Quality Engineering

ISSN: 0898-2112 (Print) 1532-4222 (Online) Journal homepage: http://www.tandfonline.com/loi/lqen20

Quality and statistical thinking in a parliament and beyond

Pedro Manuel Saraiva

To cite this article: Pedro Manuel Saraiva (2018) Quality and statistical thinking in a parliament and beyond, Quality Engineering, 30:1, 2-22, DOI: 10.1080/08982112.2017.1368557

To link to this article: https://doi.org/10.1080/08982112.2017.1368557

Accepted author version posted online: 22 Aug 2017.
Published online: 22 Aug 2017.

QUALITY ENGINEERING
2018, VOL. 30, NO. 1, 2–22
https://doi.org/10.1080/08982112.2017.1368557

Taylor & Francis
Taylor & Francis Group

Quality and statistical thinking in a parliament and beyond

Pedro Manuel Saraiva

Chemical Engineering Department, University of Coimbra, Coimbra, Portugal

ABSTRACT

After decades of an intensive academic career, with research activities having a strong statistical and quality background, I was given the opportunity of running twice for election to the National Parliament of Portugal. This article is mostly about the challenging experience of being a Member of Parliament from 2009–2015. I aim to provide sufficient evidence to convince the reader that statistical thinking, quality tools, and fact-based approaches are necessary to better governing. These tools and approaches can help to provide a better understanding of how Parliaments work and some of the strongest features of their organizational culture. Routine application of these tools can yield better results, increased efficiency and efficacy in Parliaments, politics and in public policy making.

For that purpose, I provide specific illustrations, which show how I applied statistical tests, variation analysis, clustering, and Bayesian interpretations to several situations related with the Portuguese Parliament. The goal of this article is to provide enough support to show that: (i) statistical or quality thinking and tools can help to better understand and improve Parliaments, as well as come up with better evidence-based decisions made by politicians; and (ii) Parliaments and societies are likely to get better if more people with a sound statistical background accept the challenge of becoming Members of Parliament, at least for a while.

KEYWORDS

democracy quality; evidence-based policy making; politicians utility functions; quality tools and parliaments; statistical thinking in politics

Introduction

I discovered a passion for applied statistics mostly while being a PhD student at MIT. I remember how Box, Hunter, and Hunter (2005), John Tukey (1977), and Box and Draper (1998) were amongst the classical textbooks that made it happen. They were also sources of inspiration for a Ph.D. thesis titled "Data-Driven Learning Frameworks for Continuous Process Analysis and Improvement", completed back in 1993. Since then, I have been conducting research and teaching activities at the University of Coimbra, mostly in the field of data-driven Process Systems Engineering.

With such a background, the invitation to run for election as Member of the Portuguese Parliament came as a surprise. At the time, 2009, I was Vice-rector of the University of Coimbra and advisor to the President of Portugal for higher education. Unexpectedly, I was then invited to represent my district (Coimbra) and party, the Social Democratic Party (PSD), in the National Parliament. PSD is one of the two leading parties in the country. The other leading party is the Socialist Party or PS.

At the time, I had to make a choice, about whether to keep things pretty much as they stood, or give it a try. I wanted to do more than just blame politics and politicians for not being concerned or able to address societal challenges in an efficient or useful way. Three societal paradoxes, at the crossroads of quality, statistics, politics and politicians, played a major role in making my decision. First, we are not applying enough quality or statistical thinking and tools to domains where they might provide important benefits and lead to large impacts, regarding politics and politicians. Second, the world is becoming ever more complex and data intensive, but politics and politicians are moving to less and less content in their speeches and decisions. They often rely on sound-bites to convey messages. Third, society would benefit from having more Members of Parliament and politicians with a good statistical or quality background, but not many such professionals are willing to serve.

Furthermore, it is important not to forget about the monetary impact related to the quality of policies and political decisions. Quality professionals, are

CONTACT Pedro Manuel Saraiva pas@eq.uc.pt Chemical Engineering Department, University of Coimbra, Rua Silvio Lima, 3030–790 Coimbra, Portugal.
Color versions of one or more of the figures in the article can be found online at www.tandfonline.com/lqen.

accustomed to working on projects with potential monetary impact generally less than a million euros. By contrast, when we deal with policy making and decisions made by Governments and Parliaments, their non-quality costs and impacts can be much larger. Here wrong or suboptimal solutions can often correspond to societal losses on the order of billions of euros. Indeed as recognized by Zonnenshain, Naveh, and Halevy (1998), "the cost of non-quality is one of the major sources of national waste" and public policies have a lot to do with such losses.

All of the above arguments lead to a strong hypothesis that tremendous potential societal gains and impacts can derive from approaches leading politics and politicians to avoid simplistic solutions at the surface of any debate. Rather politicians better serve their constituents by addressing increasingly complex problems in a complex world by: (i) searching for deeper knowledge, root causes and corresponding solutions; (ii) adopting evidence, science and fact-based approaches to decision making; (iii) applying systems quality and statistical thinking, data collection and analysis; and (iv) making speeches having much more substance.

