      | -0.380***            |
|                                                                       | <b>Log (light growth) 2005 - 2010</b> | 2.012**        | 2.591**        | 0.091          | 0.293                |
| <b>Total impact</b>                                                   | <b>Log (light) 2010</b>               | -0.560***      | -0.771***      | -0.589***      | -0.347***            |
|                                                                       | <b>Log (light growth) 2005 - 2010</b> | 1.964*         | 2.374**        | 0.217          | 0.295                |
| Robust standard errors in parentheses, *** p<0.01, ** p<0.05, * p<0.1 |                                       |                |                |                |                      |

### *6.5.3 Factors underlying the regional differences in the impact of economic development on water pollution across regions*

To explore the robustness of the pattern and the interpretations of the GWR maps, I run a regression that includes the average economic level in 2010, economic growth from 2005 to 2010, a dummy for the eastern region and the interaction between this dummy and the two former variables in model (6.4)<sup>n</sup> (Table 6-3).

The results show that the eastern dummy variables are negative and significant, implying that eastern water management areas have, on average, fewer pollutants in their surface water than the western regions (see Table 6-10 in the Appendix). However, the interactions of the eastern dummy with the average level of development have a positive and significant direct impact on three out of the four water pollutants. Therefore, economic development has a higher propensity to pollute water in the east than in the west (Table 6-3). This confirms the GWR results and implies that there is pollution related to the level of development in the eastern part of the country.

When including the land uses in the model to check whether the nature of activities could be the cause of higher pollution in the east, it yields essentially identical result (Table 6-11 in the Appendix). It suggests that land use is not the main driver of the difference in the

---

<sup>n</sup> We still include the environmental variables (river code, runoff, and water management area), which are necessary for an unbiased estimation of the spatial lag of the dependent variables in the model.

impact of the average economic level on water quality between the east and the west. This potentially implies that the western part of the country has some better water policies.

Second, I check for the hypothesis that the difference in the impact of economic development on water quality between the eastern and western regions could be the consequence of apartheid's inequitable water policies. I use the racial distribution map in 1970 in South Africa from the University of Texas<sup>o</sup> to create a dummy variable of former white-dominated lands. With this dummy, I regress model (6.4) with night-time light, night-time light growth, and their respective interactions with the dummy.

The results show that the dummy variable of former white-dominated areas is positive for all the regression and significant for phosphate and sulfate (see Appendix Table 6-12). Hence, water areas that were white dominated have, unexpectedly, more water pollutants than other areas of the country, all else being equal. Subsequently, regarding the interaction terms in former white-dominated areas, the economic development is positive but not significant, while economic growth is negative and significantly different from growth in other areas (Table 6-4). Hence, based on the assumption of economic development having a dual impact, economic growth (through efficient water policies) in former white-dominated areas decreases water pollutant concentrations more than in other parts of the country. This result is related to two interpretations.

---

<sup>o</sup> [https://legacy.lib.utexas.edu/maps/south\\_africa.html](https://legacy.lib.utexas.edu/maps/south_africa.html) on 06/17/2019 - Map No. 503971 1979 originally produced by the U.S. Central Intelligence Agency.

First, using insights of EKC theory, this result may show that former white-dominated areas are in a stage of development where economic growth improves water quality. This is the result of an average larger economic development than other regions of the country and the enforcement of water policies for a longer period. Second, these results relate to poor water quality in the informal settlements in former white-dominated areas as being a former white-dominated area is linked to lower water quality, while economic growth in there improves surface water quality. Therefore, townships (in former white-dominated areas) with economic growth lower than in downtown areas, have poorer water quality. This result rejects the null hypothesis (2) and suggests that economic growth improves water quality more in former white-dominated areas than in other parts of the country.

On the other hand, although this relationship is significant only for the phosphate concentration, the indirect impact of the economic level in former white-dominated areas is related to a higher concentration of water pollutants. Whereas the indirect impact in other areas of the country is related to lower concentration of water pollutants. Similarly, the indirect impact of economic growth on the water pollution in former white-dominated areas is related to higher pollutant concentration than that of other parts of the country, with the result for TDS and sulfate being significant. Although, the results are not consistent for all the water pollutant indices, this finding may suggest that economic development in former white-dominated areas does not lead to better water policies in neighboring areas.

These relationships suggest two intuitions. First, the dual economy in South Africa persists several years after apartheid (Acemoglu & Robinson, 2012). Hence, the result describes the presence of economically developed and dynamic former white-dominated areas with good water policies surrounded by other water management areas with significantly less effective water policies. Second, water-related investments made in municipalities (at a local and district level) and provinces benefit former white-dominated areas more than they benefit other areas of the same province or municipality. Thus, these results represent significant evidence of the government's inability to effectively provide good water policy outside the former white-dominated areas at the lowest spatial level of water policy enforcement.

Table 6-3 The average impact of economic development on water quality in eastern regions and western regions without control variables

|                                                                      |                                       | Log TDS   | Log SO4   | Log PO4    | Log NO3-NO2-N |
|----------------------------------------------------------------------|---------------------------------------|-----------|-----------|------------|---------------|
| Direct impact                                                        | Log (light) 2010                      | -0.037    | -0.009    | -0.007     | -0.0001       |
|                                                                      | East * Log (light) 2010               | 0.515***  | 0.711***  | 0.4256*    | 0.666**       |
|                                                                      | Log (light growth) 2005 - 2010        | -0.233    | -0.220    | -0.002     | -0.102        |
|                                                                      | East * Log (light growth) 2005 - 2010 | 0.332**   | 0.069     | 0.1975     | 0.237         |
| Indirect impact                                                      | Log (light) 2010                      | -0.450*** | -0.608**  | -0.867***  | -0.223        |
|                                                                      | East * Log (light) 2010               | -0.059    | 0.591     | 1.709***   | 0.387         |
|                                                                      | Log (light growth) 2005 - 2010        | 2.094     | 3.477*    | 2.697***   | 1.176         |
|                                                                      | East * Log (light growth) 2005 - 2010 | -0.907    | -2.871    | -4.899***  | -2.784        |
| Total impact                                                         | Log (light) 2010                      | -0.488*** | -0.617*** | -0.8741*** | -0.223        |
|                                                                      | East * Log (light) 2010               | 0.455     | 3.256*    | 2.134***   | 1.074         |
|                                                                      | Log (light growth) 2005 - 2010        | 1.862     | 1.302*    | 2.695***   | 1.053         |
|                                                                      | East * Log (light growth) 2005 - 2010 | -0.575    | -2.801    | -4.701***  | -2.547        |
| Robust standard errors in parentheses *** p<0.01, ** p<0.05, * p<0.1 |                                       |           |           |            |               |

Table 6-4 The average impact of economic development on water quality in former black-dominated lands with the control variables

|                                                                      |                                        | Log TDS    | Log SO4   | Log PO4   | Log NO3-NO2-N |
|----------------------------------------------------------------------|----------------------------------------|------------|-----------|-----------|---------------|
| Direct impact                                                        | Log (light) 2010                       | -0.004     | 0.0427    | 0.012     | 0.0433        |
|                                                                      | White * Log (light) 2010               | 0.011      | 0.0525    | 0.223**   | 0.2509        |
|                                                                      | Log (light growth) 2005 - 2010         | -0.0158    | -0.203    | 0.107     | 0.023         |
|                                                                      | White * Log (light growth) 2005 - 2010 | -1.478***  | -1.084*** | -0.628    | -1.688***     |
| Indirect impact                                                      | Log (light) 2010                       | -0.5091*** | -0.665    | -0.793*** | -0.534***     |
|                                                                      | White * Log (light) 2010               | 0.7329     | 1.696     | 2.78***   | 1.51          |
|                                                                      | Log (light growth) 2005 - 2010         | 1.7320*    | 1.983     | -0.161    | 0.132         |
|                                                                      | White * Log (light growth) 2005 - 2010 | 9.580*     | 21.03**   | 7.554     | 4.299         |
| Total impact                                                         | Log (light) 2010                       | -0.514***  | -0.622    | -0.781*** | -0.490***     |
|                                                                      | White * Log (light) 2010               | 0.744      | 1.748     | 3.008***  | 1.760         |
|                                                                      | Log (light growth) 2005 - 2010         | 1.716*     | 1.7804    | -0.054    | 0.155         |
|                                                                      | White * Log (light growth) 2005 - 2010 | 8.101      | 19.948**  | 6.926     | 2.611         |
| Robust standard errors in parentheses *** p<0.01, ** p<0.05, * p<0.1 |                                        |            |           |           |               |



## **6.6. Conclusions and policy implications**

In sum, initial estimations of the GWR suggest that to analyze data across regions, it is necessary to first recognize the varying relationships between water quality and economic development in each region. Such variation implies a double impact of economic development on water quality. Using fine-scale data of water management areas reveals that economic development through urban settlements, mining, industrial and agricultural activities poses a risk for water quality. On the other hand, economic development can strengthen water policies, thereby significantly reducing water pollution. However, after controlling for the effect of land use and the effect of water pollution from neighboring areas, we found a significant positive relationship between water quality and economic development in Western Cape and Eastern Cape and a negative relationship in the provinces of Gauteng, Free State, and North West. In addition, I found that economic growth in former white-dominated areas (the most developed cities in South Africa) locally improves water quality, but this effect does not benefit neighbouring areas.

