INSTITUTO
UNIVERSITÁRIO
DE LISBOA

# A Regional Development Index for Portugal revisited

Manuel João Duarte Cruz

Master in Political Economy

Supervisor:
Doctor Ana Margarida Mendes Camelo Oliveira Brochado,
Associate Professor (with Aggregation) at
ISCTE - Instituto Universitário de Lisboa

September, 2023

Departament of Political Economy

# A Regional Development Index for Portugal revisited

Manuel João Duarte Cruz

Master in Political Economy

Supervisor:
Doctor Ana Margarida Mendes Camelo Oliveira Brochado,
Associate Professor (with Aggregation) at
ISCTE - Instituto Universitário de Lisboa

September, 2023

# Acknowledgements

This dissertation marks (for now, at least) the end of my school path. With such an important step that is set to be concluded, I shall never forget the people that always stood behind my back.

My nuclear basis is aware of how much I value it. My first recognition would always have to go to my parents. Throughout my whole life, their common love, extreme support, and encouragement were vital to this path, regardless of all the divergences between them. I am also deeply grateful to Laura, not only for the love, but also for the tough love, both much needed and much appreciated in this phase. To all my friends, thank you for both for helping me on my dissertation and in leisure times. I will see you in my patio for a proper celebration.

Then, I would like to thank to the people that ISCTE-IUL brought into my life and made my stay in Lisbon such a positive one. Mainly, to my supervisor, professor Ana Brochado, whose expertise and guidance were fundamental to the development of this dissertation. It was an absolute pleasure to be supervised by her right since from the very beginning. At another level, I am thankful to professor Paulo Marques for being a permanent source of trust and dialogue; and to professor Fátima Suleman, the one to whom I presented my theme for the first time and told me to go ahead. This dissertation could not see the light of the day if it was not a brainstorm together with Laura Ramires. Therefore, this one is also for her.

Finally, it is also important to honour the life of those who are not physically with us. In this way, this dissertation is my homage to professor José da Silva Costa, a prominent name of Portuguese Regional and Public Economics, for the excellent personal and academic contact that I had with him during my bachelor years.

# Abstract

Whether we consider policymaking at any degree or at the academia, regional development has been an extensive topic for decades. This subject is, perhaps, more important for countries like Portugal, as its chronic problems of regional asymmetries between the littoral and interior regions bring increasingly damaging impacts: while urban coastal pression is faced, desertification of the countryside seems to be unstoppable as well. Moreover, this decade has brought a wide panoply of phenomena: from the increase of public investment in areas that had not been seen as relevant before (e.g., energetic transition) to the new shape of NUTS since 2015, and even the more recent COVID-19 pandemic that alerted us to the significant asymmetries that prevail. We aim to conceptualise regional development, to analyse regional asymmetries in Portugal, and to understand what are its main determinants of regional variability. Gathering data from 2013 to 2021, this study employs a two-step approach: first, it updates a Portuguese Regional Development Index augmented with new governance and environment dimensions, whose results reveal that the new dimensions are of utmost importance for the chronically more depressed regions. Second, it uses our index scores to provide panel data and neural network estimations, whose findings advocate that economic dimensions, such as unemployment and productivity, are key to understand the current picture. As far as we know, this study is the first one to use index calculations, panel data, and neural networks in the study of regional development in Portugal.

**JEL-Codes:** O15, P25, R11, R58

**Keywords:** human development index, regional development and asymmetries, Portugal

# Resumo

Quer se considere na elaboração de políticas ou na academia, desenvolvimento regional tem sido um tópico extenso há décadas. Este assunto é, talvez, mais importante para países como Portugal, visto que os seus problemas crónicos de assimetrias regionais entre as regiões do litoral e as do interior trazem impactos cada vez mais prejudiciais: enquanto se enfrenta uma pressão urbana costeira, a desertificação do interior parece também ser imparável. Além do mais, esta década trouxe uma vasta panóplia de fenómenos: desde ao aumento de investimento público em áreas que não eram anteriormente vistas como relevantes (e.g., transição energética) até ao novo desenho das NUTS desde 2015, finalizando com a pandemia de COVID-19 que nos alertou para as assimetrias significantes que prevalecem. Pretendemos conceptualizar o desenvolvimento regional, analisar assimetrias regionais em Portugal, e perceber quais são os principais determinantes de variabilidade regional. Compilando dados de 2013 a 2021, este estudo emprego uma abordagem de dois passos: primeiro, atualiza um Índice de Desenvolvimento Regional Português aumentado com novas dimensões de governança e ambiente, cujos resultados revelam que as novas dimensões são de mais elevada importância para as regiões cronicamente mais deprimidas. Segundo, usa os resultados do nosso índice para fornecer estimações de dados em painel e de redes neurais, cujas descobertas advogam que dimensões económicas, como desemprego e produtividade, são chave para entender a imagem atual. Até onde sabemos, este estudo é o primeiro a usar cálculos de um índice, dados em painel, e redes neurais no estudo de desenvolvimento regional em Portugal.

**Códigos-JEL:** O15, P25, R11

**Palavras-chave:** índice de desenvolvimento humano, desenvolvimento e assimetrias regionais, Portugal

# Contents


Introduction ........ 1
Chapter 1. Literature review ........ 3
1.1 A kickstart – the concept of development ........ 3
1.2 Stances on regional development ........ 4
1.2.1 Regional development as a field of study ........ 4
1.2.2 Pivotal cases of regional policy ........ 6
1.2.3 Regional development and intergovernmental organisations ........ 7
1.3 Ground-breaking measurements – the Human Development Index ........ 9
1.4 Not up to par – critiques to the Human Development Index ........ 10
1.5 Measuring alternatives to the Human Development Index ........ 11
1.6 Drivers of regional development ........ 13
1.7 Answering the “where?” question – the concepts of region and NUTS ........ 14
1.8 Research context: what progress has been done up to now? ........ 15
Chapter 2. Methodology ........ 17
2.1 A composite index step by step ........ 17
2.2 The model, research design, and secondary data collection ........ 20
2.3 A second estimation method: neural network ........ 26
Chapter 3. Results and discussion ........ 27
3.1 Regional asymmetries in Portugal – results and discussion ........ 27
3.2 Panel data models. What shapes a developed region? ........ 33
3.3 Neural network: a second perspective ........ 36
Chapter 4. Conclusions ........ 38
References ........ 41
Annexes ........ 47

# List of abbreviations

| | |
|---|---|
| **APA** | *Agência Portuguesa do Ambiente* |
| **ATM** | Automated teller machines |
| **EU** | European Union |
| **EIB** | European Investment Bank |
| **GDP** | Gross Domestic Product |
| **GNI** | Gross National Income |
| **HDI** | Human Development Index |
| **HDR** | Human Development Report |
| **ICC** | Intraclass Correlation Coefficient |
| **ICT** | Information and Communications Technology |
| **IMF** | International Monetary Fund |
| **INE** | *Instituto Nacional de Estatística* (Statistics Portugal) |
| **ISDR** | *Índice Sintético de Desenvolvimento Regional* (*Regional Development Synthetic Index*) |
| **MASST** | Macroeconomic, Sectorial, Social, Territorial |
| **NUTS** | Nomenclature of Territorial Units for Statistics |
| **OECD** | Organization for Economic Co-operation and Development |
| **OLS** | Ordinary Least Squares |
| **pc** | *per capita* |
| **PPP** | purchasing power parity |
| **PRDI** | Portuguese Regional Development Index |
| **SDGs** | Sustainable Development Goals |
| **UN** | United Nations |
| **UNCDR** | United Nations Centre for Regional Development |
| **UNDP** | United Nations Development Programme |

# List of figures

Figure 3.1: PRDI per NUTS III without the environmental and governance dimensions – 2013, 2017, and 2021. ........ 29
Figure 3.2: PRDI per NUTS III with the environmental and governance dimensions – 2013, 2017, and 2021. ........ 31

# List of tables

Table 2.1: dimensions of the PRDI, their respective indicators and subindices, and reference values for the forming of the PRDI. 19

Table 2.2: categories, their respective indicators, identification of the belonging models, and reference values. 21

Table 2.3: variables and respective hypotheses per econometric models. 25

Table 3.1: PRDI per NUTS III without environment and governance ranking differentials from 2013 to 2021. 30

Table 3.2: PRDI per NUTS III with environment and governance ranking differentials from 2013 to 2021. 32

Table 3.3: regression table. 35

Table 3.4: independent variables, their respective importance, and normalised importance for regional development in Portugal. 36

# List of equations

Equation 2.1: dimension subindex ........................................................................... 18
Equation 2.2: income subindex ................................................................................ 18
Equation 2.3: PRDI formula ...................................................................................... 18
Equation 2.4: panel data regression model ................................................................ 20

# Introduction

In a world increasingly marked by globalisation, global warming, and political change, it becomes even more necessary to consider all the dimensions that human action and policymaking influence. As such, the United Nations (UN) are prone to put into practice the 2030 Agenda for Sustainable Development, which encompasses 17 Sustainable Development Goals (SDGs) to promote prosperity and sustainable development in its three core dimensions - economic, social, and environmental - considering the rise of material life quality for the present generations without compromising life quality of the next ones, diminishing every type of social inequalities, and respecting nature and the ecosystems (United Nations, 2015).

Not only the world is diverse and complex, but so as countries hide intrinsic differences and disparities. Namely, regions are shaped by multiple aspects, such as amenities, accessibility, industry, or agglomeration effects. It thus become relevant to understand economic, social, political, and environmental phenomena at the regional level. In this sense, international organizations, such as the Organization for Economic Co-operation and Development (OECD), are committed to implement place-based policies that give thought to specific territorial assets and that can foster regional cohesion (OECD, 2011).

It becomes very relevant to analyse regional development inside Portugal, both in what has been happening in the last decades and in what is being planned for the upcoming ones. Every ten years, Statistics Portugal presents the Census and the 2021 edition shows that the country is struggling with an ageing and declining population, a demographic structure that puts pressure on the coastal regions – highlighting the two metropolitan areas of Lisboa and Porto and the region of Algarve – and that empties municipalities from the interior ones (INE, 2022). Moreover, *Relatório do Estado do Ambiente 2022* shows that Portugal is highly exposed to climate change, especially with drought and fires (APA, 2023).

Yes, the reality is challenging, but not everything is bad news: for instance, the Portuguese population is more educated (INE, 2022), more than half of the electricity that was produced in the country came from renewable sources (APA, 2023), and European structural funds are on their way to be implemented for next years in several vital areas, such as digital and green transitions or socioeconomic resilience (*Ministério do Planeamento*, 2021). These are some of the very important attributes that regional development in Portugal will (have to) rely on.

Given this context, our key objectives for this dissertation are the following: first, to conceptualise regional development and to understand why it matters to see it with today's eyes; then, to demonstrate what is the current Portuguese panorama of regional development for the 25 NUTS III; in the end, to understand what are the dimensions – in Economics terms, the drivers or explanatory variables – behind regional performance.

This study revisited the composite index designated Portuguese Regional Development Index (PRDI), created by Silva and Ferreira-Lopes (2014), an index that aggregates the dimensions used by the United Nations – income, health, and education – and is augmented in two new dimensions that, according to the vast literature (see, for instance, Alkire, 2010) deserve more attention – environment and governance; after this, we use these regional-level scores and 29 independent variables to create panel data estimations and a neural network. These two last models count with four dimensions – Macroeconomic, Sectorial, Social, Territorial – inspired by Capello (2007). In this way, this dissertation points to suggest to policymakers what should be more or less considered for achieving more regional cohesion in Portugal.

The most important research questions that we try to answer to are the following: first, what is the current shape of Portuguese asymmetries, taking into account the five dimensions of the Portuguese Regional Development Index and, particularly, what is the impact of environment and governance dimensions? Second, what are the drivers behind such distribution?

Silva and Ferreira-Lopes (2014) are, undeniably, the main starting source of inspiration for this dissertation. But what were, after all, the literature gaps that we majorly consider in this dissertation? The authors clearly state that research could be made in both the update of their index on time and the use of econometric estimations to find the reasons why Portugal has such development path.

This dissertation is structured as follows. Chapter 1 presents a literature review of key concepts regarding (regional) development, development indices - taking into account their critiques and limitations -, drivers of regional development, and regional development in the Portuguese landscape. Chapter 2 displays the methodological procedures, first, to update and calculate the index, and, second, to create the estimations. Chapter 3 addresses results; particularly, the Portuguese Regional Development Index scores are used for determining the factors behind regional development, which is made via panel data models and a neural network. Chapter 4 presents concluding remarks with policy implications, limitations of this work, and possibilities for future research.

CHAPTER 1

# Literature review

For a proper understanding of "regional development" as a field of study and policymaking, it is vital to acknowledge its constituents. Thus, we kick off by defining "development" (section 1.1). Then, we explore important information regarding the historical setting of regional development (section 1.2), as seen from the perspective of the academia (section 1.2.1), of national governments (section 1.2.2), and of international organisations (section 1.2.3).

Following the understanding of these ideas and this field of study, it is also key to learn how development was – and is – measured in a first place, introducing the Human Development Index (section 1.3), the drawbacks of its methodology (section 1.4), and possible other suggestions to provide knowledge in this topic (section 1.5).

Since this is a general goal of this dissertation. we also provide a brief literature review focused on the determining factors of developed regions (section 1.6), based on the dimensions considered in the PRDI.

After a conceptual overview on regional development and on the Human Development Index, we aim to describe the general outlook of regional asymmetries between the Portuguese NUTS III. Therefore, our first step is to present some of the major aspects regarding regions as a concept and their several typologies, from which the contributes of Costa and Nijkamp (2009), Perroux (1950), and McCann (2013) are considered, addressing the Portuguese NUTS case (section 1.7). We end with some past Portuguese evaluation frameworks (section 1.8).

## 1.1 A kickstart – the concept of development

*Development* can be defined, in a nutshell, as the set of processes of changing prosperity (Medeiros, 2022). It is the process of expansion of capacities that contribute to the advancement of society through the realisation of individuals', firms', and communities' potential. It implies the expansion of human capital in its broadest sense, such that individuals can undertake active economic, social, and cultural roles (Feldman et al., 2016).

This is, not seldom, a concept that brings confusion to the academia and to the general public, due to the "easy" substitution of economic development by economic growth: the latter can be characterised as a quantitative concept measured through a global indicator, usually the real GDPpc change rate; the former is, mostly, a qualitative and normative concept, as it implies a structural transformation and leads to real satisfaction of individuals, not only through the material goods arising from economic growth, but also through non-tradeable, market means (Figueiredo et al., 2008). Thus, whereas economic growth is easily quantified as the increase of aggregate output, economic development is the means to create prosperity and increase citizens' quality of life (Feldman et al., 2016).

Sen (1983) affirms that a clear distinction between growth and development exists. In his point of view, growth is just means to get the other goal, which does not mean that it is not relevant, but that its importance comes with the benefits that emerge throughout the process of economic growth. To sustain his statement, he highlights that, for instance, the same levels of average life expectancy, literacy, health, and high education – 4 development indicators – can be seen in countries with different pc income levels. Moreover, he highlights that, beyond the increase of national output, aggregate income, and the supply of goods in an economy, the process of economic development happens through the expansion of peoples' *capabilities* and *rights*: whereas the former is related to what individuals can or cannot do, the latter represents the set of alternative goods that they can control with the use of all the opportunities that are facing them.

In sum, as stated by Brinkman (1995), the process of growth implies the replication of the same economic pattern, whereas development implies a structural transformation. The economic evolution gathers growth and development, but growth alone will not bring development.

## 1.2 Stances on regional development

In this section, regional development, the analysis of "complex space-time dynamics of regions" (Nijkamp and Abreu, 2009; Medeiros, 2022), is seen from the point of view of the academia and of policymaking, both at the national and at the international levels, taking into consideration an adequate multidimensional perspective.

### 1.2.1 Regional development as a field of study

It is true that space, surfaces, and distances support every economic activity. Economic geography is remerging with an important statement: although geography is not always a necessary condition for innovation, some environments are more likely than others to bring forth a stream of innovations (Desrochers, 1998, p. 79), and, thus, more likely to assure more processes of *change* or *evolution* – two Schumpeterian terms for the obtainment of economic growth and consequent development (Graça Moura, 2018).

Notwithstanding, spatial issues were often ignored by economists and economic policymakers[1]. Overall, space (and even time) is absent from mainstream, neoclassical Economics – economic theory, empirical analysis, and economic policy –, considering the whole economic system as assumed to take place on a "spaceless wonderland" (Desrochers, 1998, p. 63). Why is that so? The relevance of mainstream economic modelling asks for simple and predictable models with the assumptions of perfect competition, perfect factor mobility, and perfect price flexibility, aiming to achieve an equilibrium solution. Krugman (1992) put the non-consideration of space it this way:

> "*In international economics we take as our base case a world in which resources are completely immobile but in which goods can be costlessly traded. We may then modify the model to introduce transport costs or non-traded goods, on one side, or mobile factors, on the other, but the modelling style is clearly determined by the base case. And as anyone who has done economic theory knows, the style of our models strongly determines their content – issues that are awkward to address are generally speaking not addressed.*" (Krugman, 1992, p. X).

Regional economics was, after all, out of sight of mainstream economics. However, adequate tools to deal with increasing returns, external economies, and imperfect competition arrived. The author adds, though:

> "*... Urban and regional economists who worry about geographical issues... are almost uniformly peripheral to the economic profession. (...) They may do excellent work, but it does not inform or influence the economics profession.*" (Krugman, 1992, p. 3-4).

Even if some sides of the academia are prone to release controversies, what is known to be true is that regional economics deals with areas that are larger than cities. The central questions of regional economics focus on the reason why different parts of the country behave in different ways:

[1] There are some valuable exceptions, such as Ricardo, Von Thünen, Weber (Desrochers, 1998), and Marshall (Desrochers, 1998; Belussi and Caldari, 2008).

why some of the regions are rich, developed, and leading; why others are poor, underdeveloped or lagging; and how do they all relate between themselves (McCann, 2013).

One last example of the enormous contemporary relevance of this field of study can be exemplified by Medeiros (2022). The author provides, at a first place, a sum up of several economic and semi-economic centric regional development theories, explaining its advantages and, especially, its disadvantages. From this point, a strategic-based regional development theory is presented, being described as an alternative, integrated, and “more comprehensive theory of everything”, encompassing spatial/territorial planning and governance importance. In total, this theory rests on 6 dimensions with 5 respective cores – sustainable-, institutional-, knowledge-, place-, balanced-, and infrastructural-based.

### 1.2.2 Pivotal cases of regional policy

Despite the general awareness of international organisations towards “our common future”, regional development has been a priority for national governments throughout, at least, the last century. Two examples based on the *place prosperity-people prosperity* dualism are displayed.

Suffering deep effects due to the national economic collapse of the 1930s and being one of the most depressed agricultural regions in the United States of America, the Tennessee Valley was the target of a vital case of regional development policy. Belonging to the New Deal development program, the Tennessee Valley Authority was created in order to provide a gradual, structural transformation in the economic profile of the region, departing from a more modern and productive agriculture to boost the industrial sector, improving the general conditions of living of people, which were very low. One of the strongest measures was the massive construction of dams, which provided not only water, but also generalised energy at low costs for families and companies (Barbour, 1937; Beth, 1993). This is an accurate example of a *place prosperity* policy, a set of public expenditures – for instance, grants to local governments or businesses – that aim to increase infrastructure, private capital, education, and worker training in particular places, aiming to take advantage of the region’s comparative advantage (Bolton, 1992).