In light of all the above arguments, I had no other choice but to accept this challenge. I ended up being elected for two consecutive terms, one in the opposition and the other supporting Government. This work focusses on perspectives driven from Quality Management (Sampaio and Saraiva 2016) and Statistical Thinking (Hoerl and Snee 2012) resulting from my tenure as a member of Parliament (MP) from 2009–2015. In the following, I express most of the material in a qualitative and non-technical manner.

Quality and statistics can be quite powerful when they are also applied to political organizations or even territories. In doing so, larger scopes and ambitions can be formulated at several layers of geographical scale, ranging from local to global levels (Conti 2003), from quality in our communities to quality of countries or even in the world, as well as democracy quality.

In retrospect, these six years of my life dedicated to political positions: (i) corresponded indeed to a quite rewarding experience; (ii) have shown the potential benefits for Parliaments and politicians to adopt a stronger evidence-based approach to discussion and decision-making, going deeper into the understanding of problem root causes; (iii) make me believe that quality and statistical thinking can play an important role for Parliaments to move in such a direction; and (iv) have shown that societies will benefit a lot from having Members of Parliament with a sound quality and statistical background.

In the forthcoming sections I reflect on a number of different but interconnected topics, dealing with my parliamentary experience, and some of the inspiration coming from it. This also resulted in new ways of thinking or the identification of different types of projects that can be carried out based upon "macroquality" challenges. In this journey, this article will refer to: (i) the need and benefits of deeper analysis and evidence-based politics; (ii) some parliament organizational culture features; (iii) examples of supervised learning and classification; (iv) a Bayesian analysis of divergences; (v) definitions of clusters and unsupervised learning; (vi) illustrations of uses and misuses of data; (vii) a sample of how simple quality or statistical approaches may be helpful; (viii) the power and impacts of fact-based analysis and reasoning; (ix) "macroquality" challenges and how statistical tools can help to address them; and (x) some final thoughts and conclusions.

Unique features and challenges of Parliaments

As was mentioned before, it is paradoxical that currently societal problems are becoming more complex while discussions held in the political arena are simplistic or shallow. The 140 character limit of twitter exacerbates this problem. If not taken seriously, this paradox may lead to societies that handle issues without ever considering root causes.

To overcome these quite dangerous limitations, deeper and stronger evidence-based policy making is needed. This evidence-based paradigm can also be quite helpful in reaching broader agreements between different parties represented in Parliaments. By definition, objective facts are nonpartisan and have no right or left ideologies attached to them. My personal experience as MP has shown how broader convergences of positions can be obtained from a good data-based starting point. This way it becomes easier to focus discussions, positions, and even to achieve larger final voting consensus. In the end, one of the main conclusions I did come up with is that "in Parliaments, facts can be your best friends".

While being a MP, I recollect a sample of 20 areas (Table 1) where in particular I felt that a more

Table 1. A samples of 20 issues where a proper scientific approach and technology assessment may result in better quality of political decisions and strong positive impacts.

Incentives for promoting the use of renewable energies
Performance of low cost vs. premium gas
Approval of drugs for being supported by national health system
Use of stem cells and genomics/DNA information
Clinical trials, bioteries, animal testing
Public policies for Research, Development, Innovation, and Entrepreneurship
Medical prescription by ICD code
Medically assisted procreation
Living will
Large bandwith and digital television coverage
Electricity/gas price formation and tariffs
Regulations for specialized professional activities
Public private partnerships proper evaluation and cost/benefit analysis
Transportation alternatives for people and goods
Large infrastructure investments (such as airports or high-speed trains)
Areas for the active promotion of clusters
European patents
Positions regarding HORIZON 2020 Program
Digital contents intellectual property rights and data access
Incentives for promoting the use of renewable energies

scientific approach, including proper technology assessment efforts in the Parliament, would lead to much better results and decision making quality. This is already recognized by some Parliaments, that do have in place their own independent technology offices, or other related agencies, aimed at providing a professional analysis of relevant topics, to improve the quality of parliamentary discussions and political decision making. That is the case namely in the European Parliament, UK, Germany, Israel, or Switzerland.

Therefore, I am more than ever a strong believer that societies and Parliaments have much to gain by adopting evidence, data-driven, and fact-based approaches, with quality, systems and statistical thinking, to support political decision-making. At the same time, one should go much deeper into proper problem understanding, discussion, and identification of root causes (going from "a zero to a five consecutive whys mindset"), before moving into the definition and approval of possible solutions or decisions.