Therefore, acknowledging that all economic activities require decent surface water quality (DWAF, 1996; Verheyen et al., 2015; Yokwe, 2009) and assuming that such activities are not intentionally developed in areas with poor water quality, I suggest that variation in the relationship between economic activity and water quality can be found in the implementation and the efficacy of water policy in South Africa. In particular, the study shows that areas that

had the best water policies during apartheid seem to also have the best water policies currently. Although the results do not imply current racial discrimination, they show the inefficiency of the government in effectively rectifying spatial inequalities in water quality, which is one of their primary objectives since 1994. Furthermore, the results of the analysis imply that there is a need to tackle the issue of mining, industrial and agricultural pollution in the country, as their pollution pose important risks to water quality.

Nationally, despite the apparent equitable spending of the government on water-related issues, some water management areas do not seem to have enough expertise to efficiently apply policies (Folifac, 2007; Herselman et al., 2006). Therefore, I recommend a homogenization of water policies (between the Western and the Eastern provinces) so that water institutions and managers with superior water management records could share their expertise. Also, I recommend an adjustment of the size of water management areas according to their hydrological boundaries (basins) for effective management in cooperation with meaningful stakeholders (related to mining, industrial and urban areas) who have direct and/or spillover impacts on the surface water quality (Herrfahrdt-Pähle, 2010; Shackleton, Le Maitre, van Wilgen, & Richardson, 2016). Mining, industrial, and agricultural activities, whose pollutants spread over a wide area, must be regulated by effective policies that consider such large areas or by policies that operate in synergy across different water management areas and provinces.

Because of the inability of the Government to address the issues related to water pollution, South Africa now faces a lack of incentives in black communities to partake in the implemented water tariffs; certain policies seem disconnected to population needs (Nleya, 2008); and above all, the number of informal settlements and the legitimacy of their presences seems increasing (Makaudze & Gelles, 2015; Nleya, 2008; Smith & Hanson, 2003). Hence, the results of the analysis include themselves in a vicious circle: the high water pollution caused by the apartheid, in certain areas, reduces the confidence in the Government thereby hindering the ability of the Government to propose policies to residents' needs. Therefore, the Government in South Africa would have to go beyond cost-efficient and egalitarian policies (Makaudze & Gelles, 2015; Nleya, 2008; Smith & Hanson, 2003) and propose differentiated policies between the former white-dominated areas and the other areas. The statue-quo (in which economic growth in former white-dominated areas significantly reduces surface water pollution but not in the other areas) is likely to worsen the solidarity of the country.

Finally, this study suggests the need for further research on the nature of water policies most likely to improve water quality, as well as the identification of actions that can improve their efficiency in South Africa. Furthermore, the results show that there is a need in addressing water policy efficiencies in countries, like South Africa, that suffered from historical inequalities. Indeed, internationally improving water sanitation and services with equity is critical in a nexus approach of the Sustainable Development Goals (SDGs) (J. Liu et al., 2018).

Thus, countries in which economic development worsens water quality are very likely to diverge from a sustainable development.

## References

- Abildtrup, Jens, Serge Garcia, and Anne Stenger. 2013. "The Effect of Forest Land Use on the Cost of Drinking Water Supply: A Spatial Econometric Analysis." *Ecological Economics* 92: 126–36. <https://doi.org/10.1016/j.ecolecon.2013.01.004>.
- Acemoglu, Daron, and James Robinson. 2012. "Chapter 9: Reversing Development." In *Why Nations Fail : The Origins of Power, Prosperity and Poverty*. London, England: Crown Publisher.
- Adewumi, J. R., A. A. Ilemobade, and J. E. Van Zyl. 2010. "Treated Wastewater Reuse in South Africa: Overview, Potential and Challenges." *Resources, Conservation and Recycling*. <https://doi.org/10.1016/j.resconrec.2010.09.012>.
- Adler, Rebecca a, Marius Claassen, Linda Keren Godfrey, and Anthony R Turton. 2007. "Water, Mining, and Waste: A Historical and Economic Perspective on Conflict Management in South Africa." *Economics of Peace and Security Journal* 2 (2): 33–41. <https://doi.org/10.15355/epsj.2.2.33>.
- Ahjum, Fadiel, Bruno Merven, James Cullis, Gary Goldstein, Pat Delaquil, and Adrian Stone. 2018. "Development of a National Water-Energy System Model with Emphasis on the Power Sector for South Africa." *Environmental Progress and Sustainable Energy* 37 (1): 132–47. <https://doi.org/10.1002/ep.12837>.
- Akcil, Ata, and Soner Koldas. 2006. "Acid Mine Drainage (AMD): Causes, Treatment and Case Studies." *Journal of Cleaner Production* 14 (12-13 SPEC. ISS.): 1139–45. <https://doi.org/10.1016/j.jclepro.2004.09.006>.
- Amin, Ariane. 2016. "Exploring the Role of Economic Incentives and Spillover Effects in Biodiversity Conservation Policies in Sub-Saharan Africa." *Ecological Economics* 127: 185–91. <https://doi.org/10.1016/j.ecolecon.2016.03.018>.
- Anderson, Barbara A., John H. Romani, Heston Phillips, Marie Wentzel, and Kholadi Tlabela. 2007. "Exploring Environmental Perceptions, Behaviors and Awareness: Water and Water Pollution in South Africa." *Population and Environment* 28 (1): 133–61. <https://doi.org/10.1007/s11111-007-0038-5>.
- Ashton, Peter, David Love, Harriet Mahachi, and Paul Dirks. 2001. "An Overview of the Impact of Mining and Mineral Processing Operations on Water Resources and Water Quality in the Zambezi, Limpopo and Olifants Catchments in Southern Africa." *Contract Report to the Mining, Minerals and Sustainable Development (Southern Africa) Project, by CSIR-Environmentek, Pretoria and Geology Department, University of Zimbabwe-Harare. Report No. ENV-PC, 42*. <https://doi.org/ENV-P-C 2001-042>.

- Beck, Thessa, L. Rodina, E. Lucker, and L. Harris. 2016. "Institutional and Policy Mapping of the Water Sector in South Africa." Vancouver, Canada.  
<https://doi.org/10.13140/RG.2.2.32761.88164>.
- Best, Jim. 2019. "Anthropogenic Stresses on the World's Big Rivers." *Nature Geoscience* 12 (1): 7–21. <https://doi.org/10.1038/s41561-018-0262-x>.
- Brunsdon, Chris, A. Stewart Fotheringham, and Martin E. Charlton. 2010. "Geographically Weighted Regression: A Method for Exploring Spatial Nonstationarity." *Geographical Analysis* 28 (4): 281–98. <https://doi.org/10.1111/j.1538-4632.1996.tb00936.x>.
- Chakir, Raja, and Julie Le Gallo. 2013. "Predicting Land Use Allocation in France: A Spatial Panel Data Analysis." *Ecological Economics* 92: 114–25.  
<https://doi.org/10.1016/j.ecolecon.2012.04.009>.
- Chang, Heejun. 2008. "Spatial Analysis of Water Quality Trends in the Han River Basin, South Korea." *Water Research* 42 (13): 3285–3304.  
<https://doi.org/10.1016/j.watres.2008.04.006>.
- Chien, L, H Robertson, and J W Gerrard. 1968. "Infantile Gastroenteritis Due to Water with High Sulfate Content." *Canadian Medical Association Journal* 99 (3): 102–4.
- Constitution of the Republic of South Africa. 1996. *Constitution of the Republic of South Africa No. 108 of 1996*. Cape Town. <https://doi.org/10.1017/S0021855300011499>.
- Dikio, E. D. 2010. "Water Quality Evaluation of Vaal River, Sharpeville and Bedworth Lakes in the Vaal Region of South Africa." *Research Journal of Applied Sciences, Engineering and Technology* 2 (6): 574–79.
- Duflo, Esther, and Rohini Pande. 2007. "Dams." *Quarterly Journal of Economics*.  
<https://doi.org/10.1162/qjec.122.2.601>.
- DWA. 2008. "Water for Growth and Development in South Africa: Version 6." Cape Town, South Africa.
- DWAF. 1996. "Water Quality Guidelines Agricultural Use: Irrigation." *Water*. Vol. 4. Pretoria, South Africa.
- DWS. 2013. "Green Drop Progress Report." South Africa, Pretoria.
- . 2018. "Strategic Overview of the Water Sector in South Africa 2018." Pretoria, South Africa.
- Feris, Loretta, and Louis J. Kotzé. 2014. "The Regulation of Acid Mine Drainage in South Africa: Law and Governance Perspectives." *Potchefstroom Electronic Law Journal*.

<https://doi.org/10.17159/1727-3781/2014/v17i5a2158>.