In a similar time period and in a similar context of increasing unemployment, it was also on the rise the relocation of labour force from depressed regions to more prosperous ones. In 1928, the Industrial Transference Board was created in the United Kingdom, assuring a larger encouragement of mobility and relocation. How was this made? The Board assured training for the workers that had different functions in the past and, more importantly for relocation factors, assured the payment of all travelling costs to all the workers (Wren, 1996). This, in contrast of the former case, is an example of *people prosperity* policy, simply through supporting the transference of unemployed people from declining regions to other ones with higher job opportunities, regardless of they live (Bolton, 1992).

### 1.2.3 Regional development and intergovernmental organisations

In the last decades, we have been witnessing a change:

> *"Whereas once urban and regional issues were seen by many economists to be a minor avenue of research, nowadays many of the most important international institutions including the World Bank, the OECD, the European Commission, and the United Nations are all grappling with the economic challenges and possibilities associated with regions and cities."* (McCann, 2013, p. XVII).

In 2015, the United Nations adopted the 2030 Agenda for Sustainable Development, composed by 17 SDGs and spread by 5 domains of relevance – people, planet, prosperity, peace, and partnership (United Nations, 2015), which can be described as "a blueprint for global welfare for current and future generations" (Tsani et al., 2020, p. 572). Narrowing, the Centre for Regional Development (UNCDR)[2] addresses the importance of promoting "sustainable regional development in developing countries with a focus on development planning and management in the context of globalisation and decentralisation trends, and the growing concern towards global environmental issues and their impacts". The UNCRD is very much focused on providing research and collecting data across several thematic areas that can be used, *grosso modo*, by developing countries, those that are structurally more needed of aid.

In addition, financial institutions, such as the IMF or the World Bank, focus their activity on this field. More specifically, they also assure multiple updated reports[3] with respect to regional development. Specifically, a very recent World Bank book – *Place, Productivity, and Prosperity:*

[2] Retrieved from UNCRD | Department of Economic and Social Affairs.
[3] See, for instance, the Regional Economic Outlook of the IMF here, Regional Economic Outlook (imf.org).

*Spatially Targeted Policies for Regional Development* -, written by Grover et al. (2022), was published. They point out the major motivations as follows:

> "*Place matters for productivity and prosperity. Myriad factors support a successful place, including not only the hard infrastructure such as roads, but also the softer elements such as worker skills, entrepreneurial ability, and well-functioning institutions. History suggests that prosperous places tend to persist, while "left-behind" regions—or those hurt by climatic, technological, or commercial shocks—struggle to catch up."* (back cover)

With this book, updated analytical and empirical insights are provided on the 3 drivers of economic geography – agglomeration economies, migration, and distance – thus being possible to analyse how they are correlated. Moreover, it is highlighted that these interactions vary substantially when developing countries are compared with developed ones, as it is seen that, in the former, there has been a poorer adaptation of cities to rural exodus, not leaving opportunities to structural transformation and increasing levels of productivity to prevail. Finally, the book emphasizes the role of place-based policymaking, consequently being easier to assess where or whether the efforts might be noted the most (Grover et al., 2022).

The OECD[4] is another example of an international organisation whose focus also lies on regional development, as regions are considered by them as the optimal territorial level to implement development policies (OECD, 2010). It is put by them as "a general effort to enhance well-being and living standards in all region types, from cities to rural areas, and improve their contribution to national performance and more inclusive, resilient societies", applying their efforts on the interactions of regional development with the economic recovery after the COVID-19 pandemic, digitalisation and innovation, low-carbon economic transition, and socioeconomic inequalities.

Needless to say that this subsection could not end without mentioning the vital role of European entities on the most recent trends of regional development policymaking, as the European Union, in its several components, is putting more and more efforts into the elaboration of cohesion policies and into the allocation of structural funds. For instance, the European Commission[5] states that the European

[4] Retrieved from Regional Development Policy - OECD.
[5] Retrieved from European Regional Development Fund - Regional Policy - European Commission (europa.eu).

Regional Development Fund, established in the time period of 2021 to 2027, will strengthen Europe in indispensable areas, such as innovation and digitalisation, climate, mobility, citizenry, education, and healthcare. As for the European Investment Bank[6], its scope lies on prioritising the less developed regions, that is, those which have a GDP per inhabitant that is less than 75% of the EU average. Moreover, we can get a proper idea of the relative importance in terms of economic sectors. For instance, the largest proportions of lending between 2014 and 2020 are related to small and medium companies, transport and energy, with, respectively, 33%, 18.8%, and 14.5% (EIB, 2021).

## 1.3 Ground-breaking measurements – the Human Development Index

After acknowledging the insufficiencies of assessing development levels only by the real GDPpc (see Cornia, 2003), the abovementioned "capabilities-based approach" suggested by Sen started to be operationalised through the Human Development Index (hereafter, HDI). This index encompassed several fundamental dimensions that are much more related to the quality of life as ends, rather than unidimensional means.

It was in the first Human Development Report (HDR), in 1990, that the HDI appeared for the first time (UNDP, 1990). The report suggests three "essential elements of human life": longevity, knowledge, and decent living standards (op. cit., pp. 11-12), respectively measured by life expectancy at birth, adult literacy, and real GDPpc PPP.

Since 1990, there has been yearly publications of HDRs. Throughout the years, the HDI has suffered some key methodological changes in terms of data collection, compilation, and calculation, highlighting the income component as the most intervened one (Kovacevic, 2010). Until 2009, the old method used the arithmetic mean of suitably normalized values for life expectancy, educational attainment, and income. Many solid criticisms were made, especially in terms of the choice of the arithmetic mean and consequent perfect substitution across dimensions (see Herrero et al., 2010, for detailed explanations).

From 2010 onwards, the new method gives equal relative importance to all three dimensions, and both education elements are equally weighted. Thus, the health subindex is measured by life expectancy at birth; the education subindex is a combination of the expected schooling years of a child entering in school at a given time, on the one hand, and of the mean schooling years for adults

[6] Retrieved from Regional development & cohesion (eib.org).

over-25 years old, on the other; and the income subindex is provided by the GNIpc PPP (UNDP, 2010). Finally, the HDI is calculated as a geometric mean of the three dimensions, thus embodying imperfect substitutability across all HDI dimensions and correcting one of the most serious critiques of the linear aggregation formula, which allowed for perfect substitution (UNDP, 2010). All the formulas and steps are shown in section 2.2, with the difference of considering the five dimensions used by Silva and Ferreira-Lopes (2014).

## 1.4 Not up to par – critiques to the Human Development Index

In a nutshell, we can separate the critiques to the HDI in three main areas. When it comes to methodology, they refer the lack of subjectivity in the weighting attribution to each dimension and data, a non-desirable number of cases of similarly ranked countries that have very different results in each dimension, and a high correlation between HDI and its components, making it so that the economic indicator provides a large part of the information. Then, the HDI might face autocorrelation problems – when a score is partially dependent of the former – as the time lag of some measures is such that health and education stock variables are difficult to vary in the short run. Finally, we have the selection of areas of analysis, an eternal discussion on the most relevant, practical, and accurate ones to assess human development (Omrani et al., 2020).

Even in the first HDR, the UNDP (1990) easily was recognising some issues related to the measurement of the HDI:

> "*All three measures of human development suffer from a common failing: they are averages that conceal wide disparities in the overall population. Different social groups have different life expectancies. There often are wide disparities in male and female literacy. And income is distributed unevenly."* (op. cit., p. 12).

An argument from Ravallion (2012) states that the improvements done by the UNDP since the 2010 HDI are, to a large extent, "more complicated and more in its trade-offs across dimensions" (op. cit., p. 208). Namely, the author raises concerns on the increasing devaluation of longevity and in the incapability of the new HDI to eliminate the perfect substitutability property of the old HDI.

Lind (2019) states a critical remark on the minimum and maximum values of the four HDI components used for ranking purposes, arguing that these are unjustified, unnecessary, distort the gradings, and are, to a large extent, arbitrary and ambiguous – for instance, with the assumption that income, health, and education are equally important. If three dimensions are aggregated in the way that the HDI is, it can happen that countries wish to manipulate the weights of certain dimensions in detriment of others, as they may be perfectly substitutable (Desai, 1991; Sagar and Najim, 1998; Kovacevic, 2010).

Another critique that is commonly pointed out to the HDI is related to its elementary form of just encompassing just three human development aspects (Herrero et al., 2010). It is inevitable, of course, that all similar indices shall not capture all the scopes. However, for instance, issues of gender – income, gender, race – are not considered in these accounts, even though there are mentions in the Human Development Reports through additional indices (Lind, 2019). The vast majority of the literature puts environment and governance as the two main dimensions that are not – and should be – considered in the calculus of the HDI (Alkire, 2010, as cited in Silva and Ferreira-Lopes, 2014).

## 1.5 Measuring alternatives to the Human Development Index

After understanding some of the most important critiques that have been done in more than three decades to the HDI and seeing several potential regional development determinants, it is important to set some possible alternatives for the measurement of development[7]. The degree of variety is quite substantial, as we show several indices that have the same dimensions and number of indicators, yet differing, for instance, in terms of data collection and calculation methods[8], and we show many other, more distinguishable contributions.

Even before the 1990 HDR, an early attempt was brought by Morris (1979), through the Physical Quality of Life Index, which consisted in a weighted average of literacy, infant mortality, and life expectancy (Ravallion, 2010).

An axiomatic, generalised development index is presented by Chakravarty (2003), putting the HDI as a special case. He proposes axiomatic measures for assessing achievement, underlining that they help isolating dimensions according to their degrees of sensitivity to wellbeing, through the

[7] For the complete list of indices, see Annex B. Section 2.1 sets up alternatives implemented for the Portuguese case.
[8] To deepen this analysis does not fit in the purposes of this dissertation.

calculations of percentual contribution made by each dimension to overall achievement, which displays a big matter in terms of policymaking potential in augmenting the more fragile ones.

The OECD (2011) Better Life Index is an eleven-topic index with a more complex vision of development – actually, the concept of well-being is more considered. Opposing other indices, is not a static one. Rather, it allows everyone to calculate their own version based on the relative preferences that someone might have, thus providing a more interactive and flexible tool[9] (Greco et al., 2020).

Based on the contributions of Chakravarty (2003), Ravallion (2012) suggests an alternative index that avoids the trade-offs in the new HDI. It still considers the diminishing returns of income, but without the log transformation, and it now longer lies on the geometric mean of the two education dimensions, thus returning to their arithmetic mean.

Pereira and Mota (2016) present a Municipal HDI applied to the city of Recife through a much different method, ELECTRE TRI-C that is suitable to lower many of the problems used by the mainstream methodology, such as the compensatory effect and the calculation issues, and suitable to compare between years, reducing autocorrelation problems.

One particularly interesting contribution for the purposes of this dissertation comes from Permanyer and Smits (2018), as they display a study for the HDI with the same methodology, but applying at a subnational level.

Lind (2019) show a modified index of human development (Index H), using the same four statistics of the HDI, but altering the parameters, based on the assumption that people everywhere are free to choose how to spend their time and their income, whether it is on health, education, or something else. Income is seen as a mere means to the end of a free, long, and healthy life, thus reducing the relative importance of income to development in the Index H. Henceforth, utility is a good approximation of human development and life quality evolution occur when the statistics increase simultaneously.

The Planetary Pressures-Adjusted Human Development Index is the most recent tool of the UNDP. It consists in correcting HDI scores by the carbon dioxide emissions and material footprint, both at the pc level, making it so that, as the differentials between the two indices go lower, the better is the awareness of lowering the human pressure on the planet (UNDP, 2020).

[9] Your Better Life Index is available here: OECD Better Life Index.

## 1.6 Drivers of regional development

The literature on regional development and on the determining factors behind the former is broad. This is a very pertinent topic, as it raises awareness to decisionmakers on the lagging causes, in order to accurately address them and, henceforth, helping regions to become more prosperous. Let us make an analysis of some of the main drivers inside the five dimensions of the PRDI.

Black et al. (2013) have conducted a research on the development of the Baltic Sea Region, materialising in the core of a multidimensional innovation ecosystem for health and life sciences. The so-called “ScanBalt Health Region” considers the challenges and opportunities that this region faces for such ecosystem, arguing that there is plenty of potential for this to be a great example of successful integration in the economy, innovation, and health of the region.

Addressing the dimension of education, Balland et al. (2019) address the impact that relatedness has on technological diversification within the EU. In a logic of economies of agglomeration, they conclude that local knowledge and capabilities patterns of a region are essential to easier adaptations of specific technologies to be applied in the economy.

During the last years, we have been witnessing an increasing degree of modernisation of the economic structure. A study from Beynon-Davies (2010) positively connects this aspect with increments in regional development, as a set of EU policies for implementing eBusiness practices has shown to be a vital long-term measure for small and medium-sized enterprises from Wales.

A research that properly reflects the significance of finding a balance between environmental care and regional development is brought by Refsgaard et al. (2021), in which they demonstrate the growing value of bioeconomy and circular economy for the Nordic regions to reach environmentally and socially sustainable economic growth, specifically in the rural areas, according to the characteristics of their natural resources.

A proper governance capacity and adequacy is key for a region to become more prosperous. This said, Danson and Todeva (2016) show findings mainly from several British regions – and their respective authorities and development agencies – concluding that the *Triple Helices Constelations* interactions and policy implementations (government – industry – university) reach a very effective level when locally/regionally considered.

## 1.7 Answering the “where?” question – the concepts of region and NUTS

A *region* is defined as a contiguous spatial area that is larger than a single urban area and smaller than individual countries, and it is delimited according to a certain purpose and three-type criteria. First, a *homogeneous region* is a geographical area whose constituent elements have characteristics as close as possible to each other and, therefore, the underlying criterion is one of similarity. One example of homogeneous regions can be the municipalities[10]. There are 308 municipalities in Portugal. Then, a *polarized region* (also called *field of forces*) is built around a dominant pole, usually an urban centre, and is composed by complementary spatial units of a lower order which maintain more relations with the respective pole than with the outside or with any other pole of the same hierarchical order. One example of polarized regions can be the metropolitan areas[11]. There are 2 metropolitan areas in Portugal. Finally, a *plan region* (also called *program region*) is a geographical area where its elements are contiguous and are subordinated to the same decision centre. One example of plan regions can be the NUTS III[12] - there are 25 in Portugal - which is going to be the most important space dimension throughout this dissertation, which justifies its importance to be conceptualized.

NUTS is an acronym for *Nomenclature of Territorial Units for Statistics* (in Portuguese, *Nomenclatura das Unidades Territoriais para fins Estatísticos*). This nomenclature was created by Eurostat in the beginning of the 1970s, aiming to harmonize the several countries’ statistics in terms of collecting, compilating and publication of regional statistics. There are 3 levels of regional divisions – NUTS I, NUTS II, and NUTS III –, defined through populational, administrative, and geographical criteria, thus being subject to periodical changes. In Portugal, the last one occurred in 2013 and, even though it only started to be implemented since 2015, we see, quite often, that the new division of NUTS is designated by NUTS 2013. The number of NUTS III went down from 30 to 25 territorial units, since then called by “administrative units”: 23 intermunicipal entities and 2 autonomous regions – Madeira and Azores. There were no major changes to what concerns NUTS II and NUTS I. Hence, in Portugal, 308 municipalities are distributed between 25 NUTS III, 7 NUTS II and 3 NUTS I[13].

[10] Retrieved from CoR - Portugal intro (europa.eu).
[11] Retrieved from CoR - Portugal intro (europa.eu).
[12] Retrieved from O que são NUTS? | Pordata.
[13] For the complete distribution of NUTS 2013, see Annex A.

## 1.8 Research context: what progress has been done up to now?

"There was never a regional policy in this country, even though we have been talking for too long about the regional disequilibria and the necessity to correct them" (Lopes, 1984, p. 331). It is a fact that Portugal is a very asymmetrical country – namely, from the littoral to the interior regions (Silva and Ferreira-Lopes, 2014) –, and this is a worsening phenomenon, raising concerns to slow down this trend, which makes that the literature on Portuguese regional development is vast. Notwithstanding, the most followed methods are descriptive and theoretical, rather than quantitative.

Whereas Costa (2016) sums up the 50-year-evolution of Portuguese regional policy, mainly in an institutional/European perspective – separating the analysis in pre- and post- adherence to the European Economic Community -, several scholars have developed research on sector-specific impacts on regional development: for instance, some of the most popular ones are tourism (Ribeiro and Vareiro, 2007, Soukiazis and Proença, 2008, Novais and Antunes, 2009, Inácio and Cavaco, 2010, Costa et al., 2013, Oliveira and Diniz, 2018, and Trigo and Silva, 2022), and universities (Fernandes, 2010, Lucas et al., 2017, and Ferreira, 2019).

In a modern perspective of regional policy, lights are not only shed to the reduction of regional asymmetries, but also in assessing whether endogenous potential is being adequately used (Costa, 2016). Hence, quantifying and monitoring regional asymmetries throughout the time constitutes a crucial aspect of a good functioning of public policies implementation. In this sense, this monitorisation is done with the *Regional Development Synthetic Index (Índice Sintético de Desenvolvimento Regional* – hereafter, ISDR*)*, developed since 2006 by Statistics Portugal (hereafter, INE) on a yearly basis, having the last publication in 2022 (data related to 2020[14]). This index quantifies territorial cohesion and it is subdivided by three components, each one having a set of indicators: Competitiveness, related with indicators of economic growth and market penetration; Cohesion, related with indicators of equity, life conditions and territorial attractiveness; and Environmental Quality, related with indicators of balance between economic, social, and sustainable development. Like in the purposes of this dissertation, INE (2022b) considers the 25 Portuguese NUTS III, and results are presented at the NUTS III and also at the NUTS II level.

[14] It is important to remember that 2020 was the first year of the COVID-19 pandemic, which brought deep socioeconomic consequences to the world and, consequently, to small and highly shock-exposed economies such as the Portuguese one (Mamede et al., 2020).

Competitiveness was the most asymmetrical category. Only three NUTS III registered higher values, when compared to the national average: those were Área Metropolitana de Lisboa, Região de Aveiro, and Área Metropolitana do Porto. The Cohesion subindex had seven NUTS III with outcomes higher than the national average: those were Região de Coimbra, Cávado, Área Metropolitana de Lisboa, Área Metropolitana do Porto, Região de Aveiro, Região de Leiria, and Médio Tejo, again displaying a trend of comparative advantages of the littoral regions. In terms of the Environmental Quality subindex, results are, somehow, quite opposite to those of the former subindices, as the interior regions and archipelagos are much better ranked than the littoral ones, 17 NUTS III had higher results than the national average, and this is the least asymmetrical category. Highlights go to Região Autónoma da Madeira, Terras de Trás-os-Montes, Beiras e Serra da Estrela, Alto Alentejo, and Região Autónoma dos Açores (ibid).

The ISDR checked, against 2019, a lowering of the disparities related to Environmental Quality and Cohesion, and a rising of the disparities related to the Competitiveness subindex. Five NUTS III were better ranked than the national average in terms of global regional development: Área Metropolitana de Lisboa, Área Metropolitana do Porto, Região de Aveiro, Cávado, and Região de Coimbra, confirming a huge gap between the littoral and the interior regions (ibid).