Of course, all of these paradigms will be made much easier to adopt if further quality professionals play the roles of politicians or at least help them in coming up with an appropriate usage of quality and statistical tools. But such a change will also benefit from a more global effort to fight quantitative illiteracy and promote higher levels of statistical thinking and understanding of quantitative or qualitative data among citizens and societies. If voters end up becoming more literate and knowledgeable in these areas, there would be a much smaller acceptance for data misusages from the side of politicians to "justify" a pre-determined position. This would then also lead to better policies, political decisions, and laws. Furthermore, different attitudes, better aligned with quality engineering, will be adopted by candidates and politicians, since well-informed and trained citizens will demand for such an evolution to occur.

The adoption of quality and statistical thinking paradigms is powerful for gaining better understanding of behaviors regarding politics, politicians, as well as of some key parliamentary organizational culture features. Parliaments are unique types of entities. One of their main features, in terms of organizational culture, derives from the fact that MP and parties do have a strong orientation with regards to some key quantitative indicators (such as the number of MP needed for getting approvals, gross domestic product (GDP) growth, or unemployment rates). Although sometimes these may end up not being the most relevant ones for the creation of global societal value, or used in the best possible ways. But this implies that at least some kind of number-driven attitude is already present, as we will see next with some examples. That being the case, the creation of some introductory bridges between MP and statistical reasoning can be facilitated, if one explores such a kind of pre-existent data orientation.

A number that MP always take into account is the total number of MP, which in the case of Portugal is equal to 230. Several issues surrounding this number, with clear impacts over Parliament behaviors, are the following ones: (i) on the minds of many MP, what they do is strongly related with the final goal of keeping their 1/230 "market share" of seats available at the Parliament. They will try to do whatever is necessary to assure a reelection for the next term. Maximizing this probability becomes closer to what economists would call their utility function; (ii) similar approaches are followed by the corresponding parties, at a more aggregate level. Each one of them tries to maximize an objective function that corresponds to their number of MP in the next elections, with the hard equality constraint that the overall sum must equal 230; and (iii) also on a daily and weekly basis MP make computations to find out possible combinations of parties that will produce a majority of votes in favor or against all kinds of proposals being discussed.

Some of these features and constraints do create a difficult and complex optimization formulation. They are obvious in Parliaments, where the number of MP is often fixed by the National Constitutions. But

there are similar situations elsewhere, in other types of organizations. Namely at universities, where a relatively fixed number of faculty positions are available for certain areas of knowledge, leading also here to the lack of possible collaboration, or even unhealthy competitive processes within a single department or between departments. Generally, all departments want to grow under sometimes severe resource constraints where the number of available positions is fixed. This is similar to the situation in Parliaments where the limit on the number of MP, however, is explicit.

If one looks at such an organizational setting from a systems and operations research perspective, it is easy to understand that many individual and party related local optimum solutions are adopted in Parliaments. This often makes it difficult for Parliaments to achieve results close to what would be the corresponding global optimum options for the countries. That also explains why in Parliaments we typically do find out aggressive environments (with MP seeming to be almost always fighting verbally against each other, sometimes even using words and language that would not be considered appropriate in schools or other types of organizations). That happens because conflicting partial objective functions compete against each other. However, as was exemplified above, one must not forget that possibly similar behaviors might occur at any other type of organization under similar settings. In a company where the single goal of each department would be a maximization of its number of employees, but at the same time the overall total number of employees is kept constant, similar conflicts and behaviors emerge as well. Since in that company for one department to increase others would necessarily decrease, and given that each department does have as its mission to get as large as possible, strong internal competition will appear, and fights are likely to become quite common, just as in Parliaments.

If we examine the time evolution of the number of MP in Portugal for any given party, across several elections, it can be seen that such a variable is typically quite unstable and not by any means kept under statistical process control. This can be seen from a very simple run chart (Figure 1) for the case of Portugal. When we look at data available since democracy was reestablished, back in 1974, there are points clearly outside the control limits, and a boxplot does illustrate that the dispersion of values obtained becomes larger in the most recent times (since 1988), of larger election results

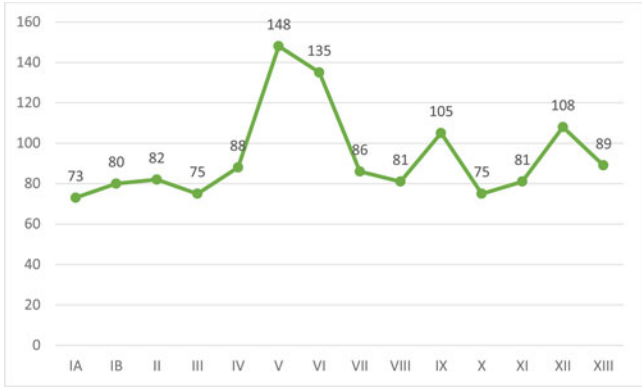

Figure 1. Run chart for the number of MP elected by the Social Democratic Party (PSD) across the different general elections that took place in Portugal between 1976 and 2015.

unpredictability. These results do show how difficult it is to manage or predict this variable. At the same time, one must deal with measurement frequencies that are quite low, since new values are obtained only every four years or so, when a new general election takes place. To make things even more difficult to manage, rather than getting a smaller variation, as is usually aimed in terms of quality management, in the most recent decades dispersions of results have actually increased, not just in Portugal but also in many other countries. Voting patterns and their translation into MP numbers now have much more uncertainties and forecasting errors associated with them, as a result of additional "noise factors" that did come into place, as well as changes in the partisan landscapes of many countries.