- Folifac, Fidelis A. 2007. "National Water Policies and Water Services at the Extremes : What Challenges Must Be Faced in Bridging the Gap ? Learning from the South Africa Experience." *African Water Journal*.
- Fotheringham, Alexander, Chris Brunsdon, and Martin Charlton. 2003. *Geographically Weighted Regression : The Analysis of Spatially Varying Relationships Data*. John Wiley & Sons. John Wiley & Sons. <https://doi.org/10.1080/13658816.2013.865739>.
- Funke, Nikki, Karen Nortje, Kieran Findlater, Mike Burns, Anthony Turton, Alex Weaver, and Hanlie Hattingh. 2007. "Redressing Inequality South Africa's New Water Policy." *Environment*. <https://doi.org/10.3200/ENVT.49.3.10-25>.
- Gibbons, Stephen, and Henry G. Overman. 2012. "Mostly Pointless Spatial Econometrics?" *Journal of Regional Science*. <https://doi.org/10.1111/j.1467-9787.2012.00760.x>.
- Giri, Subhasis, and Zeyuan Qiu. 2016. "Understanding the Relationship of Land Uses and Water Quality in Twenty-First Century: A Review." *Journal of Environmental Management* 173 (August): 41–48. <https://doi.org/10.1016/j.jenvman.2016.02.029>.
- Goldblatt, Michael. 1999. "Assessing the Effective Demand for Improved Water Supplies in Informal Settlements: A Willingness to Pay Survey in Vlakfontein and Finetown, Johannesburg." *Geoforum* 30 (1): 27–41. [https://doi.org/10.1016/S0016-7185\(98\)00034-7](https://doi.org/10.1016/S0016-7185(98)00034-7).
- Henderson, J Vernon, Adam Storeygard, and David N Weil. 2012. "Measuring Economic Growth from Outer Space." *American Economic Review* 102 (2): 994–1028. <https://doi.org/10.1257/aer.102.2.994>.
- Herrfahrdt-Pähle, Elke. 2010. "South African Water Governance between Administrative and Hydrological Boundaries." *Climate and Development* 2 (2): 111–27. <https://doi.org/10.3763/cdev.2010.0038>.
- Herselman, J. E., Carl E. Steyn, and H. G. Snyman. 2006. "Dedicated Land Disposal of Wastewater Sludge in South Africa: Leaching of Trace Elements and Nutrients." *Water Science and Technology* 54 (5): 139–46. <https://doi.org/10.2166/wst.2006.556>.
- Hess, T. M., J. Sumberg, T. Biggs, M. Georgescu, D. Haro-Monteagudo, G. Jewitt, M. Ozdogan, et al. 2016. "A Sweet Deal? Sugarcane, Water and Agricultural Transformation in Sub-Saharan Africa." *Global Environmental Change* 39: 181–94. <https://doi.org/10.1016/j.gloenvcha.2016.05.003>.
- Hopolang, Selebalo, and Webster Dennis. 2018. "Monitoring and Evaluating the Progressive

- Realisation of the Right to Water and Sanitation in South Africa.” 18. Johannesburg, South Africa.
- Kang, Joo Hyon, Seung Won Lee, Kyung Hwa Cho, Seo Jin Ki, Sung Min Cha, and Joon Ha Kim. 2010. “Linking Land-Use Type and Stream Water Quality Using Spatial Data of Fecal Indicator Bacteria and Heavy Metals in the Yeongsan River Basin.” *Water Research* 44 (14): 4143–57. <https://doi.org/10.1016/j.watres.2010.05.009>.
- LeSage, James P. 2013. “A Family of Geographically Weighted Regression Models.” In *Advances in Spatial Econometrics*, 241–64. Berlin, Heidelberg: Springer. [https://doi.org/10.1007/978-3-662-05617-2\\_11](https://doi.org/10.1007/978-3-662-05617-2_11).
- Li, Qian, Jinping Song, Enru Wang, Hao Hu, Jianhui Zhang, and Yeyao Wang. 2014. “Economic Growth and Pollutant Emissions in China: A Spatial Econometric Analysis.” *Stochastic Environmental Research and Risk Assessment* 28 (2): 429–42. <https://doi.org/10.1007/s00477-013-0762-6>.
- Lintern, A., J.A. Webb, D. Ryu, S. Liu, U. Bende-Michl, D. Waters, P. Leahy, P. Wilson, and A. W. Western. 2018. “Key Factors Influencing Differences in Stream Water Quality across Space.” *Wiley Interdisciplinary Reviews: Water* 5 (1): e1260. <https://doi.org/10.1002/wat2.1260>.
- Liu, Jianguo, Vanessa Hull, H Charles J Godfray, David Tilman, Peter Gleick, Holger Hoff, Claudia Pahl-Wostl, et al. 2018. “Nexus Approaches to Global Sustainable Development.” *Nature Sustainability* 1 (9): 466. <https://doi.org/10.1038/s41893-018-0135-8>.
- Mainali, Janardan, Heejun Chang, and Yongwan Chun. 2019. “A Review of Spatial Statistical Approaches to Modeling Water Quality.” *Progress in Physical Geography* 43 (6): 801–26. <https://doi.org/10.1177/0309133319852003>.
- Makaudze, Ephias M., and Gregory M. Gelles. 2015. “The Challenges of Providing Water and Sanitation to Urban Slum Settlements in South Africa.” In *Understanding and Managing Urban Water in Transition*, Springer, 121–33. Dordrecht, Netherlands.
- Matthews, Mark W., and Stewart Bernard. 2015. “Eutrophication and Cyanobacteria in South Africa’s Standing Water Bodies: A View from Space.” *South African Journal of Science*. <https://doi.org/10.17159/sajs.2015/20140193>.
- McCarthy, Terence S. 2011. “The Impact of Acid Mine Drainage in South Africa.” *South African Journal of Science* 107 (5/6): 1–7. <https://doi.org/10.4102/sajs.v107i5/6.712>.
- Meng, Lina, and Bo Huang. 2018. “Shaping the Relationship Between Economic Development and Carbon Dioxide Emissions at the Local Level: Evidence from Spatial



- Econometric Models.” *Environmental and Resource Economics* 71 (1): 127–56.  
<https://doi.org/10.1007/s10640-017-0139-2>.
- Mettetal, Elizabeth. 2019. “Irrigation Dams, Water and Infant Mortality: Evidence from South Africa.” *Journal of Development Economics*.  
<https://doi.org/10.1016/j.jdeveco.2018.11.002>.
- Nleya, Ndodana. 2008. “Development Policy and Water Services in South Africa: An Urban Poverty Perspective.” *Development Southern Africa*.  
<https://doi.org/10.1080/03768350802212048>.
- Peterson, Erin E., Fran Sheldon, Ross Darnell, Stuart E. Bunn, and Bronwyn D. Harch. 2011. “A Comparison of Spatially Explicit Landscape Representation Methods and Their Relationship to Stream Condition.” *Freshwater Biology* 56 (3): 590–610.  
<https://doi.org/10.1111/j.1365-2427.2010.02507.x>.
- Roux, E. H., D. H. de Jager, J. H. du Plooy, A. Nicotra, G. J. van der Linde, and P. de Waal. 1989. “Phosphate in South Africa.” *Journal of The South African Institute of Mining and Metallurgy* 89 (5): 129–39.
- Schreiner, Barbara, and Barbara Van Koppen. 2003. “Policy and Law for Addressing Poverty, Race and Gender in the Water Sector: The Case of South Africa.” *Water Policy*.
- Shackleton, Ross T., David C. Le Maitre, Brian W. van Wilgen, and David M. Richardson. 2016. “Identifying Barriers to Effective Management of Widespread Invasive Alien Trees: Prosopis Species (Mesquite) in South Africa as a Case Study.” *Global Environmental Change* 38: 183–94. <https://doi.org/10.1016/j.gloenvcha.2016.03.012>.
- Smith, Laïla, and Susan Hanson. 2003. “Access to Water for the Urban Poor in Cape Town: Where Equity Meets Cost Recovery.” *Urban Studies*.  
<https://doi.org/10.1080/0042098032000094414>.
- Torras, Mariano, and James K. Boyce. 1998. “Income, Inequality, and Pollution: A Reassessment of the Environmental Kuznets Curve.” *Ecological Economics* 25 (2): 147–60. [https://doi.org/10.1016/S0921-8009\(97\)00177-8](https://doi.org/10.1016/S0921-8009(97)00177-8).
- Tu, Jun, and Zong Guo Xia. 2008. “Examining Spatially Varying Relationships between Land Use and Water Quality Using Geographically Weighted Regression I: Model Design and Evaluation.” *Science of the Total Environment* 407 (1): 358–78.  
<https://doi.org/10.1016/j.scitotenv.2008.09.031>.
- Turpie, J. K., C. Marais, and J. N. Blignaut. 2008. “The Working for Water Programme: Evolution of a Payments for Ecosystem Services Mechanism That Addresses Both Poverty and Ecosystem Service Delivery in South Africa.” *Ecological Economics* 65 (5):

788–98. <https://doi.org/10.1016/j.ecolecon.2007.12.024>.

Verheyen, Dries, Nele Van Gaelen, Benedicta Ronchi, Okke Batelaan, Eric Struyf, Gerard Govers, Roel Merckx, and Jan Diels. 2015. “Dissolved Phosphorus Transport from Soil to Surface Water in Catchments with Different Land Use.” *Ambio*. <https://doi.org/10.1007/s13280-014-0617-5>.

Wang, Enru, Qian Li, Hao Hu, Fuli Peng, Peng Zhang, and Jianjun Li. 2018. “Spatial Characteristics and Influencing Factors of River Pollution in China.” *Water Environment Research*, 351–63. <https://doi.org/10.1002/wer.1044>.

Wang, Yong, Lingyan Kang, Xiaoqing Wu, and Yang Xiao. 2013. “Estimating the Environmental Kuznets Curve for Ecological Footprint at the Global Level: A Spatial Econometric Approach.” *Ecological Indicators* 34: 15–21. <https://doi.org/10.1016/j.ecolind.2013.03.021>.

Wood, R. A. 1990. “The Roles of Nitrogen, Phosphorus and Potassium in the Production of Sugarcane in South Africa.” *Fertilizer Research*. <https://doi.org/10.1007/BF01048746>.

World Health Organization. 2017. *Guidelines for Drinking-Water Quality*. Fourth edi. Geneva: WHO publications. [https://doi.org/10.1016/S1462-0758\(00\)00006-6](https://doi.org/10.1016/S1462-0758(00)00006-6).

Xu, Zuxin, Jin Xu, Hailong Yin, Wei Jin, Huaizheng Li, and Zhen He. 2019. “Urban River Pollution Control in Developing Countries.” *Nature Sustainability*. <https://doi.org/10.1038/s41893-019-0249-7>.