There are also some econometric studies in the Portuguese academia. The most important for the purposes of this dissertation comes from Silva and Ferreira-Lopes (2014), who present a five-topic index applied to the Portuguese NUTS III, with the three dimensions of HDI plus the two most crucial dimensions that are not included by the UNDP – environment and governance –, a structure that will be, *grosso modo*, methodologically used throughout upcoming sections. Important mentions go to Soukiazis and Proença (2008), Santos and Vieira (2020) - they make two econometric contributions addressing the importance of the touristic sector as a driver for regional development – Brás et al. (2023), by demonstrating the role of entrepreneurial universities in impacting regional competitiveness, and Viegas and Antunes (2013), by analysing convergence patterns and economic clusters of the Portuguese and Spanish NUTS III.

CHAPTER 2

# Methodology

After explaining the methodological aspects of our index and consequent changes vis-à-vis with the one made by Silva and Ferreira-Lopes (2014) (section 2.1), we present our two data mining techniques: first, the panel data models for such assessment, just as the hypothesis for each variable, are presented (section 2.2). Then, we introduce our neural network with the same variables used in the panel data models (section 2.3).

## 2.1 A composite index step by step

The first step that must be taken into the proposed inference is to present a composite index. For such, as mentioned in section 1.6, we use the methodology that is presented by Silva and Ferreira-Lopes (2014), who replicate the one used by the UNDP in the calculations of the HDI, but slightly altering the variables and with the improvement of adding the two other dimensions and respective indicators – environment and governance. Data is mainly provided by INE and, as aforementioned, we make use of data for the 25 NUTS III in Portugal.

For a better comparison between this dissertation and the work done by Silva and Ferreira-Lopes (2014), we stick to almost all the proxies that were used in their paper. This means that only the life expectancy subindex is the same that is used by the UNDP. As for the educational achievement subindex, instead of mean years of schooling and expected years of schooling, respectively, secondary school completion and gross enrollment rate in secondary school are equally weighted and used. In what stands for the income subindex, the authors choose to use the GDPpc, rather than the GNIpc PPP. As for the added variables, the governance subindex remains the same.

Special attention is given to the environment subindex. We share the struggles on selecting a proper indicator, namely, due to a scarce availability of data for NUTS III (ibid.). As they choose the percentage of population served by wastewater treatment stations, one of the specific goals of this dissertation is to find an environmental indicator that could reflect deeper concerns with climate change and a holistic view of environment[15], we thought that other indicator was better fitted for this purpose. Since INE comprises that holistic measure of the state of environment in Portugal that we seek for, we choose the indicator environmental quality from the ISDR (INE, 2022c) to fulfil our needs for an environmental subindex.

Several important steps need to be taken. As our indicators have different units of measurement, standardisation and *minmax* normalisation processes are taken, putting a common scale from 0 to 1 to all values. Table 2.1 shows the composition of each dimension and minimum and maximum values for our samples. These values are based both on important assumptions that the authors do (Silva and Ferreira-Lopes, 2014) and on the samples that we collect from INE databases. Since statistics for our environmental indicator were only available from 2013 to 2020, we calculate the average change rate to obtain values for the 2021 environmental subindex.

The standard formula for assessing each individual dimension (except the income one) for a certain NUTS III, which is as presented in the technical notes of the last HDR (UNDP, 2022) goes as in the Equation 2.1:

$$Dimension\ subindex = \frac{actual\ value - minimum\ value}{maximum\ value - minimum\ value} \tag{2.1}$$

The income dimension shall not be presented as above, as it must reflect the diminishing returns to transforming income into human capabilities, as people do not need excessive financial resources to ensure a decent living (UNDP, 1990). Therefore, the GDPpc shall be logarithmised, thus being presented as follows in the Equation 2.2:

$$Income\ subindex = \frac{Log(actual\ value) - Log(minimum\ value)}{Log(maximum\ value) - Log(minimum\ value)} \tag{2.2}$$

The final score for each NUTS III is provided by the PRDI formula from the Equation 2.3:

$$PRDI = \sqrt[5]{I_{Health} \cdot I_{Education} \cdot I_{Income} \cdot I_{Governance} \cdot I_{Environment}} \tag{2.3}$$

[15] As it is known, the vulnerability of Portugal to the harming effects of climate change is substantial (Carvalho et al., 2014), making it so that it becomes vital to protect biodiversity at all dimensions and at the whole territory.

Table 2.1: dimensions of the PRDI, their respective indicators and subindices, and reference values for the forming of the PRDI.

| **Dimensions** | **Indicators** | **Subindex of each dimension** | **Minimum** | **Mean**[16] | **Maximum**[17] | **Source** |
|---|---|---|---|---|---|---|
| Health | Life expectancy at birth (years) | Life expectancy subindex | 20 | 80 | 82 | Silva and Ferreira-Lopes (2014), INE (2022) |
| Education | Secondary school completion (%) | Educational achievement subindex | 0 | 86,3 | 95,6 | Silva and Ferreira-Lopes (2014), INE (2022) |
| | Gross enrollment rate in secondary education (%) | | 0 | 118,4 | 153,8 | Silva and Ferreira-Lopes (2014), INE (2022) |
| Income | GDPpc (€) | Income subindex | 10,291[18] | 16,229 | 27,126 | INE (2022) |
| Governance | Participation rate in elections (%) | Governance subindex | 0 | 57 | 67,4 | Silva and Ferreira-Lopes (2014), INE (2022) |
| Environment | Environmental quality (index) | Environmental subindex | 0 | 101,5 | 113,6 | Silva and Ferreira-Lopes (2014), INE (2022) |

Source: own elaboration.

[16] Arithmetic mean values in the sample.
[17] Maximum values in the sample.
[18] Minimum value in the sample.

## 2.2 The model, research design, and secondary data collection

Following a revisitation to the PRDI build up by Silva and Ferreira-Lopes (2014), we herein use the index results for each NUTS III in econometric estimations. These estimations are firstly made through an Ordinary Least Squares (OLS) multivariate regression model, for which the general expression is presented in the Equation 2.4:

$$y_{it} = \beta_0 + \beta_t X_{it} + u_{it} \tag{3.1}$$

with $i$ being the NUTS III ($i$ = 1, 2, …, 25), $t$ being a year ($t$ = 2013, …, 2021), $y_{it}$ being the PRDI score for a given NUTS III $i$ and for a given year $t$, $\alpha_i$ being the intercept for a given NUTS III $i$, $\beta_t$ being the vector of coefficients related to the explanatories for a given year $t$, $X_{it}$ being the vector of explanatory variables for a given NUTS III $i$ and for a given year $t$, and $u_{it}$ being the error term for a given NUTS III $i$ and for a given year $t$.

After describing the general lines of the econometric model, a list of potential drivers of developed regions is displayed on Table 2.2, based on the contributions of the extended literature on this matter. The general categories of potential explanatories are inspired in the framework of the MASST model – Macroeconomic, Sectorial, Social, Territorial -, proposed by Capello (2007).

An important assumption related to the index scores, mainly focusing on the governance subindex, needs to be considered in this model. Since the Portuguese municipal elections occur from four to four years - in the considered period, these have occurred in 2013, 2017, and 2021 -, some years do not have complete, five-dimension scores. Faced with this issue, the option we choose is to drag the governance results from each election year to the three following ones, making it so that the 2013 governance scores have the same value than in 2014, 2015, and 2016; the 2017 ones have the same value than in 2018, 2019, and 2020; and the time frame ends with the 2021 elections.

Prior to the model estimations, we make a multicollinearity analysis through correlation matrices with all considered variables (see Annex C), as its presence can diminish the reliability of the parameters estimates (Alin, 2010), having chosen 0.75 as a high correlation threshold. From a starting point, it is expected that some variables of our selection could be highly correlated, due to their mere nature - e.g., the pairs of "Unemployment" and "Youth unemployment", "Export share" and "Import share". Therefore, we split all variables in five models, according to low levels of correlation.

Table 2.2: categories, their respective indicators, identification of the belonging models, and reference values.

| **Category** | **Indicator** | **Abbreviation** | **Minimum** | **Mean** | **Maximum** | **Source** |
|---|---|---|---|---|---|---|
| Macroeconomic | Registered unemployment per 100 inhabitants with 15 and more years old (%) – **model 1** | *unemp* | 2,2 | 5,18 | 10,7 | INE (2023) |
| Macroeconomic | Young registered unemployment per 100 inhabitants aged between 25 and 34 years old (%) – **model 2** | *youthunemp* | 3 | 8,36 | 16,7 | INE (2023) |
| Macroeconomic | Investment rate (%) of enterprises – **model 1** | *invest* | 6,68 | 21,21 | 65,73 | INE (2023) |
| Macroeconomic | Coverage rate (%) – **model 2** | *coverage* | 42,33 | 147,28 | 1303,19 | INE (2023) |
| Macroeconomic | Proportion of exports of high technology goods (%) – **model 2** | *hitecexp* | 0 | 2,92 | 15,31 | INE (2023) |
| Macroeconomic | Housing credit per inhabitant (€) – **model 1** | *credit* | 4486 | 6501,72 | 13457 | INE (2022) |
| Macroeconomic | Apparent labour productivity (thousands of €) – **model 1** | *productivity* | 20,45 | 32,1 | 51,36 | INE (2023) |
| Macroeconomic | Corporate imported goods rate (%) – **model 2** | *impshare* | 0,1 | 3,67 | 55,7 | Pordata (2023) |
| Macroeconomic | Corporate exported goods rate (%) – **model 5** | *expshare* | 0,1 | 3,83 | 34,2 | Pordata (2023) |
| Sectorial | Proportion of primary sector enterprises (%) – **model 4** | *sector1* | 2,02 | 18,8 | 55 | INE (2023) |

| | | | | | | |
|---|---|---|---|---|---|---|
| Sectorial | Proportion of manufacturing sector enterprises (%) – **model 3** | *sector2* | 7,24 | 13,7 | 24,4 | INE (2023) |
| Sectorial | Proportion of tertiary sector enterprises (%) – **model 3** | *sector3* | 37,3 | 67,5 | 89,5 | INE (2023) |
| Sectorial | Proportion of enterprises with activities of ICT – **model 2** | *ict* | 0,23 | 0,82 | 2,52 | INE (2023) |
| Social | Beneficiaries of social integration income, of social security per 1000 inhabitants in active age (‰) – **model 3** | *socinteg* | 9,52 | 31,8 | 118,88 | INE (2022) |
| Social | Population's density (No./km$^2$) – **model 3** | *density* | 13,5 | 170,75 | 951,7 | INE (2023) |
| Social | Ageing ratio (No.) – **model 4** | *ageing* | 76 | 190,6 | 383,91 | INE (2023) |
| Social | Immigrant share (%) – **model 3** | *immigshare* | 0,39 | 3,46 | 22,84 | Own elaboration |
| Social | Enrollment rate in tertiary education (%) – **model 4** | *enrollrt* | 0 | 29,35 | 109,6 | INE (2022) |
| Social | Crime rate (‰) – **model 1** | *crime* | 19,9 | 28,88 | 51,7 | INE (2023) |
| Social | Doctors per 1000 inhabitants (No.) – **model 4** | *doctors* | 1,2 | 3,76 | 13,5 | INE (2022) |
| Territorial | Region area rate (%) – **model 5** | *area* | 0,9 | 4 | 9,3 | Pordata (2022) |
| Territorial | City area rate (%) – **model 4** | *cityarea* | 0,6 | 4 | 17 | Pordata (2022) |
| Territorial | Proportion of classified areas (%) – **model 4** | *classarea* | 5,7 | 21,62 | 59,6 | INE (2023) |

| | | | | | | |
|---|---|---|---|---|---|---|
| Territorial | Urban waste selectively collected per inhabitant (kg/inhab.) – **model 5** | *waste* | 342 | 474,31 | 926 | INE (2023) |
| Territorial | Consumption of electric energy by inhabitant (kWh/inhab.) – **model 3** | *eleccons* | 2549,9 | 4963,92 | 17008 | INE (2023) |
| Territorial | Broadband Internet accesses per 100 inhabitants (%) – **model 5** | *internet* | 12 | 30,33 | 55,7 | INE (2022) |
| Territorial | ATMs per 10 000 inhabitants (No.) – **model 5** | *atm* | 7,1 | 12,01 | 18,9 | INE (2023) |
| Territorial | Coastline region (1 – Yes, 0 – No) – **model 2** | *coastline* | - | - | - | Own observation |
| Territorial | Metropolitan area (1 – Yes, 0 – No) – **model 1** | *metropol* | - | - | - | Own observation |

Source: own elaboration.

Some phenomena may influence positive or negatively any dimension. That is why we also present the hypotheses surrounding each variable. These hypotheses merely show what type of relationships are likely to happen between the variables and the index, which, in econometric terms, means whether the sign of the coefficient is positive or negative. These hypotheses are displayed in Table 2.3 and, this time, follow the distribution of the five models, rather than the MASST framework.

In order to understand which model and which estimating method we should consider to be more accurate, our methodology consists in providing five estimators and three tests. The analysed estimators are pooled OLS estimator, between estimator, first differences estimator, within/fixed effects estimator, and random effects estimator - and three tests - two Breusch-Pagan Lagrange Multiplier tests, facing random effects to pooled OLS and fixed effects to pooled OLS, and the Hausman test (Hausman, 1978), facing random effects to fixed effects. Rejecting the null hypothesis shows that the most accurate model is the fixed effects one (Gujarati, 2004).

Related to the random effects models, we briefly shed light on the Intraclass Correlation Coefficient (ICC), marked in RStudio scripts by theta: ICC measures the extent of agreement for numerical or quantitative variables (Bujang and Barahum, 2017). Results are categorised inside the following intervals: 0 to 0,5; 0,5 to 0,75; 0,75 to 0,9; and 0,9 to 1, which represent, respectively, poor, moderate, good, and excellent reliability (Koo and Li, 2016).

Table 2.3: variables and respective hypotheses per econometric models.

| **Models** | **Hypotheses** | **Expected sign of the coefficient** |
|---|---|---|
| Model 1 | H1a – *unemp* has a negative impact on PRDI | Negative |
| Model 1 | H1b – *invest* has a positive impact on PRDI | Positive |
| Model 1 | H1c – *credit* has a positive impact on PRDI | Positive |
| Model 1 | H1d – *productivity* has a positive impact on PRDI | Positive |
| Model 1 | H1e – *crime* has a negative impact on PRDI | Negative |
| Model 1 | H1f – *metropol* has a positive impact on PRDI | Positive |
| Model 2 | H2a – *youthunemp* has a negative impact on PRDI | Negative |
| Model 2 | H2b – *impshare* has a negative impact on PRDI | Negative |
| Model 2 | H2c – *hitecexp* has a positive impact on PRDI | Positive |
| Model 2 | H2d – *ict* has a positive impact on PRDI | Positive |
| Model 2 | H2e – *coastline* has a positive impact on PRDI | Positive |
| Model 2 | H2f – *coverage* has a positive impact on PRDI | Positive |
| Model 3 | H3a – *sector 2* has a positive impact on PRDI | Positive |
| Model 3 | H3b – *sector 3* has a positive impact on PRDI | Positive |
| Model 3 | H3c – *socinteg* has a negative impact on PRDI | Negative |
| Model 3 | H3d – *density* has a positive impact on PRDI | Positive |
| Model 3 | H3e – *immigshare* has a positive impact on PRDI | Positive |
| Model 3 | H3f – *eleccons* has a positive impact on PRDI | Positive |
| Model 4 | H4a – *sector1* has a negative impact on PRDI | Negative |
| Model 4 | H4b – *ageing* has a negative impact on PRDI | Negative |
| Model 4 | H4c – *enrollrt* has a positive impact on PRDI | Positive |
| Model 4 | H4d – *doctors* has a positive impact on PRDI | Positive |
| Model 4 | H4e – *cityarea* has a positive impact on PRDI | Positive |
| Model 4 | H4f – *classarea* has a negative impact on PRDI | Negative |
| Model 5 | H5a – *expshare* has a positive impact on PRDI | Positive |
| Model 5 | H5b – *area* has a negative impact on PRDI | Negative |
| Model 5 | H5c – *waste* has a negative impact on PRDI | Negative |
| Model 5 | H5d – *internet* has a positive impact on PRDI | Positive |
| Model 5 | H5e – *atm* has a positive impact on PRDI | Positive |

Source: own elaboration.

## 2.3 A second estimation method: neural network

The goal of the neural network estimations is utterly similar to the one of the panel data ones: to provide information regarding the most influential factors behind regional development in Portugal. This second estimation is rather useful to compare both methods in their most important features, similarities, and differences. The methodological implementation was, to a great extent, inspired by the works of Brochado et al. (2019), who use a multilayer perceptron – "a type of artificial neural network that attempts to mimic the human brain by building a network of neurons" (ibid., p. 179). Multilayer perceptrons are the most commonly used neural networks, and, in its genesis, simple to understand: information commonly flows in an input-output way - something known as "feed-forward" – and this network can have multiple hidden layers (Popescu et al., 2009).

CHAPTER 3

# Results and discussion

After presenting each constituent of the methodology in the previous chapter, this one presents the main results and respective discussion. First, we address the results from the PRDI with and without the environmental and governance variables, ending with a discussion on the evolution of regional asymmetries over the years (section 3.1).

Since one of the general goals of this dissertation is to understand the drivers of regional development in Portugal, we present the results of our two chosen data mining techniques: first, the panel data results for such assessment are presented, given a significance level of 10% (section 3.2). Then, we show the neural network results with the same variables used in the panel data models, here distributing them according to their normalised importance, in a way of providing more robustness to the discussion and confirming which are the most important variables to understand regional development in Portugal (section 3.3).

## 3.1 Regional asymmetries in Portugal – results and discussion

We herein analyse the results that our indices provide. First, we assess the evolution regional asymmetries in Portugal based on the three-dimension index, represented in Figure 3.1 and Table 3.1. Then, we consider environment and governance dimensions, and assess what are the major differences with the five-dimension index, represented in Figure 3.2 and Table 3.2. For both cases, the 2013, 2017, and 2021 data are considered.

Figure 3.1 and Table 3.1 show that the first and obvious conclusion is that all Portuguese regions have become more developed throughout the time, since all indices have got higher in all NUTS III (except a very small decrease in Alentejo Litoral from 2017 to 2021), some of which got outstanding improvements. Generally, the best performing NUTS III remain in the coastline (see Annex A), a statement that goes in line with the work of Silva and Ferreira-Lopes (2014) and with the general knowledge of regional asymmetries in Portugal. Área Metropolitana de Lisboa and Alentejo Litoral were the two best ranked NUTS III, with scores higher than 0,9. In 2021, there were 10 NUTS III ranked between 0,8 and 0,9, other 10 NUTS III ranked between 0,7 and 0,8, and Alto Tâmega and Tâmega e Sousa were the two worst ranked NUTS III. Região de Aveiro, Cávado, Ave, and Terras de Trás-os-Montes registered substantial increases in their position rankings, whereas the substantial decreases were registered in Oeste, Região Autónoma da Madeira, Região Autónoma dos Açores, and Beira Baixa.