One of the most powerful and useful contributions for making bridges between parties and promoting convergences to more global optimum solutions in parliaments therefore relies on the adoption of data-driven and fact-based arguments. They end up facilitating cooperation modes that otherwise it would be much more difficult to handle, given the variety of objective or utility functions being pursued by individual Members of Parliament and their parties. Furthermore, well-known statistical tools and quality management approaches help us to come up with an easier understanding for parliamentary behaviors, as we will also show in the forthcoming paragraphs.

Statistical tools and Parliament understanding

Parliaments and MP are not only used to take into account certain quantitative indicators, but also the

Figure 2. Physical location of MP according to their parties in the plenary sessions (each zone illustrates the region corresponding to a particular party represented in the Parliament).

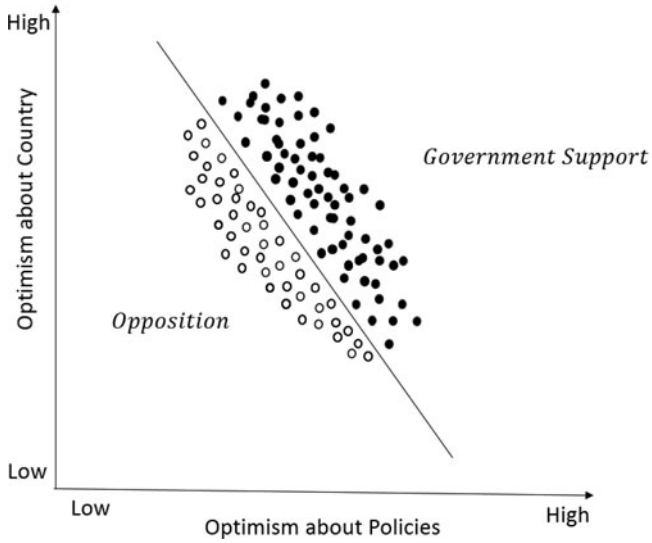

Figure 3. A more aggregate view of MP from a classification perspective and with regards to the relationships between Governments and Parliaments.

meaning of given categorical variables, so that classification and supervised learning help us also to get a better understanding of Parliaments. It may be appropriate to remember then that within the scope of supervised learning one is faced with a set of examples, including several independent variables defined in the X space, together with the corresponding known classes those examples belong to, usually expressed by means of a categorical variable, Y.

Under the context of Parliaments, it derives almost automatically from their democratic representative nature the fact they are organized according to an obvious number of categories that correspond to the parties with elected MP. This also implies that the categorical variable that corresponds to such parties, or coalitions, becomes critical for a proper understanding of behaviors. In the case of Portugal, this variable is defined as $Y = \{CDS, PSD, PS, CDU, BE\}$, with this set containing the acronyms of the five Parliamentary Groups that were elected. The relevance of such a categorical variable becomes quite clear even from the point of view of the physical location of the several MP in the plenary sessions of the Portuguese Parliament (Figure 2), with well-defined regions of seats corresponding to the different parties. That being the case, a very simple "Nearest Neighbor" classifier can predict exactly, except for the seats located in the borderlines between parties, to which party a particular MP belongs.

At a more aggregate level, and deriving also from my experience across two legislatures, one in the opposition and the other supporting the Government, one can consider and understand that there are at any given time two main groups of parties, and the corresponding MP, as defined in the Y space. This turn, they correspond to more simple and aggregate views, with $Y = \{Opposition, Government\ Support\}$. This binary variable is easy to understand. Having tried both values for this variable myself, under two different legislatures, I felt clearly how that corresponds to quite different mindsets and attitudes. But even more so at certain critical moments, like the one that corresponds to the annual public budget discussion and approval at the National Parliament. In such moments, it becomes evident, after some feature extraction, that optimistic versus pessimistic views, opinions and comments made by MP about how the country is doing and the quality of government public policies or choices allow for a conceptual statistical classification framework to be adopted (Figure 3). A clear definition can be made for the corresponding linear discriminant function that separates MP from these two classes. Under such a framework, the relationships, interactions and negotiations of Government with the Parliament and its parties can be seen, from a statistical thinking paradigm, as being focused on the definition and management of the corresponding discriminant functions. In order to assure that they lead to a combination of parties translated into having more than half of the MP supporting the budget or other initiatives presented by the Government for Parliament approval. This also implies, under this specific symbolic and conceptual representation