Yokwe, Stanley. 2009. “Water Productivity in Smallholder Irrigation Schemes in South Africa.” *Agricultural Water Management* 96 (8): 1223–28. <https://doi.org/10.1016/j.agwat.2009.03.012>.

## Appendix 6

### *Appendix 6 - Section A: Supplementary tables and figures*

Table 6-5 Summary statistics

| Variable                                              | Observations | Unit                           | Mean   | Standard Deviation | Minimum | Maximum |
|-------------------------------------------------------|--------------|--------------------------------|--------|--------------------|---------|---------|
| <b>Night-time light 2010</b>                          | 305          | Digital number                 | 6.81   | 5.93               | 3.08    | 52.07   |
| <b>Proportion of night-time light of 2010 in 2005</b> | 305          | -                              | 1.94   | 2.34               | 0.03    | 34.25   |
| <b>NO<sub>3</sub>+NO<sub>2</sub>-N</b>                | 305          | Mg/l                           | 0.31   | 0.60               | 0.02    | 5.10    |
| <b>SO<sub>4</sub></b>                                 | 305          | Mg/l                           | 50.89  | 103.66             | 0.02    | 1065.12 |
| <b>TDS</b>                                            | 305          | Mg/l                           | 389.07 | 581.08             | 6.68    | 4650.65 |
| <b>PO<sub>4</sub>-P</b>                               | 305          | Mg/l                           | 0.08   | 0.20               | 0.01    | 1.70    |
| <b>Agriculture coverage</b>                           | 305          | Inverse distance-weighted area | 120.67 | 148.29             | 0.15    | 950.49  |
| <b>Urban coverage</b>                                 | 305          | Inverse distance-weighted area | 12.62  | 21.55              | 0       | 163.19  |
| <b>Mining coverage</b>                                | 305          | Inverse distance-weighted area | 2.07   | 7.69               | 0       | 85.92   |
| <b>Industrial coverage</b>                            | 305          | Inverse distance-weighted area | 0.59   | 1.99               | 0       | 19.31   |
| <b>Population density</b>                             | 305          | People/ km <sup>2</sup>        | 71.38  | 174.60             | 0.07    | 1594.51 |
| <b>Total area</b>                                     | 305          | Thousands km <sup>2</sup>      | 0.05   | 0.06               | 0.01    | 0.43    |
| <b>Runoff</b>                                         | 305          | Thousands mm/annum             | 0.09   | 0.15               | 0.00    | 1.48    |
| <b>River order</b>                                    | 305          | -                              | 2.72   | 1.33               | 1       | 7       |

Figure 6-3 Sample repartition

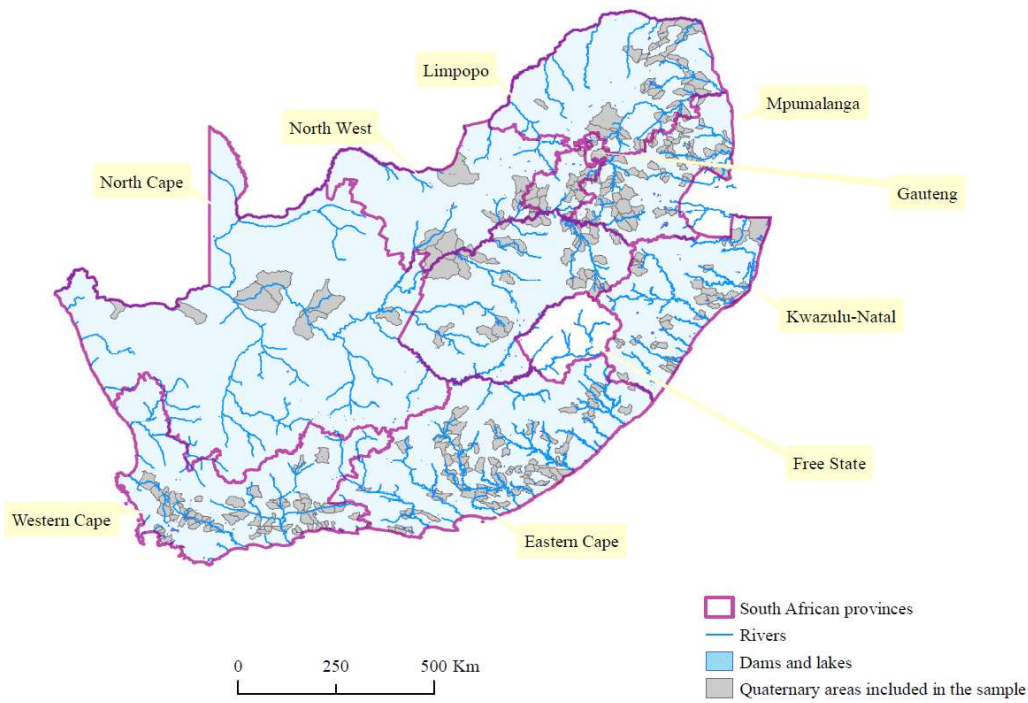


Figure 6-4Night light map in 2010

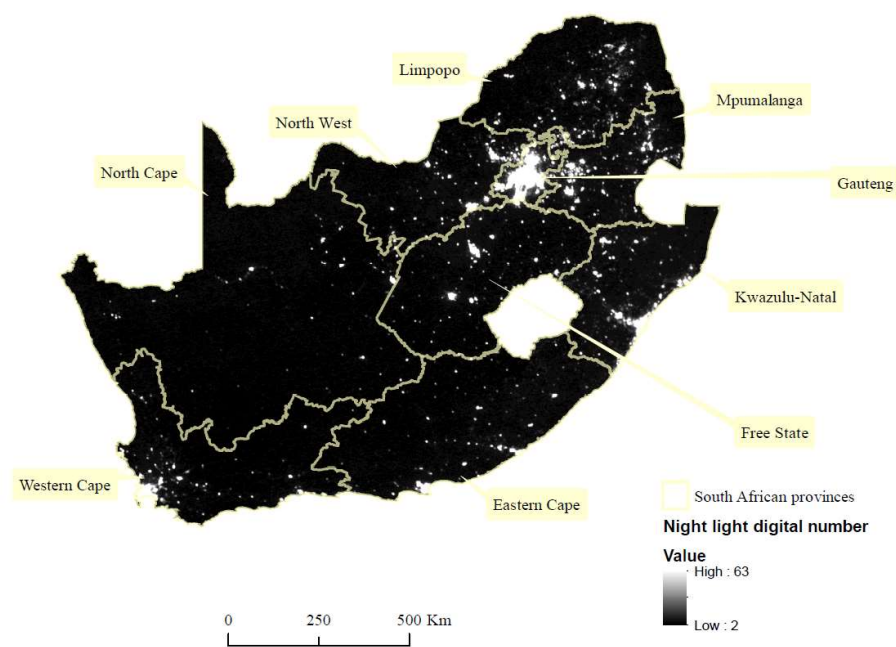
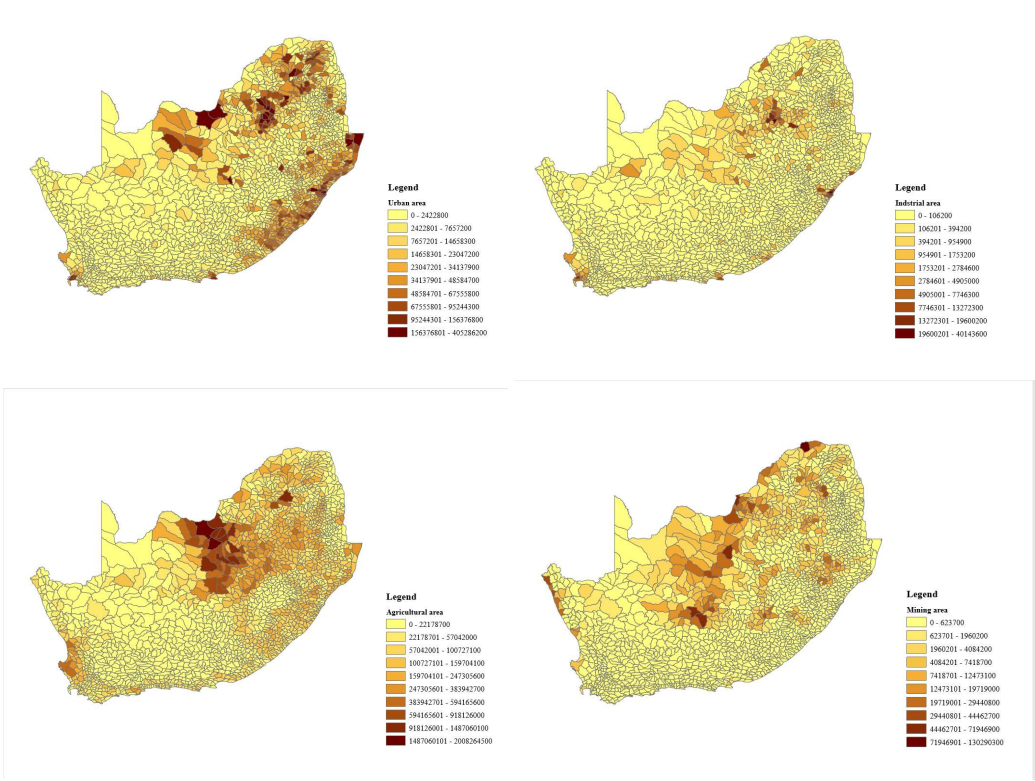


Figure 6-5 Land use repartition in South Africa



#### *Appendix 6 - Section B: Pre-estimation tests*

I tested the presence of spatial autocorrelation in the night-time light variable and the water pollutant concentration variables. The Moran's I test with the contiguity weight matrix and the Euclidean distance weight matrix was significant in all cases, rejecting the null hypothesis of nonspatial correlation in the dataset at the 1% significance level.