Figure 3.2 and Table 3.2 show that the first conclusion of the former paragraph is also verified in this case (except a very small decrease in Algarve from 2017 to 2021). However, when environment and governance are added, many interior regions are relatively better ranked when compared to other littoral ones and when compared with Figure 1, something that is also mentioned by Silva and Ferreira-Lopes (2014). In this case, all NUTS III are assessed under 0,9 and 20 of them are ranked above 0,8, Alentejo Litoral and Área Metropolitana de Lisboa switch positions at the top, and Tâmega e Sousa and Alto Tâmega switch positions at the end. An important highlight goes to the remarkable evolution of the Beiras e Serra da Estrela results, which shows that governance and environment were particularly important for this region to shorten the gap throughout the considered period. The NUTS III that climbed the highest in the ranking were Cávado, Ave, Terras de Trás-os-Montes, Alto Alentejo, and Viseu Dão-Lafões, while the biggest downfalls were registered by Algarve, Médio Tejo, and Região Autónoma da Madeira.

The insertion of the governance and environment dimensions was responsible for more similar results amongst all regions, since the distance between the best and the worst ranked NUTS III got much smaller, and many of them got very close results.

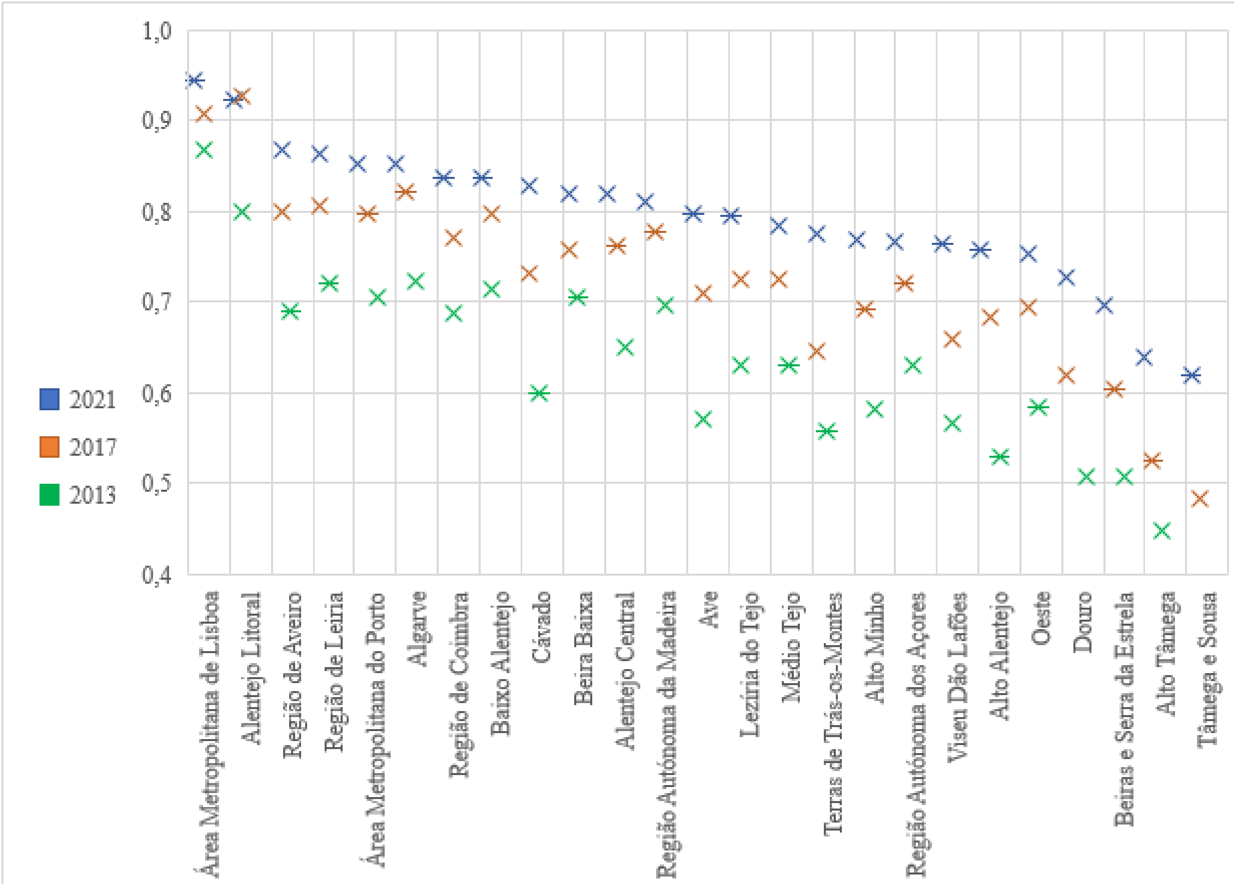


Figure 3.1: PRDI per NUTS III without the environmental and governance dimensions – 2013, 2017, and 2021. Source: own elaboration.

Table 3.1: PRDI per NUTS III without environment and governance ranking differentials from 2013 to 2021.

| **NUTS III** | **PRDI without environment and governance** | | | | | |
|---|---|---|---|---|---|---|
| | **2021** | **2017** | **2013** | **Rank 2021** | **Rank 2013** | **Ranking differentials** |
| Área Metropolitana de Lisboa | 0,945 | 0,908 | 0,867 | 1 | 1 | 0 |
| Alentejo Litoral | 0,922 | 0,926 | 0,800 | 2 | 2 | 0 |
| Região de Aveiro | 0,868 | 0,800 | 0,689 | 3 | 9 | 6 |
| Região de Leiria | 0,862 | 0,805 | 0,721 | 4 | 4 | 0 |
| Área Metropolitana do Porto | 0,852 | 0,798 | 0,704 | 5 | 7 | 2 |
| Algarve | 0,852 | 0,822 | 0,722 | 6 | 3 | -3 |
| Região de Coimbra | 0,837 | 0,770 | 0,688 | 7 | 10 | 3 |
| Baixo Alentejo | 0,837 | 0,797 | 0,714 | 8 | 5 | -3 |
| Cávado | 0,828 | 0,732 | 0,600 | 9 | 15 | 6 |
| Beira Baixa | 0,819 | 0,757 | 0,705 | 10 | 6 | -4 |
| Alentejo Central | 0,818 | 0,763 | 0,649 | 11 | 11 | 0 |
| Região Autónoma da Madeira | 0,810 | 0,777 | 0,696 | 12 | 8 | -4 |
| Ave | 0,796 | 0,709 | 0,572 | 13 | 18 | 5 |
| Lezíria do Tejo | 0,796 | 0,725 | 0,631 | 14 | 12 | -2 |
| Médio Tejo | 0,784 | 0,725 | 0,631 | 15 | 12 | -3 |
| Terras de Trás-os-Montes | 0,775 | 0,645 | 0,558 | 16 | 20 | 4 |
| Alto Minho | 0,768 | 0,692 | 0,583 | 17 | 17 | 0 |
| Região Autónoma dos Açores | 0,767 | 0,720 | 0,631 | 18 | 14 | -4 |
| Viseu Dão-Lafões | 0,763 | 0,658 | 0,566 | 19 | 19 | 0 |
| Alto Alentejo | 0,757 | 0,684 | 0,530 | 20 | 21 | 1 |
| Oeste | 0,754 | 0,693 | 0,584 | 21 | 16 | -5 |
| Douro | 0,728 | 0,620 | 0,507 | 22 | 22 | 0 |
| Beiras e Serra da Estrela | 0,697 | 0,603 | 0,507 | 23 | 22 | -1 |
| Alto Tâmega | 0,638 | 0,525 | 0,449 | 24 | 24 | 0 |
| Tâmega e Sousa | 0,620 | 0,482 | 0[19] | 25 | 25 | 0 |

Source: own elaboration.

[19] This is related to the minimum value in the sample of GDPpc.

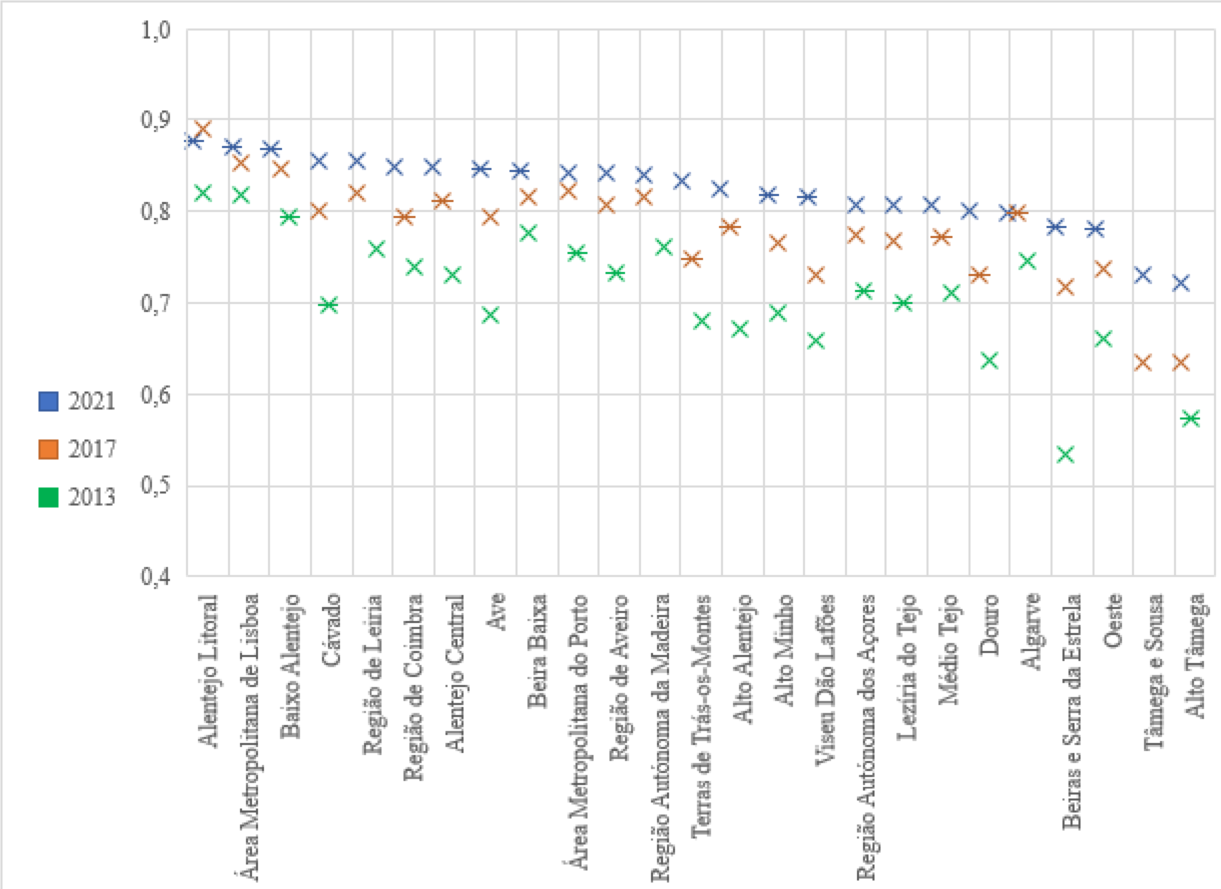


Figure 3.2: PRDI per NUTS III with the environmental and governance dimensions – 2013, 2017, and 2021. Source: own elaboration.

Table 3.2: PRDI per NUTS III with environment and governance ranking differentials from 2013 to 2021.

| **NUTS III** | **PRDI with environment and governance** | | | | | |
|---|---|---|---|---|---|---|
| | **2021** | **2017** | **2013** | **Rank 2021** | **Rank 2013** | **Ranking differentials** |
| Alentejo Litoral | 0,876 | 0,890 | 0,819 | 1 | 1 | 0 |
| Área Metropolitana de Lisboa | 0,870 | 0,853 | 0,817 | 2 | 2 | 0 |
| Baixo Alentejo | 0,869 | 0,845 | 0,794 | 3 | 3 | 0 |
| Cávado | 0,856 | 0,801 | 0,698 | 4 | 15 | 11 |
| Região de Leiria | 0,855 | 0,821 | 0,759 | 5 | 6 | 1 |
| Região de Coimbra | 0,849 | 0,794 | 0,740 | 6 | 9 | 3 |
| Alentejo Central | 0,848 | 0,811 | 0,730 | 7 | 11 | 4 |
| Ave | 0,846 | 0,794 | 0,687 | 8 | 17 | 9 |
| Beira Baixa | 0,844 | 0,816 | 0,777 | 9 | 4 | -5 |
| Área Metropolitana do Porto | 0,843 | 0,822 | 0,754 | 10 | 7 | -3 |
| Região de Aveiro | 0,841 | 0,807 | 0,734 | 11 | 10 | -1 |
| Região Autónoma da Madeira | 0,841 | 0,816 | 0,761 | 12 | 5 | -7 |
| Terras de Trás-os-Montes | 0,834 | 0,748 | 0,681 | 13 | 18 | 5 |
| Alto Alentejo | 0,824 | 0,784 | 0,672 | 14 | 19 | 5 |
| Alto Minho | 0,818 | 0,766 | 0,688 | 15 | 16 | 1 |
| Viseu Dão-Lafões | 0,817 | 0,731 | 0,659 | 16 | 21 | 5 |
| Região Autónoma dos Açores | 0,807 | 0,775 | 0,712 | 17 | 12 | -5 |
| Lezíria do Tejo | 0,806 | 0,767 | 0,700 | 18 | 14 | -4 |
| Médio Tejo | 0,806 | 0,773 | 0,711 | 19 | 13 | -6 |
| Douro | 0,800 | 0,731 | 0,637 | 20 | 22 | 2 |
| Algarve | 0,798 | 0,799 | 0,745 | 21 | 8 | -13 |
| Beiras e Serra da Estrela | 0,782 | 0,716 | 0,534 | 22 | 24 | 2 |
| Oeste | 0,781 | 0,738 | 0,662 | 23 | 20 | -3 |
| Tâmega e Sousa | 0,731 | 0,634 | 0[20] | 24 | 25 | 1 |
| Alto Tâmega | 0,721 | 0,634 | 0,574 | 25 | 23 | -2 |

Source: own elaboration.

[20] This is related to the minimum value in the sample of GDPpc.

## 3.2 Panel data models. What shapes a developed region?

We herein show the first part of our econometric findings, with the support of Table 3.3, a regression table that displays the most accurate statistics and estimation methods for each model, according to the Hausman test results.

In model 1, there is global significance for all models. Both random effects and fixed effects pass the tests against pooled OLS, and the Hausman test shows a p-value over 10%, which makes us choose the random effects estimator. The ICC equals 0.7062, which represents moderate reliability. On the one hand, the variables *unemp* and *crime* negatively impact on the region's index. On the one hand, the variables *credit* and *productivity* are linked to higher regional scores. *productivity* is a particularly important variable to understand the current demographic pattern in Portugal. The search for better economic conditions – hence, for more productive jobs – is one of the key factors behind a gigantic exodus from rural and interior regions to the coastline ones, making it so that the latter is occupied by more than 80% of the Portuguese population (Almeida, 2020). As coastline cities get more populated – especially the metropolitan areas -, housing pressures arise, which translates in increasing housing credit demand (Branco and Alves, 2018).

In model 2, we have global significance for the pooled OLS, first difference, fixed effects, and random effects models. Both random effects and fixed effects pass the tests against pooled OLS, and the Hausman test shows a p-value under 10%. This makes us choose the fixed effects estimator. The variable *youthunemp* negatively impacts on the regional development index, but the latter is positively influenced by *ict*. The dummy variable *coastline* does not appear in the estimations. Ruivo et al. (2015) evidence that the Portuguese economy is structurally changing towards a medium-high technology economy, which is proven by increasing exports in ICT outsourcing services over the last years, something that can be much helpful to reach lower youth unemployment rates – these were dramatically high during the economic crisis (Marques and Videira, 2021) -, and to mitigate the exodus of better qualified young people to leave the country (Caldas, 2012).

Model 3 has global significance for the pooled OLS, between, fixed effects, and random effects models. Both random effects and fixed effects pass the tests against pooled OLS, and the Hausman test shows a p-value under 10%. This, again, makes us choose the fixed effects estimator. All variables are statistically significant: *sector3*, *immigshare*, and *eleccons* have a positive impact on the region's index, whereas *socinteg*, *density*, and *sector2* negatively impact on the score.

About model 4, the pooled OLS, the fixed effects, and the random effects models are globally significant. Both random effects and fixed effects pass the tests against pooled OLS, and the Hausman test shows a p-value under 10%, which makes us go for the fixed effects estimator, with consistent estimates. The statistically significant variables in this case are *ageing* and *doctors*, both positively impacting on the score of each region. Just as in model 2 with *coastline*, *cityarea* does not appear in the estimations.

In model 5, we have global significance for the pooled OLS, between, fixed effects, and random effects models. Both random effects and fixed effects pass the tests against pooled OLS, and the Hausman test shows a p-value over 10%. This makes us go for the random effects estimator, which is consistent and most efficient if the appropriate model is the random effects one. The ICC is 0.7102, which, just as in model 1, represents moderate reliability. For this case, the only statistically significant variable is *internet*, which positively impacts on the index. The implementation of ICT is surely necessary to reach more economic efficiency, but they are not sufficient by themselves, as proved by Tranos (2012), who presents a study on the causal relationship between internet infrastructures and regional development on European cities.

In general, we find our results in line with what is expected from the literature and even from common sense. When comparing our results to our expectations shown in Table 2.3, the rationale behind a positive relationship between, for instance, access to credit, higher productivity levels with more developed regions is understandable. Moreover, the importance of demographic diversity in economic growth goes along with the literature – see, for instance, Ottaviano and Peri (2005) and Florida et al. (2008).

On the other side of the coin, the negative relationship between development scores and higher criminality, (youth) unemployment levels, or population also makes sense. It is expected as well that a region where a bigger proportion of the active population depends on social benefits, rather than from labour income, has a small score. This can happen due to a relatively weaker economic structure that is not adequately capable to insert all active population. In fact, poverty and social exclusion is a very important topic in Portugal, as around 20% of the population is in risk of poverty, even after receiving social benefits (Pordata, 2022).

We find the results for *ageing* peculiar and, to some extent, contradicting reality, as the most economically prosperous Portuguese regions tend to be the littoral ones and, at the same time, the ones with relatively younger population (Nunes, 2017).