(Figure 3), that the line separating both categories of MP must be such that, in the case of Portugal, at least 116 MP lie on the upper right half-plane defined by it. Government relationships with the Parliament do, therefore, under this perspective, correspond to an appropriate management of this, sometimes unstable, line slope. From one year to the other, or one vote to the other, there is always the risk of having certain parties or MP move from one to the other side of this classification. Continuous negotiations between Governments, Parliaments, and their MP are, in this context, nothing other than critical discriminant function definitions, with nonlinearities coming into place especially when certain votes are needed to come up with a majority to support certain controversial proposals. To include additional less usual sub-clusters of MP above the line that separates "No" from "Yes" positions, classification "algorithms" can become somewhat more complex, with possibly a set of nonlinear and dynamic discriminant functions that need to be defined and managed. For that purpose, other classification or statistical tools, such as nearest neighbor, support vector machines, classification trees or case-based reasoning methodologies can also be used for deeper learning and understanding of this process of interface management between Governments and Parliaments.

Keeping the same overview of parliamentary compositions as consisting mainly of two categories of MP depending on whether parties are in the opposition or supporting Governments, a Bayesian view allows us to understand why lots of divergent positions are the daily routine in a Parliament. Large consensus can also be obtained, from time to time, but that is not the case most of the time. The underlying nature for such an aggressive and almost universal split of positions, from a Bayesian perspective, can be seen as deriving from the existence of rather strong prior odds. These either depend on the simplified opposition/government support statistical categories of MP, or can come also from the strong ideological roots associated with the several parliamentary parties. To illustrate this point, in Table 2 we describe a hypothetical situation showing that even proposals coming from a very good Government are very unlikely to receive approval or be supported by MP belonging to an opposition party. This very simple Bayesian view lets us understand better why in general discussions in the Parliament tend to be of a quite strong contradictory or even aggressive nature. Because for driving MP positions, when they speak under that role, Bayes theorem comes into place and dictates their mutual behaviors, corresponding in almost every occasion to divergent views or opinions over the same reality. Different opinions and conclusions will be typically extracted from it, and justified using namely contradictory arguments, data interpretations, or even sometimes, if needed, somewhat "creative" data analysis, as we will exemplify later on.

Table 2. A hypothetical Bayesian analysis example that explains why Parliamentary divergences are so common.

Assume that a very good Government is running the country, and as such the probability of a proposal coming from it to be good for the nation is quite high, and let´s call such a probability
$P(A) = 0.80$
Let $P(B)$ represent the likelihood of having a MP that stands in opposition to the Government, to vote in favor of a proposal coming from the Government, which is quite low, given the corresponding prior odds. Let's call such a probability
$P(B) = 0.03$
Finally, consider also the probability of being under the presence of a good proposal coming from the Government, given that it was approved by opposition MP, which needs to be particularly high. Let's call such a probability
$P(A/B) = 0.99$
Therefore, the final likelihood of having any MP from opposition to support even very good Government proposals ends up being rather small, as empirical evidence shows but the Bayes´ theorem helps to understand, since we then would get
$P(B/A) = [P(A/B) \times P(B)]/P(A) = (0.99 \times 0.03)/0.80 = 0.037$

Now, that we have used supervised learning to help us understand Parliaments, it is important to stress that after a certain period of time acting as a MP (in my case on the order of two years), one can see ourselves moving quite easily from supervised to unsupervised classification modes. And we can find other MP clustering approaches, not based only upon the parties connected with each particular MP. As opposed to supervised learning, within the scope of unsupervised learning, one is faced just with a set of examples, including a number of independent variables in the X space, but without any information regarding the corresponding classes those examples belong to. Under unsupervised learning contexts, these classes, usually expressed by means of a categorical variable, Y, have to be derived just from the data made available, namely by looking at similarities and distances computed from features found in the X space.

With appropriate feature extraction, it then becomes possible to guess what party a particular MP does belong to, just by observing the contents of what they tend to say, when and how they do it. Having additional knowledge available, built from experience (for me after four years), positions of MP became more predictable. That is, I could make good estimates about

what is going to be stated or guess by anticipation the positions that are about to come from most MP. I could therefore reach good prediction capabilities, within acceptable error margins, beforehand about what a particular MP is going to say. This may also well mean that I was already reaching a saturation point in the learning curve connected with my cumulative experience as a MP. When this does happen, and one can classify, just by listening, almost all MP into the corresponding parties and main underlying motivations, it becomes self-evident that a transition has been made from supervised to unsupervised learning stages of parliamentary analysis.