Both matrices are normalized at their row, which means that each weight is divided by the sum of its respective line. Regarding the contiguity matrix, I use the queen criterion, which considers an observation as a neighbour if it shares a common edge or vertex.

Figure 6-6 Clusters among dependent variables

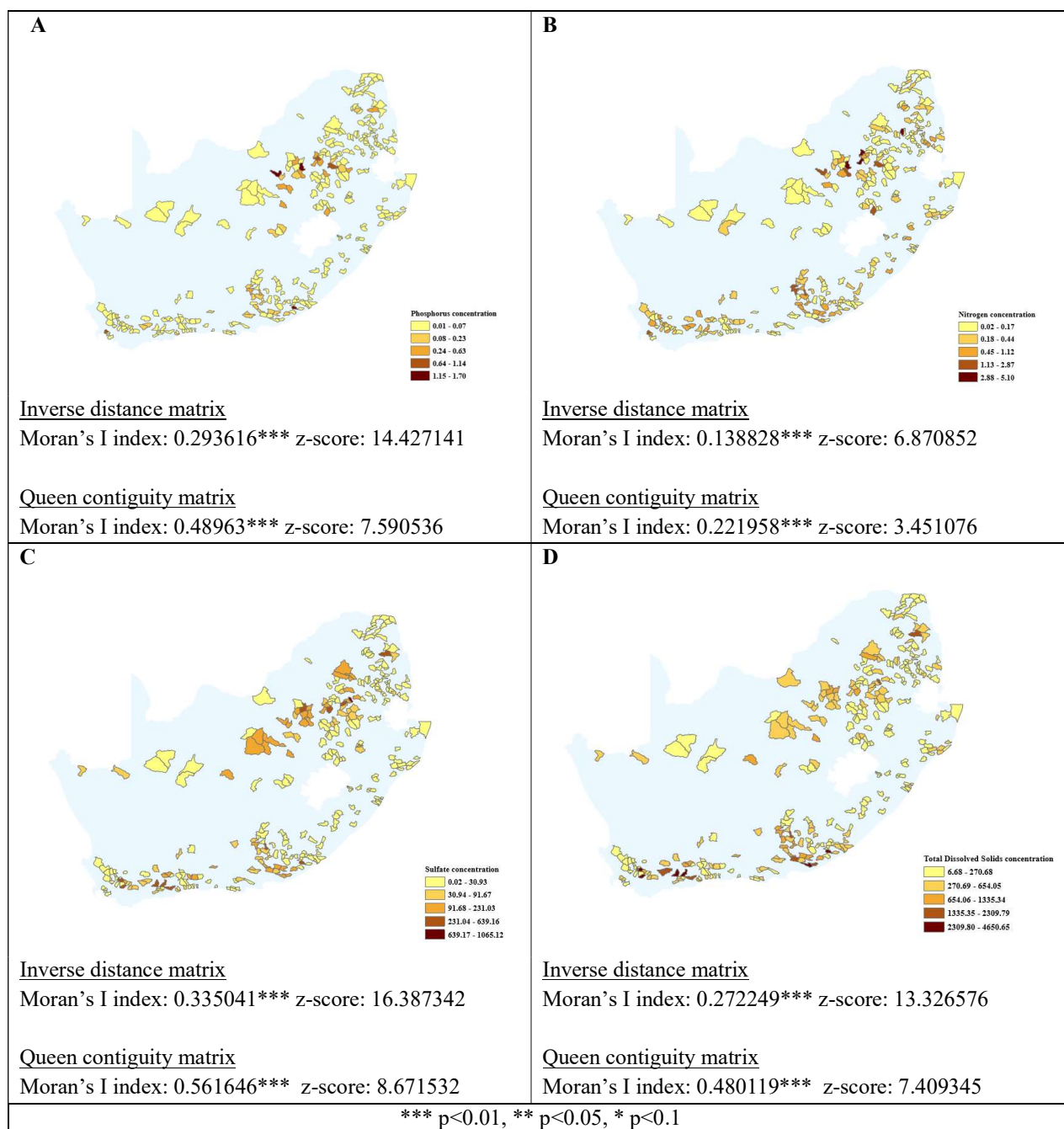
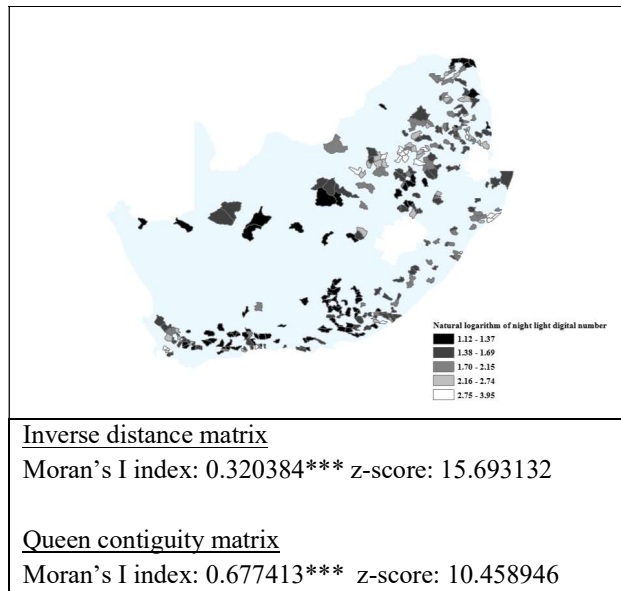




Figure 6-7 Cluster in the logarithm of night-light digital number



## *Appendix 6 - Section C: GWR results performance*

I present the model overall performance to be concise (Fotheringham, Brunsdon, & Charlton, 2003). No assumptions regarding the spatial correlation of water indices nor spatial correlations of economic development was applied when performing the GWR. The model was solely used as to investigate the spatial relationship between the economic variables and water pollution indices. Therefore, the low accuracy of the GWR is not a major issue. Nevertheless, the model provides results with no spatial cluster of the standard residuals.

On the other hand, using the cross-validation method instead of the AICc method to determine the best adaptative bandwidths of the weight matrix does not significantly change the insights of the interpretation. Although using the cross-validation method provide a higher R squared coefficient, the range of the coefficient linking the economic variables to water pollution indices is very large. It may suggest that using the bandwidth with a cross-validation method overfit the data.

Table 6-6 GWR performance with AICc estimation of bandwidth of the weight matrix

|                  | <b>Log TDS</b> | <b>Log SO4</b> | <b>Log PO4</b> | <b>Log NO3-NO2-N</b> |
|------------------|----------------|----------------|----------------|----------------------|
| <b>Neighbors</b> | 182            | 149            | 138            | 116                  |
| <b>AICc</b>      | 534.4223476    | -162.082893    | 3679.820716    | 4710.51513           |
| <b>R-squared</b> | 0.179936021    | 0.198779498    | 0.135180498    | 0.1762863            |