Table 3.3: regression table[21].

| Dependent variable: PRDI_score | | | | | |
|---|---|---|---|---|---|
| | (1) | (2) | (3) | (4) | (5) |
| X1unemp | -0.026*** | | | | |
| | (0.003) | | | | |
| X1invest | 0.0004 | | | | |
| | (0.001) | | | | |
| X1credit | 0.00001** | | | | |
| | (0.00001) | | | | |
| X1productivity | 0.005*** | | | | |
| | (0.001) | | | | |
| X1crime | -0.002* | | | | |
| | (0.001) | | | | |
| X1metropol | 0.008 | | | | |
| | (0.041) | | | | |
| X2youthunemp | | -0.017*** | | | |
| | | (0.001) | | | |
| X2impshare | | 0.003 | | | |
| | | (0.005) | | | |
| X2hitecexp | | 0.001 | | | |
| | | (0.002) | | | |
| X2ict | | 0.102*** | | | |
| | | (0.034) | | | |
| X2coverage | | 0.00001 | | | |
| | | (0.00003) | | | |
| X3sector3 | | | 1.041*** | | |
| | | | (0.354) | | |
| X3socinteg | | | -0.004*** | | |
| | | | (0.001) | | |
| X3density | | | -0.005*** | | |
| | | | (0.002) | | |
| X3immigshare | | | 0.008** | | |
| | | | (0.003) | | |
| X3eleccons | | | 0.00003** | | |
| | | | (0.00001) | | |
| X3sector2 | | | -2.193** | | |
| | | | (0.947) | | |
| X4ageing | | | | 0.002*** | |
| | | | | (0.0003) | |
| X4enrollrt | | | | -0.0003 | |
| | | | | (0.001) | |
| X4doctors | | | | 0.039*** | |
| | | | | (0.013) | |
| X4classarea | | | | -0.007 | |
| | | | | (0.015) | |
| X4sector1 | | | | 0.353 | |
| | | | | (0.238) | |
| X5expshare | | | | | 0.001 |
| | | | | | (0.002) |
| X5area | | | | | 0.007 |
| | | | | | (0.006) |
| X5waste | | | | | -0.0001 |
| | | | | | (0.0001) |
| X5internet | | | | | 0.008*** |
| | | | | | (0.001) |
| X5atm | | | | | 0.002 |
| | | | | | (0.004) |
| Constant | 0.693*** | | | | 0.528*** |
| | (0.063) | | | | (0.054) |
| Observations | 225 | 225 | 225 | 225 | 225 |
| R2 | 0.607 | 0.563 | 0.411 | 0.550 | 0.542 |
| Adjusted R2 | 0.596 | 0.498 | 0.320 | 0.483 | 0.532 |
| F Statistic | 336.813*** | 50.322*** (df = 5; 195) | 22.582*** (df = 6; 194) | 47.620*** (df = 5; 195) | 259.425*** |
| Note: | | | | | *p<0.1; **p<0.05; ***p<0.01 |

Source: own elaboration.

[21] Made with R, version 4.2.1, and RStudio, package *stargazer* (Hlavac, 2022).

## 3.3 Neural network: a second perspective

One contrast against panel data estimations – among many others - is the inclusion of the geographical and time dimensions, represented by *region_cod* and *year*, respectively, with the goal of analysing the variability between regions and throughout the time. In the panel data models, these only serve to filter data, rather than being explanatories.

We choose to address the 10 most and 10 least influential variables in this neural network, and full results are presented in Table 3.4. According to the importance and normalised importance of each independent variable. the most impactful explanatories in regional development in Portuguese NUTS III are *waste* (100%), *unemp* (92,8%), *productivity* (87,8%), *impshare* (74,1%), *credit* (64,9%), and *cityarea* (63,5%), *region_cod* (57,5%), *year* (57%), *eleccons* (56,3%), and *enrollrt* (51,8%). Oppositely, the least impactful ones are *hitecexp* (9,2%), *metropol* (13%), *invest* (20,3%), *atm* (20,7%), *doctors* (21,6%), *coastline* (22,8%), *youthunemp* (23,5%), *internet* (26,5%), *crime* (26,8%), and *immigshare* (27,5%).

This neural network confirms the relevance of several explanatories, as they have appeared as statistically significant in the panel data estimations – *unemp*, *productivity*, *impshare*, and *credit.* Both estimations are also in harmony in considering *metropol*, *invest*, *atm*, *coastline*, and *hitecexp* as lowly important for our panorama. It is also immensely relevant to note that both estimations are divided in matters of several variables: by way of illustration, *youthunemp*, *internet*, *crime*, and *immigshare* were statistically significant in the first estimations, but are not fundamental in the second ones; in a distinct sense, variables such as *waste*, *cityarea*, and *enrollrt* are highlighted in this multilayer perceptron.

Table 3.4: independent variables, their respective importance, and normalised importance for regional development in Portugal.

| Variables | Importance | Normalised importance |
|---|---|---|
| region | ,044 | 57,5% |
| coastline | ,017 | 22,8% |
| metropol | ,010 | 13,0% |
| year | ,043 | 57,0% |
| unemp | ,071 | 92,8% |

| | | |
|---|---|---|
| youthunemp | ,018 | 23,5% |
| invest | ,015 | 20,3% |
| coverage | ,022 | 29,5% |
| hitecexp | ,007 | 9,2% |
| credit | ,049 | 64,9% |
| productivity | ,067 | 87,8% |
| impshare | ,056 | 74,1% |
| expshare | ,032 | 41,8% |
| sector1 | ,032 | 42,5% |
| sector2 | ,030 | 39,3% |
| sector3 | ,029 | 38,5% |
| ict | ,030 | 39,4% |
| socinteg | ,022 | 28,7% |
| density | ,031 | 40,1% |
| ageing | ,028 | 36,9% |
| immigshare | ,021 | 27,5% |
| enrollrt | ,039 | 51,8% |
| crime | ,020 | 26,8% |
| doctors | ,016 | 21,6% |
| area | ,024 | 31,4% |
| cityarea | ,048 | 63,5% |
| classarea | ,021 | 27,8% |
| waste | ,076 | 100,0% |
| eleccons | ,043 | 56,3% |
| internet | ,020 | 26,5% |
| atm | ,016 | 20,7% |

Source: own elaboration[22].

[22] Made with SPSS.

CHAPTER 4

# Conclusions

Ultimately, this dissertation aims to raise awareness on the importance of human, sustainable development in Portuguese regions. A place-based policymaking that considers endogenous characteristics as a potential for prosperity can have larger successful implementation, thus augmenting regional resilience (Stanickova and Melecký, 2018). Our main objectives were to shed lights on the most important concepts around regional development, to display what is the current Portuguese reality in terms of regional asymmetries at the NUTS III level, and to determine the factors behind such pattern.

As regards the first research question – "what is the current shape of Portuguese asymmetries, taking into account the five dimensions of the Portuguese Regional Development Index and, particularly, what is the impact of environment and governance dimensions?" - first, we updated the PRDI initially calculated by Silva and Ferreira-Lopes (2014), which contains a total of 6 indicators grouped in 5 dimensions - income, health, education, environment, and governance. As the three "UN" dimensions put the more coastal regions in a much higher place, this index showed that considering environment and governance was helpful to put the chronically more depressed regions closer to the more developed ones. Normally, in Portugal, this distribution goes from the interior to the coastal regions. This conception thus diminishes regional asymmetries and, although shows that Portugal is still a very asymmetric country, it is not as much as it is normally considered. This index also showed in which dimension(s) each region can focus, in order to become more prosperous, healthy, and environmentally safe for its inhabitants.

Subsequently and considering the second research question – "what are the drivers behind such distribution?", these regional development scores for the Portuguese NUTS III were used as dependent variable for econometric estimations. In a first instance, 5 panel data models that encompassed 29 independent variables, and, in a second instance, a neural network with the same variables highlighted several potential explanatories that helped us to understand such abovementioned scores. These variables were distributed in two logics: the MASST framework (Capello, 2007) and reduced correlation. There are many distinct results between them, but we highlight that both sets of estimations have coincidence in confirming the relevance of several explanatories, such as unemployment per 100 inhabitants with 15 and ore years old, apparent labour productivity, housing credit per inhabitant, and percentage of corporate imported goods. These estimations have coincidence in confirming lower relevance of other variables, such as a region being a metropolitan area or not, a region being at the coastline, the investment rate of enterprises, proportion of exports of high technology goods, and the number of ATMs per 10000 inhabitants.

The results of this study are expected to have theoretical contributions. When compared to the original index created by Silva and Ferreira-Lopes (2014), our study tried to keep the core characteristics of the latter – keeping the five main dimensions and calculating the index as much as possible as following the methodology of the UN –, but also tried to innovate where there was relatively more margin for: to give examples, considering a very complex environmental subindex and substituting the percentage of population served by wastewater treatment stations, including more years of analysis, and actually including an econometric analysis that can help to understand why regions have such an asymmetries pattern.

This study is expected to offer managerial implications for policymakers. Considering the abovementioned significant variables and considering the role that European structural funds will play, an efficient and structural regional policy for shortening asymmetries in the future will need to prioritize a balance between socioeconomic resilience with climate action and energetic transition: on the one hand, focused on the rise of productivity levels and export capacity, the foster of good quality, non-precarious employment, and a wider access to housing credit; on the other hand, giving incentives to industrial decarbonization, renewable energies, and bioeconomy, thus lowering carbon footprint (*Ministério do Planeamento*, 2021). Preferably, these implementations will regard of preserving endogenous potential that exists in all regions.

As in every piece of academic work, this one is not free of having its limitations, mainly related to secondary data collection. The governance subindex relies upon the “dragging” of the participation rate from the election years to the following ones, which can bring the necessity of considering other governance variable (e.g., the dominant political party orientation in the municipalities of each region). The environmental subindex, although seen as a differentiating aspect in this dissertation, might also deserve another look, as an alternative indicator that shall not be as complex as the presented in this work could possibly be arranged, for which it can be worthy to think about a simple environmental indicator that represents high differences amongst regions, thus having decent explanatory power[23].

Another limitation that we point out is related to the removal of variables from the econometric analysis, due to unavailability of data, either for the 2013-2021 period, either for the 25 NUTS III. For instance, we were not able to find proper indicators on water losses, tolerance, technological literacy, use of clean/renewable energy sources, proportion of public housing, or corporatisation.

A final limitation about the econometric estimations can be presented as well. Choosing 0,75 as a threshold for the correlation coefficient can be too high, for which we identify the need of grouping variables that share correlations lower than 0,50, in order to obtain even more adequate estimations which can bring more or different statistically significant independent variables.

As the pertinence of these subjects is permanently growing, there is a wide panoply of possibilities for future research. Our index and our estimations can be replied in different time dimensions – for instance, in a very long term – and in different regions, municipalities, countries, or continents, which could open discussions to very rich comparisons in terms of drivers of regional development. Another hypothesis is testing the "neighbourhood effect", associated with gravity models, and which principal premise is that the level of development of a regional is correlated with the level of neighbouring regions (Nijkamp and Ratajczak, 2021).

[23] It is also important to mention that finding proper indicators for these two dimensions was already a struggle for the past literature.

# References


Alin, A. (2010). Multicollinearity. *Wiley Interdisciplinary Reviews: Computational Statistics*, 2(3), 370-374. https://doi.org/10.1002/wics.84

Alkire, S., and S. Jahan. (2018). "The New Global MPI 2018: Aligning with the Sustainable Development Goals." Human Development Research Paper. UNDP-HDRO, New York. http://hdr.undp.org/en/content/new-global-mpi-2018-aligning -sustainable-development-goals

Almeida, , M. A. (2020). The use of rural areas in Portugal: Historical perspective and the new trends. *Revista Galega De Economía*, *29*(2), 1-17. https://doi.org/10.15304/rge.29.2.6750

APA. (2023). *Relatório do Estado do Ambiente 2022*, Lisboa. Retrieved from: https://rea.apambiente.pt/sites/default/files/rea/REA%202022_pdf_vfinal_21_07_2023.pdf

Balland, P. A., Boschma, R., Crespo, J., & Rigby, D. L. (2019). Smart specialization policy in the European Union: relatedness, knowledge complexity and regional diversification. *Regional studies*, *53*(9), 1252-1268. https://doi.org/10.1080/00343404.2018.1437900

Barbour, G. B. (1937). The Tennessee Valley Project. *The Geographical Journal*, *89*(5), 393-405. https://www.jstor.org/stable/1786409

Belussi, F., & Caldari, K. (2009). At the origin of the industrial district: Alfred Marshall and the Cambridge school. *Cambridge Journal of economics*, *33*(2), 335-355. https://doi.org/10.1093/cje/ben041

Beth, M. P. (1993). Regional development: more than ever, the business of America. *Urban geography*, *14*(2), 177-193. https://doi.org/10.2747/0272-3638.14.2.177

Beynon-Davies, P. (2010), "eBusiness as a driver for regional development", *Journal of Systems and Information Technology*, 12(1) 17-36. https://doi.org/10.1108/13287261011032634

Bilbao-Ubillos, J. (2013). Another approach to measuring human development: The composite dynamic Human Development Index. *Social Indicators Research*, *111*(2), 473-484. https://doi.org/10.1007/s11205-012-0015-y

Blank, W., Frank, P., & Karopka, T. (2013). Health and life sciences as drivers for regional development and prosperity in the Baltic Sea region. *Journal of East-West Business*, *19*(1-2), 122-137. https://doi.org/10.1080/10669868.2013.780531

Bolton, R. (1992). "Place Prosperity vs People Prosperity" Revisited: An Old Issue with a New Angle. Urban Studies, 29(2), 185–203. https://doi.org/10.1080/00420989220080261

Branco, R., & Alves, S. (2020). Urban rehabilitation, governance, and housing affordability: Lessons from Portugal. *Urban Research & Practice*, *13*(2), 157-179. https://doi.org/10.1080/17535069.2018.1510540

Brás, G. R., Torres Preto, M., Daniel, A. D., & Teixeira, A. A. C. (2023). Assessing the Impact of Universities' Entrepreneurial Activity on Regional Competitiveness. *Administrative Sciences*, *13*(2), 34. MDPI AG. Retrieved from http://dx.doi.org/10.3390/admsci13020034

Bravo, G. (2014). The Human Sustainable Development Index: New calculations and a first critical analysis. *Ecological indicators*, *37*, 145-150. https://doi.org/10.1016/j.ecolind.2013.10.020

Brinkman, R. (1995). Economic Growth versus Economic Development: Toward a Conceptual Clarification. Journal of Economic Issues (Association for Evolutionary Economics), 29(4 OP-Journal of Economic Issues (Association for Evolutionary Economics). Dec95, Vol. 29 Issue 4, p1171-1188. 18p. 1 Diagram, 2 Graphs.), 1171. Retrieved from http://search.ebscohost.com/login.aspx?direct=true&site=eds-live&db=heh&AN=9601024293

Brochado, A., Rita, P., & Moro, S. (2019). Discovering patterns in online reviews of Beijing and Lisbon hostels. *Journal of China Tourism Research*, *15*(2), 172-191. https://doi.org/10.1080/19388160.2018.1543065

Bujang, M. A., & Baharum, N. (2017). A simplified guide to determination of sample size requirements for estimating the value of intraclass correlation coefficient: a review. *Archives of Orofacial Science*, *12*(1). api.semanticscholar.org/CorpusID:127106131

Caldas, J. C. (2012). *The consequences of austerity policies in Portugal* (pp. 1-5). Friedrich-Ebert-Stiftung. library.fes.de/pdf-files/id-moe/09311.pdf

Capello, R. (2007). A forecasting territorial model of regional growth: the MASST model. *The Annals of Regional Science*, *41*(4), 753. https://doi.org/10.1007/s00168-007-0146-2

Carvalho, A., Schmidt, L., Santos, F. D., & Delicado, A. (2014). Climate change research and policy in Portugal. *Wiley Interdisciplinary Reviews: Climate Change*, *5*(2), 199-217. https://doi.org/10.1002/wcc.258

Chakravarty, S. R. (2003). A generalized human development index. *Review of development economics*, *7*(1), 99-114. https://doi.org/10.1111/1467-9361.00178

Cornia, G. A. (2003). Growth Patterns, International Finance and human Development: the experience of the last two decades. *Cadernos de Estudos Africanos*, (4), 11-33. https://doi.org/10.4000/cea.1575

Costa, J. (2016). Cinco décadas de políticas de desenvolvimento regional em Portugal. *Boletim Regional, Urbano e Ambiental*, 14, 73–81. IPEA. https://repositorio.ipea.gov.br/bitstream/11058/6842/1/BRU_n14_Cinco.pdf

Costa, J., & Nijkamp, P. (2009). Compêndio de economia regional. volume I, teoria, temáticas e políticas. *Parede, Portugal: Princípia Editora*.

Costa, V., Monte, A. P., & Fernandes, P. O. (2013). Políticas de desenvolvimento regional para o setor do turismo na Região Norte de Portugal. In *III Congresso Internacional de Turismo da ESG/IPCA* (pp. 197-222). ESG-IPCA. http://hdl.handle.net/10198/10706

Danson, M., & Todeva, E. (2016). Government and Governance of Regional Triple Helix Interactions. *Industry and Higher Education*, 30(1), 13–26. https://doi.org/10.5367/ihe.2016.0293

Desai, M. (1991). Human development: concepts and measurement. *European Economic Review*, *35*(2-3), 350-357. https://doi.org/10.1016/0014-2921(91)90136-7

Desrochers, P. (1998). A geographical perspective on Austrian economics. *Quarterly Journal of Austrian Economics*, *1*(2), 63-83. https://doi.org/10.1007/s12113-998-1009-0

EIB. (2021). *Cohesion and Regional Development Overview 2021*. Luxembourg: European Investment Bank publications office. https://www.eib.org/attachments/thematic/cohesion_and_regional_development_overview_2021_en.pdf

Feldman, M., Hadjimichael, T., Lanahan, L., & Kemeny, T. (2016). The logic of economic development: A definition and model for investment. *Environment and Planning C: Government and Policy*, *34*(1), 5-21. https://doi.org/10.1177/0263774X15614653

Fernandes, J. M. S. R. (2010). *O impacto económico das instituições do ensino superior no desenvolvimento regional: o caso do Instituto Politécnico de Bragança* (Doctoral dissertation, Instituto Politécnico de Bragança (Portugal)). http://hdl.handle.net/10198/3681

Ferreira, E. C. C. (2019). *A Importância das Instituições de Ensino Superior no Desenvolvimento Regional em Portugal*. Tese de Doutoramento, Universidade de Évora, Évora, Portugal. http://hdl.handle.net/10174/25457

Figueiredo, A. M., Silva, M. R., & Pessoa, A. (2008). Crescimento económico. Escolar Editora.

Florida, R., Mellander, C., & Stolarick, K. (2008). Inside the black box of regional development - human capital, the creative class and tolerance. *Journal of Economic Geography*, 8(5), 615-649. doi:10.1093/jeg/lbn023

Foster, J. E., Lopez-Calva, L. F., & Szekely, M. (2005). Measuring the distribution of human development: methodology and an application to Mexico. *Journal of Human Development*, *6*(1), 5-25. https://doi.org/10.1080/1464988052000342220

Graça Moura, M. (2018). Joseph Schumpeter. In Conlin, J. (Ed.), *Great Economic Thinkers: An Introduction – From Adam Smith to Amartya Sen* (pp. 110-129). London. Reaktion Books.

Greco, S., Ishizaka, A., Resce, G., & Torrisi, G. (2020). Measuring well-being by a multidimensional spatial model in OECD Better Life Index framework. *Socio-Economic Planning Sciences*, *70*, 100684. https://doi.org/10.1016/j.seps.2019.01.006

Grimm, M., Harttgen, K., Klasen, S., & Misselhorn, M. (2008). A human development index by income groups. *World Development*, *36*(12), 2527-2546. https://doi.org/10.1016/j.worlddev.2007.12.001

*Grover, Arti; Lall, Somik V.; Maloney, William F.. 2022. Place, Productivity, and Prosperity : Revisiting Spatially Targeted Policies for Regional Development. Washington, DC: World Bank. © World Bank.* https://openknowledge.worldbank.org/handle/10986/36843 *License: CC BY 3.0 IGO.*

Gujarati, D. (2004). *Basic Econometrics*, (4th ed.), New York: McGraw-Hill.