Once that stage of parliamentary experience and exposure has been reached, other less obvious categories of MP, from a more global and nonpartisan perspective, can derive from a conceptual cluster analysis. Under an unsupervised learning mode, features can be extracted that may result in the creation of other classes of MP, not strictly based upon the parties they belong to. For instance, clusters of MP from different parties can be defined and identified according to the predominant driving forces that determine their behaviors. Such predominant driving forces can be associated with what in economics are usually known as being individual utility functions, similar to each other across groups of MP that belong to several parties. From my own analysis, made after around two years of experience in the Parliament, this may essentially lead to the consideration of a conceptual model comprising six clusters (from A–F in Figure 4), composed of MP having in common features that can be illustrated on a plane (Figure 4). That can be done according to the following two dimensions of main individual motivations and previous life experiences, as expressed taking into account: (i) whether they are individuals with their own professional careers or have mostly been always politicians; and (ii) their dominant motivations and goals being either to get exposure and visibility (media driven), follow very strictly party ideological rules and instructions (ideology driven), or mostly trying to identify, address, and solve the country's main structural challenges and problems (society driven). These two dimensions, taken together, result in a total number of six possible MP clusters (such as the one corresponding to MP that are simultaneously "Media Driven" and "Professional Politicians"), with strong similarities in the way MP in each of such clusters do behave and in the roles they play in the Parliament. Still referring to Figure 4, the adoption of quality-based or statistical thinking approaches does seem to be much easier to achieve when we are dealing with clusters C and F (the natural allies for breeding this type of culture in the Parliament), and particularly more difficult with regards to cluster A (MP in this cluster will require more pedagogical effort to understand the benefits from such paradigms to be adopted in the political arena). This type of cluster analysis is quite helpful in terms of letting you know who can be your quickest allies, across different parties, in promoting more evidence-based and country centered parliamentary decisions or initiatives. You can also use this conceptual framework to understand, across parties, what types of utility functions are connected with several particular groups of MP. Therefore, quite often a proper and deeper understanding of similarities and dissimilarities across MP, not based strictly upon the parties they represent, can help in coming up with better approaches, alliances for supporting certain initiatives, and ways to understand Parliaments. Rather than the

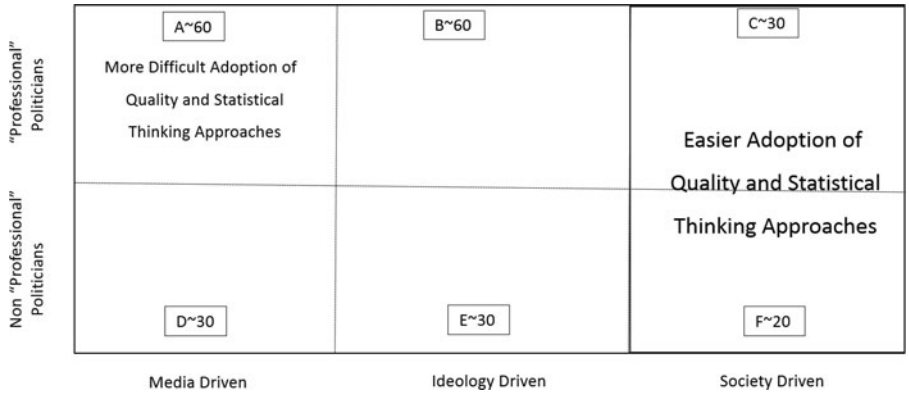

Figure 4. A conceptual cluster analysis of MP and their closeness to the adoption of quality-based and statistical thinking approaches, together with a rough subjective idea for the number of MP mostly associated with each cluster.

ones you get by using only the more obvious but sometimes shallow party-based view of MP and clusters of MP.

Uses and misuses of data in Parliaments

Parliaments are data intensive organizations, where both quantitative and qualitative information is conveyed daily. This is not to mention the large amounts of documents it does exchange with other international entities, such as the European Parliament, EUROSTAT (statistical office of the European Union), or OECD (Organisation for Economic Co-operation and Development). But the Bayesian view in parliaments often leads different parties and their MP to come up with opposed interpretations over the same basic reality, facts and data sets. According to their own utility functions, and depending on whether they are supporting the Government or stand in its opposition, MP take certain positions or arguments, sometimes regardless of what the corresponding data, evidences or realities happen to be. This sets the ground for quite peculiar, creative and sometimes strongly biased approaches for data (mis)handling. Some of the data interpretations and (mis)understandings commonly found in parliaments can be better understood under the following key "empirical principles", built from my own experience: (i) *If you torture the data long enough, they will confess to almost anything you want*; (ii) *If you take enough partial derivatives, any country will show good or bad news*; and (iii) *If you keep stratifying samples, a leading (good or bad) position can always be found.*

Therefore, Parliaments spend a lot of time discussing numbers. These discussions are even more prevalent when countries are facing times of crisis, as it happened to be the case for Portugal in recent years. But such a discussion is almost always contaminated with strong partisan biases that tend to search for good or bad news (un)related with data, often manipulated or distorted. We will next explore further such types of parliamentary data (mis)handling through a couple of examples.