Figure 6-8 Residual distribution of the GWR

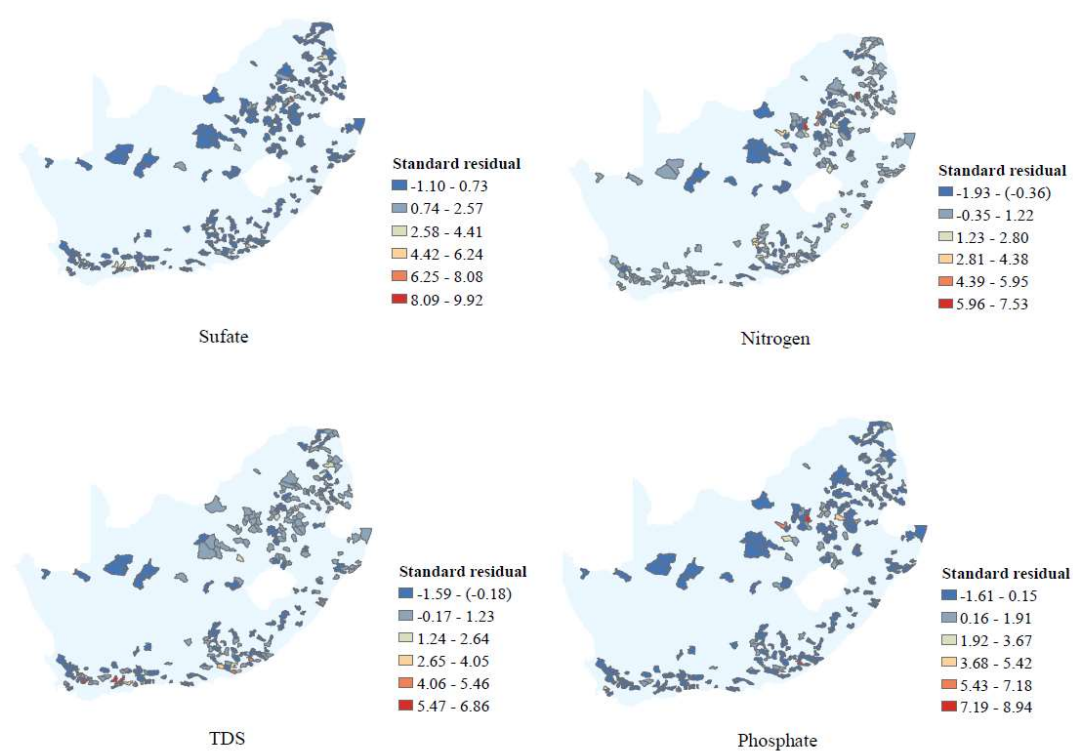


Figure 6-9 Average economic development impact on water quality

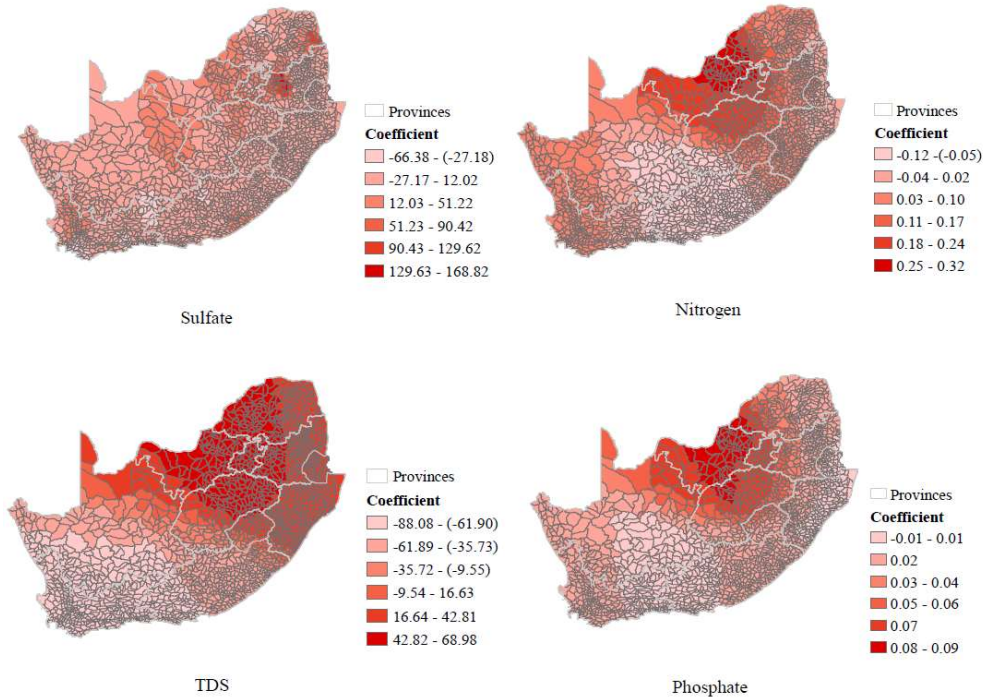


Figure 6-10 Economic growth impact on water quality

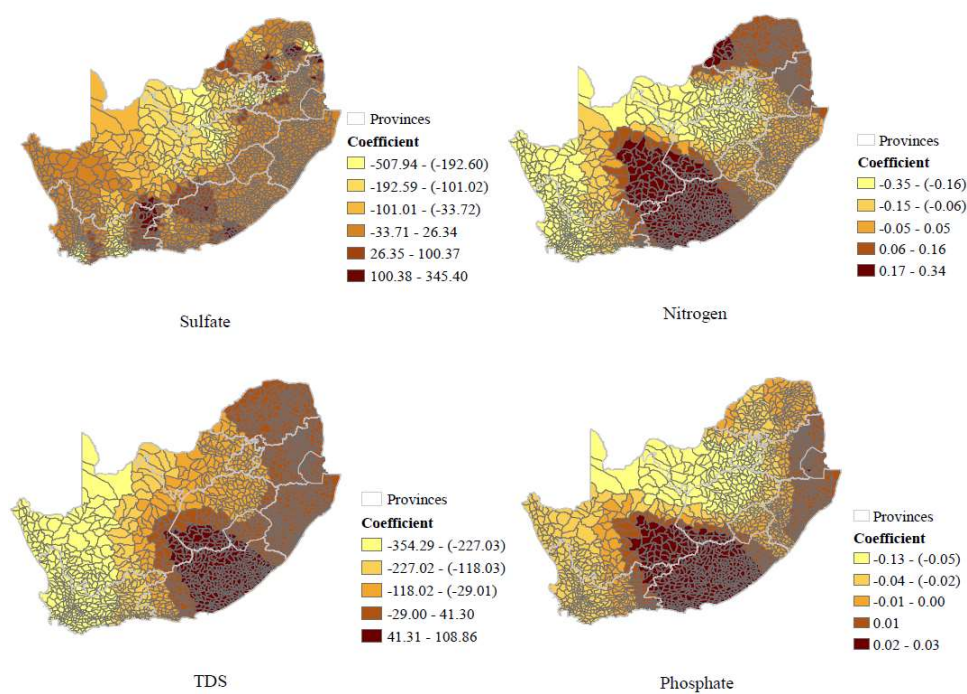


Table 6-7 GWR performance with cross-validation estimation of bandwidth of the weight matrix

|                  | <b>Log TDS</b> | <b>Log SO4</b> | <b>Log PO4</b> | <b>Log NO3-NO2-N</b> |
|------------------|----------------|----------------|----------------|----------------------|
| <b>Neighbors</b> | 141            | 18             | 118            | 97                   |
| <b>AICc</b>      | 4712.3484      | 3781.560       | -161.62093     | 535.745649175        |
| <b>R-squared</b> | 0.1936         | 0.5785         | 0.213642082    | 0.196686748702       |

Appendix 6 - Section D: Additional results

Table 6-8 SDM regression with log night-time light as the only human-related dependent variable

|                                  | Log TDS              | Log SO4              | Log PO4              | Log NO3-NO2-N        |
|----------------------------------|----------------------|----------------------|----------------------|----------------------|
| Log (light)                      | 0.033<br>(0.030)     | 0.081**<br>(0.035)   | 0.106***<br>(0.033)  | 0.101***<br>(0.036)  |
| Log (light growth)               | -0.117<br>(0.080)    | -0.278*<br>(0.148)   | -0.019<br>(0.093)    | -0.111<br>(0.100)    |
| River order                      | 0.124***<br>(0.044)  | 0.234***<br>(0.058)  | 0.099**<br>(0.048)   | 0.185***<br>(0.053)  |
| Total area (billions)            | -0.069<br>(0.772)    | 0.655<br>(1.058)     | -0.744<br>(0.734)    | -4.143***<br>(0.882) |
| Runoff (thousands)               | -2.460***<br>(0.535) | -1.584***<br>(0.558) | -0.787**<br>(0.395)  | -0.223<br>(0.382)    |
| Constant                         | 4.238***<br>(0.428)  | 0.635<br>(0.402)     | -4.217***<br>(0.337) | -2.878***<br>(0.469) |
| Spatial lag * Log (light)        | -0.475***<br>(0.099) | -0.514***<br>(0.122) | -0.510***<br>(0.103) | -0.402***<br>(0.144) |
| Spatial lag * Log (light growth) | 1.835*<br>(1.076)    | 2.746***<br>(0.977)  | 1.479**<br>(0.738)   | 1.070<br>(0.984)     |
| Spatial lag dependent variable   | 0.070***<br>(0.024)  | 0.275***<br>(0.065)  | -0.006<br>(0.032)    | -0.169**<br>(0.085)  |
| <i>R squared</i>                 | 0.36                 | 0.34                 | 0.24                 | 0.17                 |
| <i>Observations</i>              | 305                  | 305                  | 305                  | 305                  |

\*  $p < 0.1$ ; \*\*  $p < 0.05$ ; \*\*\*  $p < 0.01$

Table 6-9 The estimation results with the spatial Durbin model

|                                  | Log (TDS)            | Log (SO4)             | Log (PO4)                | Log (NO3+NO2-N)        |
|----------------------------------|----------------------|-----------------------|--------------------------|------------------------|
| Log (light)                      | -0.016<br>(0.032)    | 0.030<br>(0.037)      | -0.001<br>(0.030)        | 0.032<br>(0.037)       |
| Log (light growth)               | -0.053<br>(0.080)    | -0.236<br>(0.149)     | 0.127<br>(0.086)         | 0.003<br>(0.095)       |
| Urban coverage (thousands)       | 2.944<br>(2.211)     | -2.272<br>(2.895)     | 2.421<br>(3.624)         | 3.607<br>(3.726)       |
| Mining (thousands)               | 22.632***<br>(7.094) | 46.736***<br>(11.810) | 8.910<br>(11.595)        | -8.185<br>(7.344)      |
| Industry (thousands)             | 36.571**<br>(18.233) | 27.655<br>(30.667)    | 170.504***<br>(52.012)   | 139.726***<br>(42.171) |
| Agriculture coverage (thousands) | -0.770*<br>(0.453)   | 0.556<br>(0.453)      | 0.703**<br>(0.279)       | -0.020<br>(0.407)      |
| Population density               | 362.338<br>(289.556) | 575.165<br>(388.319)  | 1,052.65***<br>(398.069) | 515.781<br>(455.113)   |
| Total area (billions)            | 0.080<br>(0.819)     | -0.269<br>(0.970)     | -1.672**<br>(0.845)      | -4.121***<br>(1.010)   |
| Runoff (thousands)               | -2.331***<br>(0.478) | -1.262**<br>(0.490)   | -0.507<br>(0.353)        | -0.149<br>(0.368)      |
| River order                      | 0.117***<br>(0.043)  | 0.209***<br>(0.056)   | 0.073*<br>(0.044)        | 0.174***<br>(0.053)    |
| Constant                         | 4.164***<br>(0.390)  | 0.857**<br>(0.410)    | -3.749***<br>(0.302)     | -2.588***<br>(0.475)   |
| Spatial lag * log (light)        | -0.512***<br>(0.103) | -0.617***<br>(0.125)  | -0.568***<br>(0.092)     | -0.405***<br>(0.139)   |
| Spatial lag * Log (light growth) | 1.904**<br>(0.965)   | 2.044**<br>(1.009)    | 0.084<br>(0.607)         | 0.315<br>(1.027)       |
| Spatial lag dependent variable   | 0.066***<br>(0.023)  | 0.265***<br>(0.063)   | 0.038<br>(0.033)         | -0.091<br>(0.082)      |
| <i>Pseudo R-squared</i>          | 0.36                 | 0.38                  | 0.33                     | 0.17                   |
| <i>Observations</i>              | 305                  | 305                   | 305                      | 305                    |