Harttgen, K., & Klasen, S. (2011). A human development index by internal migrational status. *Journal of Human Development and Capabilities*, *12*(3), 393-424. hdrp200954pdf.pdf (undp.org)

Harttgen, K., & Klasen, S. (2012). A household-based human development index. *World Development*, *40*(5), 878-899. https://doi.org/10.1016/j.worlddev.2011.09.011

Hausman, J. A. (1978). Specification tests in econometrics. *Econometrica: Journal of the econometric society*, 46(6), 1251-1271. https://doi.org/10.2307/1913827

Herrero, C., Martínez, R., & Villar, A. (2010). *Improving the measurement of human development*. United Nations Development Programme. hdr.undp.org/system/files/documents/hdrp201012pdf.pdf

Hlavac, M. (2022). stargazer: Well-Formatted Regression and Summary Statistics Tables. R package version 5.2.3. https://CRAN.R-project.org/package=stargazer

Hou, J., Walsh, P. P., & Zhang, J. (2015). The dynamics of human development index. *The Social Science Journal*, *52*(3), 331-347. https://doi.org/10.1016/j.soscij.2014.07.003

Inácio, A. I., & Cavaco, C. (2010). Enoturismo em Portugal: forma de desenvolvimento regional e afirmação cultural local. *Revista Turismo & Desenvolvimento*, *2*(13/14), 761-769. https://doi.org/10.34624/rtd.v2i13/14.12357

INE. (2022). *Índice Sintético de Desenvolvimento Regional*, Lisboa, pp. 1- 12.

INE. (2022). Censos 2021. XVI Recenseamento Geral da População. VI Recenseamento Geral da Habitação: Resultados definitivos. Lisboa: INE, 2022. Disponível na www: <url:https://www.ine.pt/xurl/pub/65586079>. ISSN 0872-6493. ISBN 978-989-25-0619-7

Koo, T. K., & Li, M. Y. (2016). A guideline of selecting and reporting intraclass correlation coefficients for reliability research. *Journal of chiropractic medicine*, *15*(2), 155-163. DOI: 10.1016/j.jcm.2016.02.012

Kovacevic, M. (2010). Review of HDI critiques and potential improvements. *Human development research paper*, *33*, 1-44. hdrp201033pdf.pdf (undp.org)

Krugman, P. (1992). *Geography and trade*. MIT press.

Legatum Institute. (2021). *The Legatum Prosperity Index. A tool for transformation*, London. https://docs.prosperity.com/3616/3544/3967/The_2021_Legatum_Prosperity_Index_Overview_-_Web.pdf

Lind, N. (2019). A development of the human development index. *Social Indicators Research*, *146*(3), 409-423. https://doi.org/10.1007/s11205-019-02133-9

Lopes, A. S. (1984). *Desenvolvimento Regional: Problemática, Teoria, Modelos*, 2ª edição, Lisboa: Fundação Calouste Gulbenkian.

Lucas, E., Mangas, N., Marques, J., & Nicolau, A. (2017). *A importância do ensino superior politécnico em Portugal para o desenvolvimento regional – o caso do Politécnico de Leiria*, in: http://www.aforges.org/wpcontent/uploads/2017/03/11ELucas_NMangas_JMarques_ANicolau_A-import_ncia-do-ensino-superior.pdf.

Ma, W., Renwick, A., & Zhou, X. (2020). The relationship between farm debt and dairy productivity and profitability in New Zealand. *Journal of dairy science*, *103*(9), 8251-8256. https://doi.org/10.3168/jds.2019-17506

Mamede, R. P.; Pereira, M.; Simões, A. (2020), *Portugal: uma análise rápida do impacto da COVID-19 na economia e no mercado de trabalho.* Lisboa: OIT. wcms_754606.pdf (ilo.org)

Marques, P., & Videira, P. (2021). Labor market policies to fight youth unemployment in Portugal: between statism and experimentalism. In *The Growing Challenge of Youth Unemployment in Europe and America* (pp. 127-164). Bristol University Press. https://doi.org/10.2307/j.ctv1rnpjgq.11

McCann, P. (2013). *Modern urban and regional economics*. Oxford University Press.

Medeiros, E. (2022). Strategic-based regional development: Towards a theory of everything for regional development?. European Journal of Spatial Development, 19(5), 1-26. http://dx.doi.org/10.5281/zenodo.6805455

Ministério do Planeamento. (2021). *PRR – Recuperar Portugal, Construindo o Futuro*, Lisboa. Retrieved from: recuperarportugal.gov.pt/wp-content/uploads/2021/10/PRR.pdf

Morris, M. D. (1978). A physical quality of life index. *Urban Ecology*, *3*(3), 225-240. https://doi.org/10.1016/0304-4009(78)90015-3

Mota, C., and Pereira, D. (2016). Human development index based on ELECTRE TRI-C multicriteria method: an application in the city of Recife. *Social Indicators Research*, *125*(1), 19-45. http://dx.doi.org/10.1007/s11205-014-0836-y

Nathwani, J. S., Lind, N. C., & Pandey, M. D. (1997). Affordable Safety by Choice. *Institute for Risk Research, University of Waterloo, Canada*. http://irr.uwaterloo.ca/pdf_files/LQI_Summary.pdf

Nijkamp, P., & Abreu, M. A. (2009). *Regional development theory*. Amsterdam, The Netherlands: Vrije Universiteit, Faculty of Economics and Business Administration. 20090029.pdf (vu.nl)

Nijkamp, P., & Ratajczak, W. (2021). Gravitational Analysis in Regional Science and Spatial Economics: A Vector Gradient Approach to Trade. International Regional Science Review, 44(3-4), 400-431. doi:10.1177/0160017620980519

Noorbakhsh, F. (1998). A modified human development index. *World Development*, *26*(3), 517-528. https://doi.org/10.1016/S0305-750X(97)10063-8

Novais, C. B., and Antunes, J. (2009), "*O contributo do enoturismo para o desenvolvimento regional: o caso das rotas dos vinhos*", in Piaget University (Ed.), Cabo Verde: Redes e Desenvolvimento Regional, Cabo Verde, pp. 1253-1280. https://www.researchgate.net/profile/Joaquim-Antunes-2/publication/267253794_O_contributo_do_Enoturismo_para_o_desenvolvimento_regional_ o_caso_das_Rotas_dos_Vinhos/links/56323d0908aefa44c36806de/O-contributo-do-Enoturismo-para-o-desenvolvimento-regional-o-caso-das-Rotas-dos-Vinhos.pdf

Nunes, A. M. (2017). Demografia, envelhecimento e saúde: uma análise ao interior de Portugal. *Revista Kairós-Gerontologia*, *20*(1), 133-154. https://doi.org/10.23925/2176-901X.2017v20i1p133-154

OECD (2010). *Regions matter*, OECD.

OECD. (2011). *OECD better life initiative compendium of OECD well-being indicators*. Portugal: OECD. oecd.org/sdd/47917288.pdf

OECD. (2011). *OECD Regional Outlook 2011: Building resilient regions for stronger economies*, OECD Publishing. http://dx.doi.org/10.1787/9789264120983-en

Oliveira, I. D., & Diniz, F. (2018). Turismo e Desenvolvimento Regional – Uma perspetiva do Turismo em Espaço Rural na Serra do Marão, em Portugal. *Turismo e Sociedade*, *11*(1). http://dx.doi.org/10.5380/tes.v11i1.58479

Omrani, H., Alizadeh, A., & Amini, M. (2020). A new approach based on BWM and MULTIMOORA methods for calculating semi-human development index: An application for provinces of Iran. *Socio-Economic Planning Sciences*, *70*, 100689. https://doi.org/10.1016/j.seps.2019.02.004

Ottaviano, G., & Peri, G. (2005). Cities and Culture. *Journal of Urban Economics*, 58:304-337. https://doi.org/10.1016/j.jue.2005.06.004

Permanyer, I. & Smits, J. (2018). The subnational human development index: Moving beyond country-level averages. UNDP. http://hdr.undp.org/en/content/subnational-human-development-index-moving-beyon d-country-level-averages.

Perroux, F. (1950). Economic space: theory and applications. *The quarterly journal of economics*, *64*(1), 89-104. https://doi.org/10.2307/1881960

Popescu, M. C., Balas, V. E., Perescu-Popescu, L., & Mastorakis, N. (2009). Multilayer perceptron and neural networks. *WSEAS Transactions on Circuits and Systems*, *8*(7), 579-588. 29-485-libre.pdf (d1wqtxts1xzle7.cloudfront.net)

Pordata. (2022). *Pordata divulga dados que retratam a pobreza em Portugal e na Europa*. https://www.ffms.pt/sites/default/files/2022-10/CI_09_Pordata_Dia%20Erradica%C3%A7%C3%A3o%20da%20Pobreza_17.10.22.pdf

Ravallion, M. (2010). Mashup Indices of Development, *Policy Research Working Paper 5432*, World Bank. https://doi.org/10.1596/1813-9450-5432

Ravallion, M. (2012). Troubling tradeoffs in the human development index. *Journal of Development Economics*, *99*(2), 201-209. https://doi.org/10.1016/j.jdeveco.2012.01.003

Refsgaard, K., Kull, M., Slätmo, E., & Meijer, M. W. (2021). Bioeconomy – A driver for regional development in the Nordic countries. *New Biotechnology*, *60*, 130-137. DOI: 10.1016/j.nbt.2020.10.001

Ribeiro, J. C., & Vareiro, L. C. (2007). *Turismo e Desenvolvimento Regional: o Espaço Rural como Destino Turístico*. *In* Atas do 1º. Congresso Internacional Casa Nobre: um património para o futuro, pp. 470-486. Edição da Câmara Municipal de Arcos de Valdevez. Arcos de Valdevez. https://hdl.handle.net/1822/6520

Ruivo, P., Rodrigues, J., Neto, M., Oliveira, T., & Johansson, B. (2015). Defining a framework for the development of ICT services “nearshoring” in Portugal. *Procedia Computer Science*, *64*, 140-145. DOI: 10.1016/j.procs.2015.08.474

Sagar, A. D., & Najam, A. (1998). The human development index: a critical review. *Ecological economics*, *25*(3), 249-264. https://doi.org/10.1016/S0921-8009(97)00168-7

Santos, L. D., & Vieira, A. C. (2020). Tourism and regional development: a spatial econometric model for Portugal at municipal level. *Portuguese Economic Journal*, *19*(3), 285-299. [339765.pdf (up.pt)]

Sen, A. (1983). DEVELOPMENT: WHICH WAY NOW? Economic Journal, 93(372 OP-Economic Journal. Dec83, Vol. 93 Issue 372, p745-762. 18p. 2 Charts.), 745. Retrieved from http://search.ebscohost.com/login.aspx?direct=true&site=eds-live&db=bth&AN=4536189

Silva, R., & Ferreira-Lopes, A. (2014). A Regional Development Index for Portugal. *Social Indicators Research*, *118*(3), 1055–1085. http://www.jstor.org/stable/24721051

Soukiazis, E., & Proença, S. (2008). Tourism as an alternative source of regional growth in Portugal: a panel data analysis at NUTS II and III levels. *Portuguese Economic Journal*, *7*(1), 43-61. DOI: 10.1007/s10258-007-0022-0

Stanickova, M., & Melecký, L. (2018). Understanding of resilience in the context of regional development using composite index approach: the case of European Union NUTS-2 regions. *Regional Studies, Regional Science*, *5*(1), 231-254. DOI: 10.1080/21681376.2018.1470939

Tsani, S., Koundouri, P., & Akinsete, E. (2020). Resource management and sustainable development: A review of the European water policies in accordance with the United Nations' Sustainable Development Goals. *Environmental Science & Policy*, *114*, 570-579. https://doi.org/10.1016/j.envsci.2020.09.008

The Economist. (2012). *The lottery of life: where to be born in 2013*. New York. https://www.economist.com/news/2012/11/21/the-lottery-of-life

Tranos, E. (2012). The causal effect of the internet infrastructure on the economic development of European city regions. *Spatial Economic Analysis*, *7*(3), 319-337. https://doi.org/10.1080/17421772.2012.694140

Trigo, A., & Silva, P. (2022). Sustainable Development Directions for Wine Tourism in Douro Wine Region, Portugal. *Sustainability*, *14*(7), 3949. MDPI AG. Retrieved from http://dx.doi.org/10.3390/su14073949

United Nations. (2015). *Transforming our World: The 2030 Agenda for Sustainable Development.* Retrieved from: https://sustainabledevelopment.un.org/content/documents/21252030%20Agenda%20for%20Sustainable%20Development%20web.pdf

UNDP, 1990. Human Development Report 1990. Oxford University Press, New York. hdr.undp.org/system/files/documents/hdr1990encompletenostatspdf.pdf

UNDP (United Nations Development Programme). 2010. Human Development Report 2010: The Real Wealth of Nations: Pathways to Human Development. New York. human-development-report-2010-complete-english.human-development-report-2010-complete-english (undp.org)

UNDP (United Nations Development Programme). 2020. Human Development Report 2020: The Next Frontier: Human Development and the Anthropocene. New York. hdr.undp.org/system/files/documents/hdr2020pdf.pdf

UNDP (United Nations Development Programme). 2022. Human Development Report 2021-22: Uncertain Times, Unsettled Lives: Shaping our Future in a Transforming World. New York. hdr.undp.org/system/files/documents/global-report-document/hdr2021-22pdf_1.pdf

Viegas, M., & Antunes, M. (2013). Convergence in the Spanish and Portuguese NUTS 3 regions: An exploratory spatial approach. *Intereconomics*, *48*(1), 59-66. https://doi.org/10.1007/s10272-013-0445-z

Wren, C. (1996). *Industrial subsidies: The UK experience*. Springer.

# Annexes

Annex A: map and list of NUTS 2013 in Portugal.

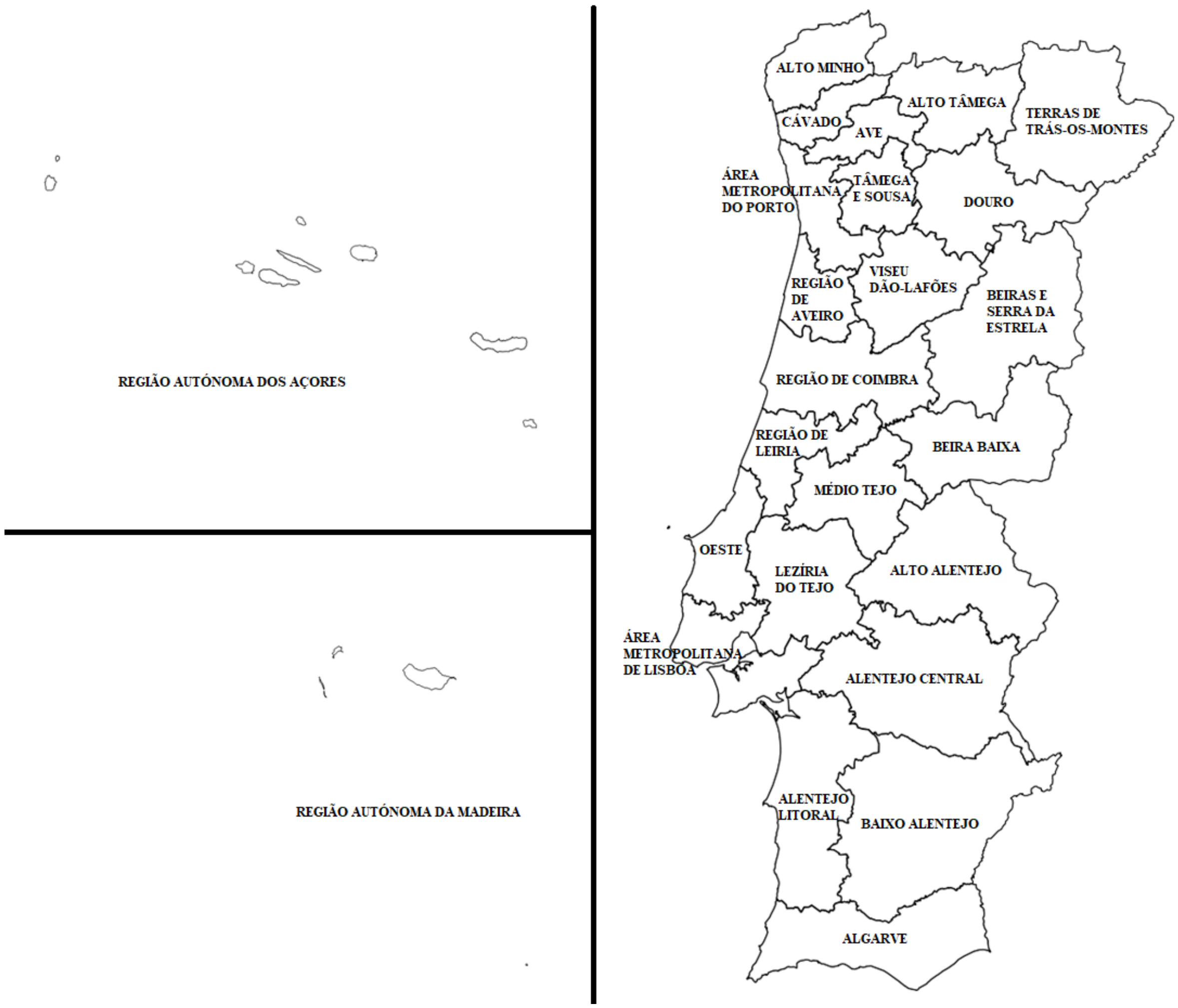


Fig. 3 – Map of Portugal with NUTS 2013. Source: own elaboration[24].

[24] Made with R, version 4.2.1, and RStudio, package rgdal. For accessing the Portuguese NUTS III shapefiles, see: NUTS - GISCO - Eurostat (europa.eu).

List of NUTS 2013 in Portugal. Source: INE.

*Continente (NUTS I – 5 NUTS II – 23 NUTS III)*

*Norte (NUTS II – 8 NUTS III)*

Alto Minho

Cávado

Ave

Área Metropolitana do Porto

Alto Tâmega

Tâmega e Sousa

Douro

Terras de Trás-os-Montes

*Centro (NUTS II – 8 NUTS III)*

Oeste

Região de Aveiro

Região de Coimbra

Região de Leiria

Viseu Dão-Lafões

Beira Baixa

Médio Tejo

Beiras e Serra da Estrela

*Área Metropolitana de Lisboa (NUTS II – 1 NUTS III)*

*Alentejo (NUTS II – 5 NUTS III)*

Alentejo Litoral

Baixo Alentejo

Lezíria do Tejo

Alto Alentejo

Alentejo Central

*Algarve (NUTS II – 1 NUTS III)*

*Região Autónoma dos Açores (NUTS II – 1 NUTS III)*

*Região Autónoma da Madeira (NUTS II – 1 NUTS III*

Annex B: list of HDI from the UNDP, alternative composite development indices, and assessment frameworks over the years. Source: own elaboration.

| **Name** | **Dimensions** | **Number of indicators** | **Source** |
|---|---|---|---|
| Physical Quality of Life Index | Health, education | 3 | Morris (1979) |
| Human Development Index | Income, health, education | 4 | UNDP (1990 – …) |
| Gender Development Index | Income, health, education | 8 | UNDP (1995 – …) |
| Life Quality Index | Income, health | 2 | Nathwani et al. (1997) |
| Modified Human Development Index | Income, health, education | 4 | Noorbakhsh (1998) |
| Generalised Development Index | Income, health, education | 4 | Chakravarty (2003) |
| Inequality-Adjusted Human Development Index | Income, health, education | 4 | Foster et al. (2005), UNDP (2010 – …) |
| Legatum Prosperity Index | Safety & security, personal freedom, governance pillar, social capital, investment environment, enterprise conditions, infrastructure & market access, economic quality, living conditions, health, education, natural environment | 67 | Legatum Institute (2007 – …) |
| Human Development Index by Income Groups | Income, health, education | 4 | Grimm et al. (2008) |
| Alternative Human Development Index | Standard of living, health, education | 4 | Kovacevic (2010) |
| Multidimensional Poverty Index | Income, health, education | 10 | UNDP (2010 – …), Alkire and Jahan (2018) |
| Human Development Index by Internal Migration Status | Income, health, education | 4 | Harttgen and Klasen (2011) |
| Better Life Index | Housing, income, jobs, community, education, environment, civic engagement, health, life satisfaction, safety, work-life balance | 24 | OECD (2011 – …) |
| Alternative Index | Income, health, education | 4 | Ravallion (2012) |

| | | | |
|---|---|---|---|
| Household-Based Human Development Index | Income, health, education | 4 | Harttgen and Klasen (2012) |
| Composite Dynamic Human Development Index | Health, education, economic welfare, inequality, poverty, gender situation, sustainability, personal safety | 11 | Bilbao-Ubillos (2013) |
| Where To Be Born Index | Material well-being, life expectancy at birth, quality of family life, state of political freedom, job security, climate, personal physical security, quality of community life, governance, gender equality | 12 | The Economist (2012) |
| Good Country Index | Science & technology, culture, international peace & security, world order, planet & climate, prosperity & equality, health & wellbeing | 35 | Anholt (2014 – …)[25] |
| Human Sustainable Development Index | Income, health, education, environment | 5 | Bravo (2014) |
| Portuguese Regional Development Index | Income, health, education, governance, environment | 6 | Silva and Ferreira-Lopes (2014) |
| Flow-Based Human Development Index | Income, health, education | 3 | Hou et al. (2015) |
| Municipal Human Development Index | Income, health, education | 4 | Pereira and Mota (2016) |
| Subnational Human Development Index | Income, health, education | 4 | Permanyer and Smits (2018) |
| Index H | Income, health, education | 4 | Lind (2019) |
| Semi-Human Development Index | Income, health, education | 25 | Omrani et al. (2020) |
| Planetary Pressures-Adjusted Human Development Index | Income, health, education, environment | 6 | UNDP (2020 – …) |

[25] Retrieved from Good Country.