For instance, consider the results obtained for Portugal in terms of innovation, according to the annual European Innovation Scoreboard results (European Commission 2016). The evolution of the SII-Summary Innovation Index (Figure 5), based upon which country innovation rankings are produced, both for Portugal and the European Union (EU) average, does show positive progress. But not much convergence

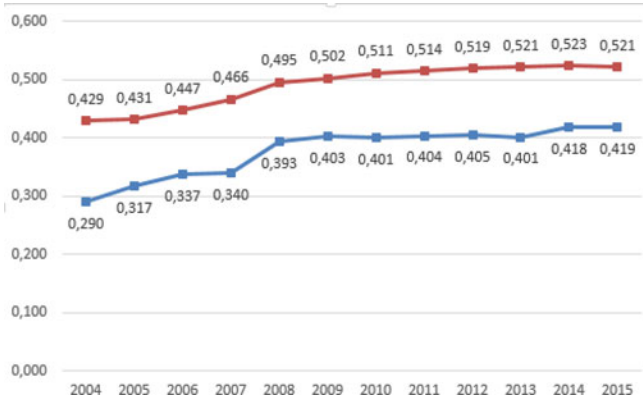

Figure 5. Evolution of the European Innovation Scoreboard results corresponding to the Summary Innovation Index for both Portugal (lower curve) and the European Union average (upper curve).

was achieved, since the gap between Portugal and the EU average remains, in 2015, similar to the one found back in 2004. However, if you are a MP looking for some data-based arguments to say positive things about national innovation policies, and their results, you can always claim that Portugal has been improving. If you need to go one step further, and want to find a "good" comparative performance, after several attempts, it is almost always possible to come up with numbers to support such an argument. For instance, by stating that Portugal was the EU country with the largest annual SII increase (over 8%) between 2006 and 2010. If you are "lucky" enough, just by forcing this alternative view over the same data set, headline news may even end up saying that Portugal is the country with the best innovation performance in the EU (not mentioning that what is at stake is just the first derivative average over 2006 and 2010). In the political arena, if the original data does not support your goals, one can always try to take derivatives over a certain period of time, and maybe by doing so you will be able to get values closer to what you "need".

If one looks at the historical roots of statistics, it becomes evident that official statistics have always played a major role, as one of the key driving forces for the discipline in government. The noble aim of official statistics is to provide politicians with appropriate data to support their decisions. As an example, the first census may have taken place as early as 3800 BC in Babylonia, to allow for proper estimation of food needs. Despite such a long experience with respect to the production of official statistics, as well as all the knowledge available in many well-known national or international statistical agencies, there is still room for significant

improvements in this area. Although some exceptions can be found, for instance regarding the PISA (Programme for International Student Assessment) OECD results (OECD, 2016), where statistically significant differences, or the absence of them, are pointed out, in most occasions national performances are presented without making the corresponding uncertainties available. This fact quite often encourages political speeches in the Parliament to comprise rather fuzzy statements about the evolution or performance of any given indicator. For instance, rather than saying with statistical significance that Portugal is above or stands below a certain average value for the EU countries, other less meaningful wordings are employed. For instance, if the national performance for a given indicator stands below the average, but the goal of a certain MP is to provide an optimistic view, he may rather say, in a vague way, that the value of this indicator for Portugal is staying "aligned" with international performances, without ever defining clearly what the fuzzy wording "staying aligned with an average" is supposed to mean. Although the message that is conveyed, without any sound statistical evidence, corresponds to trying to make others believe that we are performing at the level of the other countries, even if sounder evidence may point otherwise.

Furthermore, and even though no measurement uncertainties are usually taken under consideration, very strict rules, with "zero tolerances", are often assumed by politicians, national and international agencies. For instance, under the scope of the EU there are penalties associated with State Members with a public deficit (Public Revenues – Public Expenses) equal or above 3% of GDP. This threshold is assumed without considering measurement uncertainties corresponding to public deficit computations or even questioning the rationale for such a strict threshold of 3% to be applied. That being the case, many celebrations, criticisms or discussions end up having no sound statistical basis. As may well happen around an increase or decrease of 0.1% on the country unemployment rates, if the corresponding measurement uncertainty happens to be on the order of let's say 0.2%.

On a similar note, a recent and interesting article (Nolan, Perla, and Provost 2016) performed a proper statistical process control analysis regarding the U.S. quarterly GDP growth rates ranging from 1996–2016. A characterization of common causes of variation over this variable does show that in the corresponding time series of values only one observation stands outside control limits, in the fourth quarter of 2008. However, even a well-known specialized newspaper, after looking into these numbers in a less qualified way and without a proper understanding of this variable variation, came up with quarterly comments for this indicator such as the following ones, from 2013–2016: (i) "Economy Shows Signs of Gearing Up"; (ii) "Growth Rebound Stokes Fed Debate"; and (iii) "GDP Growth is Disappointing". As the authors of the above article appropriately do point out, "the economic losses associated with the misinterpreted variation in quarter-to-quarter data include the consequences of actions taken by individuals and institutions based on nonexistent trends such as potentially raising or lowering the U.S. interest rate, which carries profound economic implications", with costs to societies that are many times on the order of several billions of euros.