\*  $p < 0.1$ ; \*\*  $p < 0.05$ ; \*\*\*  $p < 0.01$



Table 6-10 SDM with a dummy variable for the East region without the control variables  
(results associated to Table 6-3)

|                                                  | Log (TDS)            | Log (SO4)            | Log (PO4)            | Log (NO3+NO2-N)      |
|--------------------------------------------------|----------------------|----------------------|----------------------|----------------------|
| Log (light)                                      | -0.037<br>(0.037)    | -0.013<br>(0.041)    | 0.020<br>(0.032)     | 0.014<br>(0.040)     |
| East dummy * Log (light)                         | 0.743***<br>(0.133)  | 0.936***<br>(0.193)  | 0.900***<br>(0.216)  | 0.837***<br>(0.232)  |
| Log (light growth)                               | -0.240<br>(0.156)    | -0.215<br>(0.202)    | -0.064<br>(0.119)    | -0.154<br>(0.204)    |
| East dummy * Log (light growth)                  | 0.333**<br>(0.160)   | 0.096<br>(0.248)     | 0.230<br>(0.143)     | 0.252<br>(0.211)     |
| East dummy                                       | -1.232***<br>(0.256) | -1.676***<br>(0.343) | -1.750***<br>(0.347) | -1.598***<br>(0.403) |
| Total area (billions)                            | 0.820<br>(0.840)     | 2.124**<br>(1.066)   | 0.423<br>(0.813)     | -2.452***<br>(0.882) |
| Runoff (thousands)                               | -2.229***<br>(0.536) | -1.590***<br>(0.523) | -0.394<br>(0.313)    | -0.204<br>(0.408)    |
| River order                                      | 0.128***<br>(0.043)  | 0.237***<br>(0.056)  | 0.097**<br>(0.045)   | 0.182***<br>(0.051)  |
| Constant                                         | 4.562***<br>(0.661)  | 0.534<br>(0.664)     | -5.454***<br>(0.483) | -2.765***<br>(0.752) |
| Spatial lag * Log (light)                        | -0.417***<br>(0.157) | -0.431**<br>(0.192)  | -0.725***<br>(0.131) | -0.197<br>(0.213)    |
| Spatial lag * East dummy * Log<br>(light)        | -0.188<br>(0.519)    | 0.452<br>(0.562)     | 1.647***<br>(0.425)  | 0.410<br>(0.668)     |
| Spatial lag * Log (light growth)                 | 1.553<br>(1.680)     | 3.445**<br>(1.509)   | 3.779***<br>(0.858)  | 1.666<br>(1.383)     |
| Spatial lag * East dummy * Log<br>(light growth) | -0.508<br>(1.898)    | -2.796<br>(1.828)    | -5.647***<br>(1.275) | -3.272<br>(2.140)    |
| Spatial lag dependent variable                   | 0.057**<br>(0.024)   | 0.230***<br>(0.062)  | 0.016<br>(0.031)     | -0.127*<br>(0.076)   |
| <i>Pseudo R-squared</i>                          | 0.36                 | 0.34                 | 0.24                 | 0.17                 |
| Observations                                     | 305                  | 305                  | 305                  | 305                  |

Table 6-11 SDM with a dummy variable for the eastern region and control variables (Results associated to Table 6-11)

|                                               | Log (TDS)            | Log (SO <sub>4</sub> ) | Log (PO <sub>4</sub> ) | Log (NO <sub>3</sub> +NO <sub>2</sub> -N) |
|-----------------------------------------------|----------------------|------------------------|------------------------|-------------------------------------------|
| Log (light)                                   | -0.037<br>(0.036)    | -0.004<br>(0.040)      | -0.006<br>(0.031)      | -0.001<br>(0.040)                         |
| East dummy * Log (light)                      | 0.515***<br>(0.170)  | 0.691***<br>(0.228)    | 0.422*<br>(0.227)      | 0.666**<br>(0.260)                        |
| Log (light growth)                            | 0.000<br>(0.000)     | 0.000<br>(0.000)       | 0.000<br>(0.000)       | 0.000<br>(0.000)                          |
| East dummy * Log (light growth)               | 0.334**<br>(0.159)   | 0.093<br>(0.252)       | 0.206<br>(0.139)       | 0.227<br>(0.200)                          |
| East dummy                                    | -0.876***<br>(0.302) | -1.277***<br>(0.394)   | -0.991***<br>(0.360)   | -1.321***<br>(0.439)                      |
| Urban coverage (thousands)                    | 1.642<br>(2.370)     | -4.432<br>(2.800)      | 0.130<br>(3.291)       | 2.104<br>(3.383)                          |
| Mining (thousands)                            | 16.072**<br>(6.791)  | 38.918***<br>(11.470)  | -0.442<br>(12.094)     | -16.611<br>(8.441)                        |
| Industry (thousands)                          | 19.750<br>(21.226)   | 6.244<br>(31.817)      | 161.441***<br>(53.290) | 112.856**<br>(44.583)                     |
| Agriculture coverage (thousands)              | -0.719<br>(0.452)    | 0.543<br>(0.499)       | 0.675**<br>(0.274)     | -0.055<br>(0.391)                         |
| Population density                            | 33.101<br>(319.982)  | 162.021<br>(393.867)   | 713.604*<br>(396.738)  | 50.796<br>(506.912)                       |
| Total area (billions)                         | 0.920<br>(0.925)     | 1.236<br>(1.065)       | -0.812<br>(0.936)      | -2.530**<br>(1.067)                       |
| Runoff (thousands)                            | -2.234***<br>(0.517) | -1.411***<br>(0.510)   | -0.238<br>(0.335)      | -0.208<br>(0.395)                         |
| River order                                   | 0.125***<br>(0.043)  | 0.216***<br>(0.055)    | 0.081*<br>(0.042)      | 0.180***<br>(0.052)                       |
| Constant                                      | 4.399***<br>(0.618)  | 0.821<br>(0.675)       | -5.165***<br>(0.475)   | -2.661***<br>(0.753)                      |
| Spatial lag * Log (light)                     | -0.427***<br>(0.162) | -0.489**<br>(0.193)    | -0.837***<br>(0.110)   | -0.237<br>(0.205)                         |
| Spatial lag * East dummy * Log (light)        | -0.082<br>(0.492)    | 0.349<br>(0.555)       | 1.636***<br>(0.421)    | 0.455<br>(0.673)                          |
| Spatial lag * Log (light growth)              | 2.005<br>(1.537)     | 2.845*<br>(1.582)      | 2.606***<br>(0.863)    | 1.241<br>(1.429)                          |
| Spatial lag * East dummy * Log (light growth) | -0.880<br>(1.776)    | -2.328<br>(1.843)      | -4.740***<br>(1.328)   | -2.937<br>(2.200)                         |
| Spatial lag dependent variable                | 0.058**<br>(0.024)   | 0.226***<br>(0.062)    | 0.041<br>(0.032)       | -0.074<br>(0.074)                         |
| <i>Pseudo R-squared</i>                       | 0.39                 | 0.39                   | 0.34                   | 0.2                                       |
| <i>Observations</i>                           | 305                  | 305                    | 305                    | 305                                       |

\*  $p < 0.1$ ; \*\*  $p < 0.05$ ; \*\*\*  $p < 0.01$

Table 6-12 The impacts of economic development on water quality in the eastern regions with the control variables

|                                                                      |                                       | Log TDS   | Log SO4   | Log PO4    | Log NO3-NO2-N |
|----------------------------------------------------------------------|---------------------------------------|-----------|-----------|------------|---------------|
| Direct<br>impact                                                     | Log (light) 2010                      | -0.037    | -0.009    | -0.007     | -0.0001       |
|                                                                      | East * Log (light) 2010               | 0.515***  | 0.711***  | 0.4256*    | 0.666**       |
|                                                                      | Log (light growth) 2005 - 2010        | -0.233    | -0.220    | -0.002     | -0.102        |
|                                                                      | East * Log (light growth) 2005 - 2010 | 0.332**   | 0.069     | 0.1975     | 0.237         |
| Indirect<br>impact                                                   | Log (light) 2010                      | -0.450*** | -0.608**  | -0.867***  | -0.223        |
|                                                                      | East * Log (light) 2010               | -0.059    | 0.591     | 1.709***   | 0.387         |
|                                                                      | Log (light growth) 2005 - 2010        | 2.094     | 3.477*    | 2.697***   | 1.176         |
|                                                                      | East * Log (light growth) 2005 - 2010 | -0.907    | -2.871    | -4.899***  | -2.784        |
| Total<br>impact                                                      | Log (light) 2010                      | -0.488*** | -0.617*** | -0.8741*** | -0.223        |
|                                                                      | East * Log (light) 2010               | 0.455     | 3.256*    | 2.134***   | 1.074         |
|                                                                      | Log (light growth) 2005 - 2010        | 1.862     | 1.302*    | 2.695***   | 1.053         |
|                                                                      | East * Log (light growth) 2005 - 2010 | -0.575    | -2.801    | -4.701***  | -2.547        |
| Robust standard errors in parentheses *** p<0.01, ** p<0.05, * p<0.1 |                                       |           |           |            |               |