Annex C: correlation matrices of the explanatory variables. Source: own elaboration[26].

| Model 1 | unemp | invest | credit | productivity | crime | metropol |
|---|---|---|---|---|---|---|
| unemp | 1,00 | | | | | |
| invest | -0,10 | 1,00 | | | | |
| credit | 0,05 | 0,03 | 1,00 | | | |
| productivity | -0,34 | 0,27 | 0,55 | 1,00 | | |
| crime | 0,07 | 0,03 | 0,52 | 0,33 | 1,00 | |
| metropol | 0,12 | -0,07 | 0,69 | 0,38 | 0,28 | 1,00 |

| Model 2 | youthunemp | impshare | hitecexp | ict | coastline | coverage |
|---|---|---|---|---|---|---|
| youthunemp | 1,00 | | | | | |
| impshare | -0,08 | 1,00 | | | | |
| hitecexp | 0,07 | 0,20 | 1,00 | | | |
| ict | -0,32 | 0,72 | 0,39 | 1,00 | | |
| coastline | -0,18 | 0,26 | 0,32 | 0,49 | 1,00 | |
| coverage | 0,07 | -0,18 | -0,15 | -0,20 | -0,16 | 1,00 |

| Model 3 | sector2 | sector3 | socinteg | density | immigshare | eleccons |
|---|---|---|---|---|---|---|
| sector2 | 1,00 | | | | | |
| sector3 | 0,28 | 1,00 | | | | |
| socinteg | -0,39 | -0,20 | 1,00 | | | |
| density | 0,01 | 0,62 | 0,02 | 1,00 | | |
| immigshare | -0,26 | 0,36 | -0,19 | 0,10 | 1,00 | |
| eleccons | -0,08 | 0,11 | -0,21 | -0,16 | 0,39 | 1,00 |

| Model 4 | sector1 | ageing | enrollrt | doctors | cityarea | classarea |
|---|---|---|---|---|---|---|
| sector1 | 1,00 | | | | | |
| ageing | 0,54 | 1,00 | | | | |
| enrollrt | 0,05 | 0,36 | 1,00 | | | |
| doctors | -0,28 | 0,00 | 0,65 | 1,00 | | |
| cityarea | -0,40 | -0,34 | 0,14 | 0,36 | 1,00 | |
| classarea | 0,34 | -0,01 | 0,00 | -0,09 | -0,03 | 1,00 |

| Model 5 | expshare | area | waste | internet | atm |
|---|---|---|---|---|---|
| expshare | 1,00 | | | | |
| area | -0,27 | 1,00 | | | |
| waste | -0,03 | 0,34 | 1,00 | | |
| internet | 0,28 | 0,01 | 0,12 | 1,00 | |
| atm | -0,15 | 0,58 | 0,28 | 0,38 | 1,00 |

[26] Made with R, version 4.3.0, and RStudio.

Annex D: full panel data estimations. Source: own elaboration[27].

Model 1

```
Pooling Model

Call:
plm(formula = Y ~ X, data = pdata, model = "pooling")

Balanced Panel: n = 25, T = 9, N = 225

Residuals:
      Min.    1st Qu.     Median    3rd Qu.       Max.
-0.5879783 -0.0226475  0.0058881  0.0325465  0.0836547

Coefficients:
                 Estimate  Std. Error t-value  Pr(>|t|)
(Intercept)    5.4489e-01  3.6179e-02 15.0609 < 2.2e-16 ***
Xunemp        -1.3673e-02  2.4038e-03 -5.6878 4.111e-08 ***
Xinvest        2.0079e-03  5.2466e-04  3.8271 0.0001694 ***
Xcredit        1.0219e-05  3.9520e-06  2.5857 0.0103686 *
Xproductivity  8.1634e-03  1.0324e-03  7.9072 1.302e-13 ***
Xcrime        -2.7123e-03  7.6819e-04 -3.5308 0.0005055 ***
Xmetropol     -4.4978e-03  2.0333e-02 -0.2212 0.8251363
---
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1

Total Sum of Squares:    1.7564
Residual Sum of Squares: 0.74352
R-Squared:      0.57668
Adj. R-Squared: 0.56503
F-statistic: 49.4962 on 6 and 218 DF, p-value: < 2.22e-16


Oneway (individual) effect Between Model

Call:
plm(formula = Y ~ X, data = pdata, model = "between")

Balanced Panel: n = 25, T = 9, N = 225
Observations used in estimation: 25

Residuals:
     Min.   1st Qu.    Median   3rd Qu.      Max.
-0.121404 -0.015376  0.005320  0.021814  0.061266

Coefficients:
                 Estimate  Std. Error t-value  Pr(>|t|)
(Intercept)    4.4611e-01  9.5854e-02  4.6541 0.0001973 ***
Xunemp        -1.0280e-03  8.3279e-03 -0.1234 0.9031293
Xinvest        2.1633e-03  1.5641e-03  1.3831 0.1835551
Xcredit        1.1911e-05  1.0044e-05  1.1859 0.2511000
Xproductivity  8.1932e-03  2.7087e-03  3.0248 0.0072823 **
Xcrime        -2.0308e-03  2.0061e-03 -1.0123 0.3248044
Xmetropol     -2.5883e-02  5.0555e-02 -0.5120 0.6148847
---
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1

Total Sum of Squares:    0.10214
Residual Sum of Squares: 0.035912
R-Squared:      0.64839
Adj. R-Squared: 0.53119
F-statistic: 5.53227 on 6 and 18 DF, p-value: 0.0021152
```

[27] Made with R, version 4.3.0, and RStudio, package plm.

```
Oneway (individual) effect First-Difference Model

Call:
plm(formula = Y ~ X, data = pdata, model = "fd")

Balanced Panel: n = 25, T = 9, N = 225
Observations used in estimation: 200

Residuals:
      Min.    1st Qu.     Median    3rd Qu.       Max. 
-0.0409309 -0.0092245 -0.0027254  0.0034376  0.4348041 

Coefficients:
                 Estimate  Std. Error t-value Pr(>|t|)   
(Intercept)    9.1380e-03  3.3546e-03  2.7241 0.007037 **
Xunemp        -1.4127e-02  4.8403e-03 -2.9185 0.003932 **
Xinvest        9.1984e-05  4.9472e-04  0.1859 0.852693   
Xcredit       -1.3372e-07  6.2249e-06 -0.0215 0.982884   
Xproductivity  4.1630e-03  1.8882e-03  2.2047 0.028651 * 
Xcrime         7.6973e-05  1.2404e-03  0.0621 0.950584   
---
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1

Total Sum of Squares:    0.24604
Residual Sum of Squares: 0.21802
R-Squared:      0.11387
Adj. R-Squared: 0.091033
F-statistic: 4.98597 on 5 and 194 DF, p-value: 0.00025544


Oneway (individual) effect Within Model

Call:
plm(formula = Y ~ X, data = pdata, model = "within")

Balanced Panel: n = 25, T = 9, N = 225

Residuals:
       Min.     1st Qu.      Median     3rd Qu.        Max. 
-0.44371251 -0.01102997  0.00052009  0.01328594  0.10349479 

Coefficients:
                 Estimate  Std. Error  t-value Pr(>|t|)    
Xunemp        -3.2780e-02  3.1775e-03 -10.3162  < 2e-16 ***
Xinvest       -6.3630e-04  6.3616e-04  -1.0002  0.31844    
Xcredit        1.4428e-05  6.9325e-06   2.0813  0.03871 *  
Xproductivity  2.3461e-03  1.9108e-03   1.2278  0.22101    
Xcrime        -1.4467e-03  1.3990e-03  -1.0341  0.30237    
---
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1

Total Sum of Squares:    0.83717
Residual Sum of Squares: 0.30221
R-Squared:      0.63901
Adj. R-Squared: 0.58532
F-statistic: 69.0352 on 5 and 195 DF, p-value: < 2.22e-16
```

```
Oneway (individual) effect Random Effect Model
   (Swamy-Arora's transformation)

Call:
plm(formula = Y ~ X, data = pdata, model = "random")

Balanced Panel: n = 25, T = 9, N = 225

Effects:
                   var  std.dev share
idiosyncratic 0.001550 0.039368  0.46
individual    0.001823 0.042696  0.54
theta: 0.7062

Residuals:
      Min.    1st Qu.     Median    3rd Qu.       Max.
-0.4862418 -0.0122954  0.0011466  0.0171086  0.0750762

Coefficients:
                 Estimate  Std. Error z-value  Pr(>|z|)
(Intercept)    6.9309e-01  6.2707e-02 11.0527 < 2.2e-16 ***
Xunemp        -2.6252e-02  2.7652e-03 -9.4938 < 2.2e-16 ***
Xinvest        3.5968e-04  5.8367e-04  0.6162 0.5377317
Xcredit        1.2520e-05  5.6756e-06  2.2059 0.0273934 *
Xproductivity  5.4880e-03  1.4919e-03  3.6784 0.0002347 ***
Xcrime        -1.9542e-03  1.1035e-03 -1.7709 0.0765721 .
Xmetropol      7.8102e-03  4.0862e-02  0.1911 0.8484185
---
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1

Total Sum of Squares:    0.91651
Residual Sum of Squares: 0.36012
R-Squared:      0.60708
Adj. R-Squared: 0.59626
Chisq: 336.813 on 6 DF, p-value: < 2.22e-16


> #LM test for random effects versus OLS
> plmtest(pooling)

        Lagrange Multiplier Test - (Honda)

data:  Y ~ X
normal = 13.142, p-value < 2.2e-16
alternative hypothesis: significant effects

> #LM test for fixed effects versus OLS
> pFtest(fixed, pooling)

        F test for individual effects

data:  Y ~ X
F = 12.38, df1 = 23, df2 = 195, p-value < 2.2e-16
alternative hypothesis: significant effects

> #Hausman test for fixed effects versus random effects
> phtest(random, fixed)

        Hausman Test

data:  Y ~ X
chisq = 9.0827, df = 5, p-value = 0.1058
alternative hypothesis: one model is inconsistent
```

Model 2

Pooling Model

```
Call:
plm(formula = Y ~ X, data = pdata, model = "pooling")

Balanced Panel: n = 25, T = 9, N = 225

Residuals:
     Min.   1st Qu.    Median   3rd Qu.      Max.
-0.671779 -0.029742  0.003485  0.033857  0.139051

Coefficients:
               Estimate  Std. Error t-value  Pr(>|t|)
(Intercept)  7.6664e-01  2.5720e-02 29.8073 < 2.2e-16 ***
Xyouthunemp -9.5157e-03  1.8746e-03 -5.0763 8.242e-07 ***
Ximpshare   -8.0402e-04  7.5523e-04 -1.0646   0.28823
Xhitecexp    3.9521e-03  1.7518e-03  2.2561   0.02506 *
Xict         8.8288e-02  2.1753e-02  4.0586 6.880e-05 ***
Xcoastline  -2.5058e-02  1.1787e-02 -2.1258   0.03464 *
Xcoverage    7.5568e-05  3.9935e-05  1.8923   0.05978 .
---
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1

Total Sum of Squares:    1.7564
Residual Sum of Squares: 1.228
R-Squared:      0.30085
Adj. R-Squared: 0.2816
F-statistic: 15.6343 on 6 and 218 DF, p-value: 6.5927e-15
```

Oneway (individual) effect Between Model

```
Call:
plm(formula = Y ~ X, data = pdata, model = "between")

Balanced Panel: n = 25, T = 9, N = 225
Observations used in estimation: 25

Residuals:
      Min.    1st Qu.     Median    3rd Qu.       Max.
-0.1685431 -0.0128464  0.0073812  0.0186739  0.1269012

Coefficients:
               Estimate  Std. Error t-value  Pr(>|t|)
(Intercept)  0.60848807  0.09164888  6.6393 3.126e-06 ***
Xyouthunemp  0.00537368  0.00784343  0.6851   0.50200
Ximpshare   -0.00085477  0.00199218 -0.4291   0.67297
Xhitecexp    0.00257974  0.00534534  0.4826   0.63519
Xict         0.10558297  0.06060593  1.7421   0.09855 .
Xcoastline  -0.00871911  0.03029606 -0.2878   0.77679
Xcoverage    0.00017792  0.00014486  1.2282   0.23520
---
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1

Total Sum of Squares:    0.10214
Residual Sum of Squares: 0.068833
R-Squared:      0.32608
Adj. R-Squared: 0.10144
F-statistic: 1.45157 on 6 and 18 DF, p-value: 0.25002
```

```
Oneway (individual) effect First-Difference Model

Call:
plm(formula = Y ~ X, data = pdata, model = "fd")

Balanced Panel: n = 25, T = 9, N = 225
Observations used in estimation: 200

Residuals:
      Min.    1st Qu.     Median    3rd Qu.       Max. 
-0.0588346 -0.0095824 -0.0025071  0.0047683  0.4279678 

Coefficients:
               Estimate  Std. Error t-value  Pr(>|t|)    
(Intercept)  1.3987e-02  3.4526e-03  4.0512 7.366e-05 ***
Xyouthunemp -6.7842e-03  1.9309e-03 -3.5135 0.0005503 ***
Ximpshare    1.1492e-03  5.7137e-03  0.2011 0.8408010    
Xhitecexp   -7.4161e-04  1.5441e-03 -0.4803 0.6315711    
Xict        -2.5438e-02  4.9377e-02 -0.5152 0.6070186    
Xcoverage   -1.2381e-06  2.2068e-05 -0.0561 0.9553155    
---
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1

Total Sum of Squares:    0.24604
Residual Sum of Squares: 0.22871
R-Squared:      0.07044
Adj. R-Squared: 0.046482
F-statistic: 2.94016 on 5 and 194 DF, p-value: 0.013931


Oneway (individual) effect Within Model

Call:
plm(formula = Y ~ X, data = pdata, model = "within")

Balanced Panel: n = 25, T = 9, N = 225

Residuals:
       Min.     1st Qu.      Median     3rd Qu.        Max. 
-4.7587e-01 -1.2576e-02  8.8005e-05  1.1813e-02  1.3261e-01 

Coefficients:
               Estimate  Std. Error  t-value Pr(>|t|)    
Xyouthunemp -1.6690e-02  1.4968e-03 -11.1502  < 2e-16 ***
Ximpshare    2.6276e-03  5.2088e-03   0.5045  0.61451    
Xhitecexp    7.1858e-04  1.8832e-03   0.3816  0.70319    
Xict         1.0233e-01  3.3876e-02   3.0206  0.00286 ** 
Xcoverage    8.2884e-06  3.1721e-05   0.2613  0.79415    
---
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1

Total Sum of Squares:    0.83717
Residual Sum of Squares: 0.36553
R-Squared:      0.56338
Adj. R-Squared: 0.49845
F-statistic: 50.3225 on 5 and 195 DF, p-value: < 2.22e-16
```

```
Oneway (individual) effect Random Effect Model
   (Swamy-Arora's transformation)

Call:
plm(formula = Y ~ X, data = pdata, model = "random")

Balanced Panel: n = 25, T = 9, N = 225

Effects:
                   var   std.dev share
idiosyncratic 0.001874 0.043295 0.341
individual    0.003616 0.060131 0.659
theta: 0.7666

Residuals:
      Min.    1st Qu.     Median    3rd Qu.       Max.
-0.5148039 -0.0137147 -0.0011399  0.0169670  0.0979599

Coefficients:
               Estimate  Std. Error  z-value  Pr(>|z|)
(Intercept)  8.3847e-01  3.1784e-02  26.3803 < 2.2e-16 ***
Xyouthunemp -1.6244e-02  1.4348e-03 -11.3217 < 2.2e-16 ***
Ximpshare   -8.5230e-04  1.4155e-03  -0.6021  0.547081
Xhitecexp    1.3932e-03  1.7836e-03   0.7811  0.434737
Xict         9.3955e-02  2.8197e-02   3.3321  0.000862 ***
Xcoastline  -3.1513e-02  2.7161e-02  -1.1602  0.245951
Xcoverage    1.3103e-05  3.1390e-05   0.4174  0.676371
---
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1

Total Sum of Squares:    0.88724
Residual Sum of Squares: 0.42125
R-Squared:      0.52521
Adj. R-Squared: 0.51214
Chisq: 241.152 on 6 DF, p-value: < 2.22e-16


> #LM test for random effects versus OLS
> plmtest(pooling)

        Lagrange Multiplier Test - (Honda)

data:  Y ~ X
normal = 17.472, p-value < 2.2e-16
alternative hypothesis: significant effects

> #LM test for fixed effects versus OLS
> pFtest(fixed, pooling)

        F test for individual effects

data:  Y ~ X
F = 20.005, df1 = 23, df2 = 195, p-value < 2.2e-16
alternative hypothesis: significant effects

> #Hausman test for fixed effects versus random effects
> phtest(random, fixed)

        Hausman Test

data:  Y ~ X
chisq = 64.197, df = 5, p-value = 1.645e-12
alternative hypothesis: one model is inconsistent
```