Table 6-13 The average impact of economic development on water quality in former black-dominated areas with control variables (results associated to Table 6-4)

|                                                           | Log<br>(TDS)         | Log (SO <sub>4</sub> ) | Log (PO <sub>4</sub> ) | Log<br>(NO <sub>3</sub> +NO <sub>2</sub> -N) |
|-----------------------------------------------------------|----------------------|------------------------|------------------------|----------------------------------------------|
| Log (light)                                               | -0.004<br>(0.032)    | 0.048<br>(0.038)       | 0.013<br>(0.032)       | 0.041<br>(0.039)                             |
| White dominance dummy * Log (light)                       | 0.010<br>(0.077)     | 0.035<br>(0.094)       | 0.220**<br>(0.100)     | 0.255*<br>(0.142)                            |
| Log (light growth)                                        | -0.020<br>(0.078)    | -0.216<br>(0.148)      | 0.108<br>(0.087)       | 0.023<br>(0.094)                             |
| White dominance dummy * Log (light growth)                | -1.501***<br>(0.303) | -1.251***<br>(0.307)   | -0.634<br>(0.438)      | -1.670***<br>(0.540)                         |
| White dominance dummy                                     | 0.605<br>(0.372)     | 0.754*<br>(0.418)      | 0.958*<br>(0.527)      | 0.966<br>(0.624)                             |
| Urban coverage (thousands)                                | 2.408<br>(2.222)     | -3.437<br>(2.699)      | -0.096<br>(3.145)      | 1.540<br>(2.953)                             |
| Mining (thousands)                                        | 21.534***<br>(7.003) | 45.763***<br>(11.857)  | 9.075<br>(11.496)      | -7.324<br>(6.841)                            |
| Industry (thousands)                                      | 37.593<br>(23.901)   | 17.930<br>(33.423)     | 129.634**<br>(51.626)  | 108.927**<br>(43.057)                        |
| Agriculture coverage (thousands)                          | -0.935**<br>(0.452)  | 0.288<br>(0.451)       | 0.742***<br>(0.274)    | 0.028<br>(0.396)                             |
| Population density                                        | 173.980<br>(275.680) | 285.220<br>(360.769)   | 758.854*<br>(388.977)  | 257.326<br>(484.533)                         |
| Total area (billions)                                     | 0.383<br>(0.821)     | 0.215<br>(0.966)       | -1.777**<br>(0.821)    | -4.121***<br>(0.990)                         |
| Runoff ('000)                                             | -2.374***<br>(0.477) | -1.378***<br>(0.512)   | -0.345<br>(0.339)      | 0.008<br>(0.379)                             |
| River order                                               | 0.110***<br>(0.042)  | 0.211***<br>(0.055)    | 0.085**<br>(0.043)     | 0.171***<br>(0.052)                          |
| Constant                                                  | 4.203***<br>(0.393)  | 0.887**<br>(0.418)     | -3.739***<br>(0.290)   | -2.478***<br>(0.457)                         |
| Spatial lag of Log (light)                                | -0.483***<br>(0.135) | -0.533***<br>(0.160)   | -0.781***<br>(0.112)   | -0.562***<br>(0.181)                         |
| Spatial lag of light growth                               | 1.646*<br>(0.964)    | 1.604<br>(1.024)       | -0.160<br>(0.621)      | 0.142<br>(1.057)                             |
| Spatial lag of Log (light) * White dominance dummy        | 0.695<br>(0.730)     | 1.327<br>(0.967)       | 2.740***<br>(0.920)    | 1.614<br>(1.106)                             |
| Spatial lag of Log (light growth) * White dominance dummy | 9.174*<br>(4.900)    | 16.801**<br>(7.243)    | 7.448<br>(6.280)       | 4.443<br>(7.426)                             |
| Spatial lag of dependent variable                         | 0.060***<br>(0.023)  | 0.246***<br>(0.062)    | 0.019<br>(0.032)       | -0.072<br>(0.077)                            |
| <i>Pseudo R-squared</i>                                   | 0.39                 | 0.41                   | 0.35                   | 0.2                                          |
| Observations                                              | 305                  | 305                    | 305                    | 305                                          |

\*  $p < 0.1$ ; \*\*  $p < 0.05$ ; \*\*\*  $p < 0.01$

## Chapter 7 Conclusion

This Thesis tackles the need for national and international governments to address spatial economic inequality. These needs stem from the fact that spatial economic inequalities are linked to corruptions (Jong-sung & Khagram, 2005), risks of civil wars (Østby et al., 2009) and lower national economic development (Alesina et al., 2016).

This Thesis aimed at providing the relevance of satellite data in the creation and the assessment of policies tackling spatial inequalities since efforts to address these inequalities are faced with issues of collection of reliable and comparable data across regions and countries. Hence, this Thesis, composed of five case studies, used satellite images to address spatial inequalities issues whose disentanglements were prevented by a lack of data at the fine scale. In the first study, I estimate the multinational and heterogeneous impact of air pollution on well-being when air pollution is measured at a small scale to advise international policies that attempt to put a price on air pollution across nations. Furthermore, I use satellite generated data to estimate the variation of capital stock in Africa at the sub-national level to assess any subnational inequity of sustainability. Afterwards, controlling for differences in economic activities between slums and rural areas in slum, I investigated whether the differences of well-being between these two types area remain in Greater Mumbai. Then while taking South Africa as a case study, the fourth and fifth studies are assessments of (1) spatial rebalancing policy

performed to improve water quality nationwide and (2) place-based policy to leverage the benefits of the FIFA World Cup in host cities respectively. In addition to support the relevance of satellite data, findings provide insights for the creations and the assessment of policy.

In **Chapter 2**, the Thesis shows that living in developing countries, and in countries with high income inequality may reduce access to preventive measures against air pollution, because only a small share of the population can afford them. Thus, the results suggest that the unique use of the economic valuation of clean air to create atmospheric policies may be ineffective to tackle the heterogeneous burden of air pollution. Within countries, low-income populations that live in highly polluted areas bear the highest impact of air pollution on both their well-being and their health. Moreover, the results show that populations living in highly polluted and productive countries may be reluctant to engage in international measures against air pollution because national environment participates in habituation to high levels of pollution on their well-being. This habituation to pollution may be a threat to international collaboration against pollution. Thus, to fight air pollution at the international level effectively, incentives that extend beyond well-being are necessary.

In **Chapter 3**, the thesis shows that although the wealth of African provinces increased on average by 3% between 2012 and 2018, wealth in provinces with African national capital cities increases faster than that of secondary provinces by 8%, on a per capita basis. This higher-

than-national-average IW growth is connected to a difference in ecosystem wealth and in part to a difference in levels of income. It does not imply governmental preference for investment in these provinces. This result can be used, in general, to improve accountability in need-based investment strategies from national to local level. Thus, it helps appraising several sustainable development goals (SDGs) developed by the UN. The higher sustainability performance in provinces with capital cities suggest that African countries follow a development path inconsistent to SDG 10 (Reduce inequality within and among countries). Thus, African countries will need to reconsider their development path to achieve these SDG.

In **Chapter 4**, the Thesis shows that slum residents are happier and more satisfied than rural populations in Greater Mumbai. This result may suggest that living near wealthy formal urban areas enables slum dwellers to near higher socio-economic lifestyles, and higher opportunity of upward social mobility. This higher happiness in slums may mean that rural populations will be willing to migrate toward slums if where to considerate well-being as proxy of individual's utility. Our results, exhibiting a higher SWB in slums relative to rural areas, explain why slum-dwellers refrain from relocating in rural areas and instead remain in slums. Thus, policies aiming at relocating slum-dwellers in rural areas are likely to fail.

In **Chapter 5**, the Thesis shows that cities that hosted the FIFA World Cup had considerably lower light growth than other South African cities, but this does not reflect a

difference in overall income growth between host cities and other South African cities. This finding is in line with most of the scholarly ex-post analyses on mega-events at the city-level suggesting that the World Cup is not linked to higher economic growth at the city-level. Subsequently, I found that host cities whose local government spent the most on stadium renovation and construction had a lower income growth than other host cities. It means that large investment on sport facilities would have been better invested in other part of the economies of host cities. In summary, the FIFA World Cup does not appear to be effective for place-based policies aiming to boost economic growth.

In **Chapter 6**, by using fine-scale data of water pollution and satellite monitored land cover, the Thesis reveals that economic development through urban settlements, mining, industrial and agricultural activities pose a risk for water quality in South Africa. However, it shows that economic development can strengthen water policies, thereby significantly reducing water pollution with satellite night light data used as an index of economic development. After controlling for the effect of land use and the effect of water pollution from neighbouring areas, I found that areas that had the best water policies during the apartheid era seem to also have the best water policies currently. Although the results do not imply current racial discrimination, they show the inefficiency of the government in totally rectifying spatial inequalities in water quality, which is one of their primary objectives since 1994.



## References

- Alesina, Alberto, Stelios Michalopoulos, and Elias Papaioannou. 2016. "Ethnic Inequality." *Journal of Political Economy* 124 (2): 428–88. <https://doi.org/10.1086/685300>.
- Jong-sung, You, and Sanjeev Khagram. 2005. "A Comparative Study of Inequality and Corruption." *American Sociological Review* 70 (1): 136–57. <https://doi.org/10.1177/000312240507000107>.
- Østby, Gudrun, Ragnhild Nordås, and Jan Ketil Rød. 2009. "Regional Inequalities and Civil Conflict in Sub-Saharan Africa." *International Studies Quarterly* 53 (2): 301–24. <https://doi.org/10.1111/j.1468-2478.2009.00535.x>.