Model 3

```
Pooling Model

Call:
plm(formula = Y ~ X, data = pdata, model = "pooling")

Balanced Panel: n = 25, T = 9, N = 225

Residuals:
      Min.    1st Qu.     Median    3rd Qu.       Max.
-0.6422194 -0.0340494  0.0079749  0.0407847  0.1230244

Coefficients:
               Estimate  Std. Error t-value  Pr(>|t|)
(Intercept)  6.5390e-01  4.0637e-02 16.0913 < 2.2e-16 ***
Xsector3     2.8065e-01  6.8671e-02  4.0870 6.144e-05 ***
Xsocinteg   -6.5146e-04  2.9011e-04 -2.2455   0.02574 *
Xdensity     2.7240e-06  2.9014e-05  0.0939   0.92529
Ximmigshare -1.1410e-03  1.7395e-03 -0.6559   0.51255
Xeleccons    1.1229e-05  2.0772e-06  5.4056 1.689e-07 ***
Xsector2    -7.8917e-01  1.4442e-01 -5.4643 1.265e-07 ***
---
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1

Total Sum of Squares:    1.7564
Residual Sum of Squares: 1.137
R-Squared:      0.35268
Adj. R-Squared: 0.33486
F-statistic: 19.7953 on 6 and 218 DF, p-value: < 2.22e-16
```

```
Oneway (individual) effect Between Model

Call:
plm(formula = Y ~ X, data = pdata, model = "between")

Balanced Panel: n = 25, T = 9, N = 225
Observations used in estimation: 25

Residuals:
      Min.    1st Qu.     Median    3rd Qu.       Max.
-0.1097547 -0.0232814  0.0010059  0.0345010  0.0641248

Coefficients:
               Estimate  Std. Error t-value  Pr(>|t|)
(Intercept)  6.3840e-01  8.4329e-02  7.5703 5.331e-07 ***
Xsector3     2.8338e-01  1.4462e-01  1.9595   0.06573 .
Xsocinteg   -3.5182e-04  6.1618e-04 -0.5710   0.57508
Xdensity     5.0298e-06  5.9956e-05  0.0839   0.93407
Ximmigshare -2.5617e-03  3.9220e-03 -0.6532   0.52191
Xeleccons    1.1568e-05  4.3029e-06  2.6883   0.01502 *
Xsector2    -7.3835e-01  3.0801e-01 -2.3972   0.02759 *
---
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1

Total Sum of Squares:    0.10214
Residual Sum of Squares: 0.043826
R-Squared:      0.57091
Adj. R-Squared: 0.42788
F-statistic: 3.99159 on 6 and 18 DF, p-value: 0.010259
```

```
Oneway (individual) effect First-Difference Model

Call:
plm(formula = Y ~ X, data = pdata, model = "fd")

Balanced Panel: n = 25, T = 9, N = 225
Observations used in estimation: 200

Residuals:
      Min.    1st Qu.     Median    3rd Qu.       Max.
-0.0617037 -0.0107588 -0.0010867  0.0071425  0.4255896

Coefficients:
               Estimate  Std. Error t-value  Pr(>|t|)
(Intercept)  1.6616e-02  3.6165e-03  4.5946 7.816e-06 ***
Xsector3     1.0260e-03  3.3401e-01  0.0031    0.9976
Xsocinteg   -1.1264e-03  1.3248e-03 -0.8502    0.3962
Xdensity    -1.2175e-03  1.7249e-03 -0.7058    0.4811
Ximmigshare -7.5525e-03  5.8021e-03 -1.3017    0.1946
Xeleccons    1.6781e-05  1.2776e-05  1.3135    0.1906
Xsector2    -4.6860e-01  9.5812e-01 -0.4891    0.6253
---
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1

Total Sum of Squares:    0.24604
Residual Sum of Squares: 0.23475
R-Squared:      0.045894
Adj. R-Squared: 0.016233
F-statistic: 1.54728 on 6 and 193 DF, p-value: 0.16475


Oneway (individual) effect Within Model

Call:
plm(formula = Y ~ X, data = pdata, model = "within")

Balanced Panel: n = 25, T = 9, N = 225

Residuals:
      Min.    1st Qu.     Median    3rd Qu.       Max.
-0.4698944 -0.0172565  0.0032793  0.0204559  0.1076681

Coefficients:
               Estimate  Std. Error t-value  Pr(>|t|)
Xsector3     1.04114416  0.35410556  2.9402 0.0036785 **
Xsocinteg   -0.00397771  0.00108835 -3.6548 0.0003313 ***
Xdensity    -0.00507376  0.00160607 -3.1591 0.0018358 **
Ximmigshare  0.00809090  0.00334150  2.4213 0.0163839 *
Xeleccons    0.00002714  0.00001279  2.1220 0.0351079 *
Xsector2    -2.19267965  0.94704356 -2.3153 0.0216430 *
---
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1

Total Sum of Squares:    0.83717
Residual Sum of Squares: 0.49292
R-Squared:      0.41121
Adj. R-Squared: 0.32016
F-statistic: 22.5819 on 6 and 194 DF, p-value: < 2.22e-16
```

```
Oneway (individual) effect Random Effect Model
   (Swamy-Arora's transformation)

Call:
plm(formula = Y ~ X, data = pdata, model = "random")

Balanced Panel: n = 25, T = 9, N = 225

Effects:
                   var  std.dev share
idiosyncratic 0.002541 0.050406 0.541
individual    0.002152 0.046395 0.459
theta: 0.6595

Residuals:
      Min.    1st Qu.     Median    3rd Qu.       Max.
-0.5368410 -0.0311255  0.0043419  0.0319499  0.1342290

Coefficients:
               Estimate  Std. Error z-value  Pr(>|z|)
(Intercept)  7.0268e-01  8.8085e-02  7.9772 1.497e-15 ***
Xsector3     3.4801e-01  1.4327e-01  2.4291   0.01514 *
Xsocinteg   -2.1852e-03  5.6012e-04 -3.9012 9.570e-05 ***
Xdensity    -2.6835e-05  6.5213e-05 -0.4115   0.68071
Ximmigshare  2.7945e-03  2.6780e-03  1.0435   0.29671
Xeleccons    1.0921e-05  4.5776e-06  2.3858   0.01704 *
Xsector2    -1.1714e+00  2.8842e-01 -4.0613 4.881e-05 ***
---
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1

Total Sum of Squares:    0.94376
Residual Sum of Squares: 0.72167
R-Squared:      0.23532
Adj. R-Squared: 0.21428
Chisq: 67.0878 on 6 DF, p-value: 1.6149e-12


> plmtest(pooling)

        Lagrange Multiplier Test - (Honda)

data:  Y ~ X
normal = 8.3656, p-value < 2.2e-16
alternative hypothesis: significant effects

> pFtest(fixed, pooling)

        F test for individual effects

data:  Y ~ X
F = 10.562, df1 = 24, df2 = 194, p-value < 2.2e-16
alternative hypothesis: significant effects

> phtest(random, fixed)

        Hausman Test

data:  Y ~ X
chisq = 85.182, df = 6, p-value = 3.027e-16
alternative hypothesis: one model is inconsistent
```

## Model 4

Pooling Model

```
Call:
plm(formula = Y ~ X, data = pdata, model = "pooling")

Balanced Panel: n = 25, T = 9, N = 225

Residuals:
     Min.   1st Qu.    Median   3rd Qu.      Max.
-0.724137 -0.032438  0.012232  0.046246  0.145442

Coefficients:
               Estimate  Std. Error t-value  Pr(>|t|)
(Intercept)  0.68386465  0.02687611 25.4451 < 2.2e-16 ***
Xageing      0.00024238  0.00012292  1.9718 0.0498984 *
Xenrollrt    0.00029422  0.00033645  0.8745 0.3828204
Xdoctors     0.00779839  0.00385437  2.0233 0.0442674 *
Xcityarea   -0.00031820  0.00189498 -0.1679 0.8668056
Xclassarea   0.00176053  0.00047156  3.7335 0.0002411 ***
Xsector1    -0.20416936  0.05825543 -3.5047 0.0005547 ***
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Total Sum of Squares:    1.7564
Residual Sum of Squares: 1.4658
R-Squared:      0.16544
Adj. R-Squared: 0.14247
F-statistic: 7.20272 on 6 and 218 DF, p-value: 5.0288e-07
```

Oneway (individual) effect Between Model

```
Call:
plm(formula = Y ~ X, data = pdata, model = "between")

Balanced Panel: n = 25, T = 9, N = 225
Observations used in estimation: 25

Residuals:
      Min.    1st Qu.     Median    3rd Qu.       Max.
-0.1969249 -0.0124935  0.0038215  0.0270196  0.1175800

Coefficients:
               Estimate  Std. Error t-value Pr(>|t|)
(Intercept)  0.75723464  0.07092297 10.6769 3.23e-09 ***
Xageing     -0.00011731  0.00032957 -0.3560   0.7260
Xenrollrt    0.00072133  0.00086772  0.8313   0.4167
Xdoctors     0.00171429  0.00988158  0.1735   0.8642
Xcityarea   -0.00040282  0.00466743 -0.0863   0.9322
Xclassarea   0.00136570  0.00117122  1.1661   0.2588
Xsector1    -0.12774432  0.14787534 -0.8639   0.3990
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Total Sum of Squares:    0.10214
Residual Sum of Squares: 0.081387
R-Squared:      0.20317
Adj. R-Squared: -0.062447
F-statistic: 0.764895 on 6 and 18 DF, p-value: 0.60689
```

```
Oneway (individual) effect First-Difference Model

Call:
plm(formula = Y ~ X, data = pdata, model = "fd")

Balanced Panel: n = 25, T = 9, N = 225
Observations used in estimation: 200

Residuals:
      Min.    1st Qu.     Median    3rd Qu.       Max. 
-0.0640863 -0.0117282 -0.0015737  0.0075383  0.4283184 

Coefficients:
               Estimate  Std. Error t-value  Pr(>|t|)    
(Intercept)  0.01888122  0.00544608  3.4669 0.0006482 ***
Xageing      0.00039576  0.00036165  1.0943 0.2751696    
Xenrollrt   -0.00054877  0.00126977 -0.4322 0.6660923    
Xdoctors    -0.02335242  0.02523648 -0.9253 0.3559367    
Xclassarea   0.00406742  0.01687218  0.2411 0.8097534    
Xsector1     0.19669738  0.21236713  0.9262 0.3554856    
---
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1

Total Sum of Squares:    0.24604
Residual Sum of Squares: 0.24169
R-Squared:      0.017679
Adj. R-Squared: -0.0076388
F-statistic: 0.698281 on 5 and 194 DF, p-value: 0.62536
```

```
Oneway (individual) effect Within Model

Call:
plm(formula = Y ~ X, data = pdata, model = "within")

Balanced Panel: n = 25, T = 9, N = 225

Residuals:
      Min.    1st Qu.     Median    3rd Qu.       Max. 
-0.4889870 -0.0103397  0.0024215  0.0138067  0.1152538 

Coefficients:
              Estimate  Std. Error t-value  Pr(>|t|)    
Xageing     0.00174075  0.00032610  5.3380 2.597e-07 ***
Xenrollrt  -0.00033402  0.00069997 -0.4772  0.633758    
Xdoctors    0.03938328  0.01293423  3.0449  0.002649 ** 
Xclassarea -0.00704909  0.01541487 -0.4573  0.647971    
Xsector1    0.35270079  0.23833805  1.4798  0.140531    
---
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1

Total Sum of Squares:    0.83717
Residual Sum of Squares: 0.37693
R-Squared:      0.54976
Adj. R-Squared: 0.4828
F-statistic: 47.6196 on 5 and 195 DF, p-value: < 2.22e-16
```

```
Oneway (individual) effect Random Effect Model
   (Swamy-Arora's transformation)

Call:
plm(formula = Y ~ X, data = pdata, model = "random")

Balanced Panel: n = 25, T = 9, N = 225

Effects:
                   var  std.dev share
idiosyncratic 0.001933 0.043966  0.31
individual    0.004307 0.065626  0.69
theta: 0.7821

Residuals:
      Min.    1st Qu.     Median    3rd Qu.       Max.
-0.5289128 -0.0161212  0.0054476  0.0219567  0.1002039

Coefficients:
               Estimate  Std. Error z-value  Pr(>|z|)
(Intercept)  0.41370168  0.05017443  8.2453 < 2.2e-16 ***
Xageing      0.00138572  0.00019408  7.1398 9.347e-13 ***
Xenrollrt   -0.00074082  0.00057567 -1.2869  0.198133
Xdoctors     0.02942100  0.00748511  3.9306 8.473e-05 ***
Xcityarea    0.00104196  0.00500366  0.2082  0.835041
Xclassarea   0.00259202  0.00121609  2.1314  0.033054 *
Xsector1    -0.32121479  0.12442748 -2.5815  0.009836 **
---
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1

Total Sum of Squares:    0.88084
Residual Sum of Squares: 0.5059
R-Squared:      0.42566
Adj. R-Squared: 0.40985
Chisq: 161.565 on 6 DF, p-value: < 2.22e-16


> plmtest(pooling)

        Lagrange Multiplier Test - (Honda)

data:  Y ~ X
normal = 14.556, p-value < 2.2e-16
alternative hypothesis: significant effects

> pFtest(fixed, pooling)

        F test for individual effects

data:  Y ~ X
F = 24.492, df1 = 23, df2 = 195, p-value < 2.2e-16
alternative hypothesis: significant effects

> phtest(random, pooling)

        Hausman Test

data:  Y ~ X
chisq = 261.89, df = 6, p-value < 2.2e-16
alternative hypothesis: one model is inconsistent
```

## Model 5

Pooling Model

```
Call:
plm(formula = Y ~ X, data = pdata, model = "pooling")

Balanced Panel: n = 25, T = 9, N = 225

Residuals:
       Min.     1st Qu.      Median     3rd Qu.        Max.
-0.58531456 -0.02123347 -0.00049635  0.03228586  0.11896583

Coefficients:
               Estimate  Std. Error t-value  Pr(>|t|)
(Intercept)  5.0093e-01  2.6302e-02 19.0450 < 2.2e-16 ***
Xexpshare    6.6740e-04  6.9257e-04  0.9636 0.3362861
Xarea        6.0889e-03  2.4963e-03  2.4392 0.0155159 *
Xwaste      -2.8177e-04  7.2198e-05 -3.9027 0.0001266 ***
Xinternet    9.0442e-03  8.0998e-04 11.1660 < 2.2e-16 ***
Xatm         8.1468e-03  3.0442e-03  2.6762 0.0080093 **
---
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1

Total Sum of Squares:    1.7564
Residual Sum of Squares: 0.83432
R-Squared:      0.52499
Adj. R-Squared: 0.51414
F-statistic: 48.408 on 5 and 219 DF, p-value: < 2.22e-16
```

Oneway (individual) effect Between Model

```
Call:
plm(formula = Y ~ X, data = pdata, model = "between")

Balanced Panel: n = 25, T = 9, N = 225
Observations used in estimation: 25

Residuals:
      Min.    1st Qu.     Median    3rd Qu.       Max.
-0.1010315 -0.0159627  0.0099129  0.0265462  0.0710772

Coefficients:
               Estimate  Std. Error t-value  Pr(>|t|)
(Intercept)  0.47470449  0.07101851  6.6842 2.166e-06 ***
Xexpshare   -0.00067003  0.00260122 -0.2576   0.79950
Xarea        0.00872181  0.00785967  1.1097   0.28098
Xwaste      -0.00044166  0.00025643 -1.7224   0.10124
Xinternet    0.01321513  0.00624721  2.1154   0.04783 *
Xatm         0.00565943  0.01083083  0.5225   0.60734
---
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1

Total Sum of Squares:    0.10214
Residual Sum of Squares: 0.048426
R-Squared:      0.52588
Adj. R-Squared: 0.40111
F-statistic: 4.21483 on 5 and 19 DF, p-value: 0.0095488
```

```
Oneway (individual) effect First-Difference Model

Call:
plm(formula = Y ~ X, data = pdata, model = "fd")

Balanced Panel: n = 25, T = 9, N = 225
Observations used in estimation: 200

Residuals:
      Min.    1st Qu.     Median    3rd Qu.       Max.
-0.0550681 -0.0110971 -0.0017757  0.0071858  0.4360299

Coefficients:
               Estimate  Std. Error t-value Pr(>|t|)
(Intercept)  1.3232e-02  9.0175e-03  1.4673   0.1439
Xexpshare    3.1227e-03  4.4865e-03  0.6960   0.4872
Xwaste       6.7618e-05  1.7822e-04  0.3794   0.7048
Xinternet    1.7339e-03  3.8479e-03  0.4506   0.6528
Xatm        -4.5645e-03  5.1138e-03 -0.8926   0.3732

Total Sum of Squares:    0.24604
Residual Sum of Squares: 0.24323
R-Squared:      0.011401
Adj. R-Squared: -0.0088784
F-statistic: 0.562187 on 4 and 195 DF, p-value: 0.69038


Oneway (individual) effect Within Model

Call:
plm(formula = Y ~ X, data = pdata, model = "within")

Balanced Panel: n = 25, T = 9, N = 225

Residuals:
      Min.    1st Qu.     Median    3rd Qu.       Max.
-0.4800960 -0.0112770  0.0023839  0.0123156  0.1130245

Coefficients:
             Estimate  Std. Error t-value  Pr(>|t|)
Xexpshare -5.6717e-04  5.3321e-03 -0.1064    0.9154
Xwaste    -5.9494e-05  2.3011e-04 -0.2586    0.7963
Xinternet  7.7281e-03  1.1611e-03  6.6557 2.759e-10 ***
Xatm       7.2844e-05  5.8420e-03  0.0125    0.9901
---
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1

Total Sum of Squares:    0.83717
Residual Sum of Squares: 0.37752
R-Squared:      0.54906
Adj. R-Squared: 0.48464
F-statistic: 59.6614 on 4 and 196 DF, p-value: < 2.22e-16
```

```
Oneway (individual) effect Random Effect Model
   (Swamy-Arora's transformation)

Call:
plm(formula = Y ~ X, data = pdata, model = "random")

Balanced Panel: n = 25, T = 9, N = 225

Effects:
                   var  std.dev share
idiosyncratic 0.001926 0.043887 0.452
individual    0.002335 0.048319 0.548
theta: 0.7102

Residuals:
      Min.    1st Qu.     Median    3rd Qu.       Max.
-0.5167533 -0.0086482  0.0033756  0.0149872  0.0751571

Coefficients:
               Estimate  Std. Error z-value  Pr(>|z|)
(Intercept)  0.52832380  0.05388177  9.8052   <2e-16 ***
Xexpshare    0.00074635  0.00152890  0.4882   0.6254
Xarea        0.00728190  0.00557017  1.3073   0.1911
Xwaste      -0.00014636  0.00012345 -1.1855   0.2358
Xinternet    0.00814649  0.00077486 10.5135   <2e-16 ***
Xatm         0.00236121  0.00427473  0.5524   0.5807
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Total Sum of Squares:    0.91436
Residual Sum of Squares: 0.41855
R-Squared:      0.54225
Adj. R-Squared: 0.5318
Chisq: 259.425 on 5 DF, p-value: < 2.22e-16


> #LM test for random effects versus OLS
> plmtest(pooling)

        Lagrange Multiplier Test - (Honda)

data:  Y ~ X
normal = 14.346, p-value < 2.2e-16
alternative hypothesis: significant effects

> #LM test for fixed effects versus OLS
> pFtest(fixed, pooling)

        F test for individual effects

data:  Y ~ X
F = 10.311, df1 = 23, df2 = 196, p-value < 2.2e-16
alternative hypothesis: significant effects

> #Hausman test for fixed versus random effects
> phtest(random, fixed)

        Hausman Test

data:  Y ~ X
chisq = 2.0257, df = 4, p-value = 0.731
alternative hypothesis: one model is inconsistent
```