All the above evidence shows that there seems to be a large room for bringing "metrology"- based ideas and concepts into the ways macroeconomic indicators are defined, presented, or discussed. Otherwise, many discussions or decisions may end up being meaningless. Therefore, further work needs to be regularly conducted in the field of official statistics, with additional insights regarding the proper handling and presentation, whenever appropriate, not just of data, but also of the corresponding uncertainties. As pointed out by Kenett and Shmueli (2016, 2017), there is a lot to be gained if official statistics reinforce their levels of "information quality", since they need to be appropriately defined and used to be useful, and "official statistics play a critical role" in the context of "increased availability of data sources and ubiquity of analytic technologies".

As was already mentioned, often many MP are "torturing" data, to come up with apparently good or bad news, depending upon their parties. One possible way for doing so is also to explore seasonal trends, without comparing homologous periods of time. For instance, if one looks into time series of unemployment rate values for Portugal (Figure 6), it can be seen that they follow a volatile and unstable evolution across the years, with seasonality on top of that, expressed at several time scales. A more detailed analysis, obtained by zooming into monthly values, does point out that it is the case that unemployment rates go up from August to September, due to peak summer tourism activities in August as well as large numbers of people

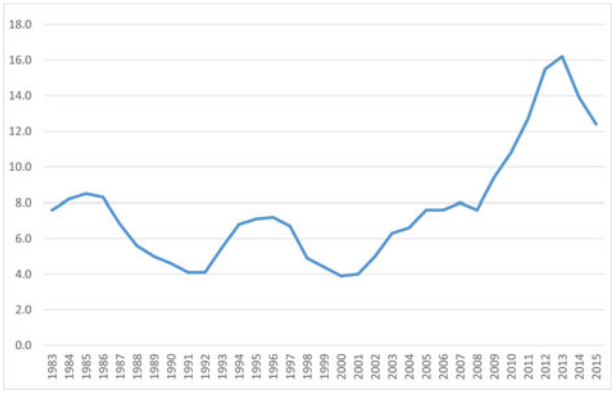

Figure 6. Evolution of unemployment rates (%) in Portugal across the years.

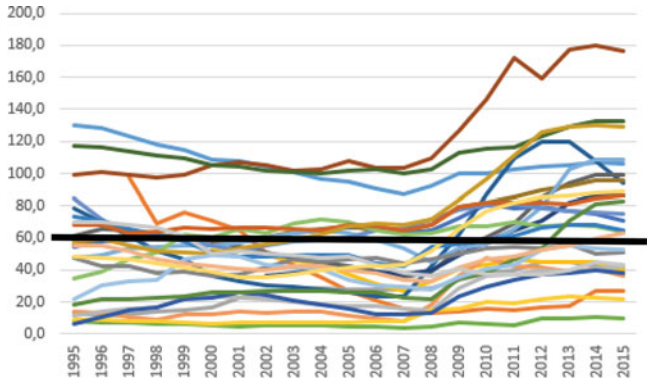

Figure 7. Evolution of public debt as a percentage of gross domestic product for the European Union countries, and its comparison with the assumed specification (staying below 60%).

(e.g., teachers) that start looking for jobs every September. By examining monthly unemployment data over a decade, between 2004 and 2014, in a total of 11 observations comprising the differences of unemployment rates between September and August, just one of them was negative, and only by 0.1%. This means that the probability of going up between these two months is above 0.9, and the corresponding hypothesis test does confirm that its average change value is positive, reflecting a structural seasonal phenomenon. However, it is not uncommon to get comments in this regard every Autumn, when MP do use these values to convey the message that unemployment is increasing in the country, as numbers "apparently" will support, and media coverage will repeat at least 90% of the time.

From a quality control point of view, sometimes specifications are defined and taken for granted in a very rigorous but somewhat strange way, at the national as well as the EU zones. One of such specifications corresponds to stating that all countries must have a level of public debt placed below 60% of the corresponding GDP. However, reality shows that the majority of the EU countries have been and still are clearly falling outside of these specification limits (Figure 7). With clear signs that we are dealing here with what might be called a process that is both highly unstable and incapable. In terms of process capabilities, we are thus on the opposite extreme of conformance, away from any sort of six sigma or reasonable AQL performance, with a 60% nonconformity rate. But rather than having this reality lead to a deeper discussion about the underlying process capabilities, deviations, tolerances, differentiation between special and common causes of variation, ways for improving performance or eventually revising specifications, such a lack of conformance is just repeated every year. No serious systemic approach or efforts to understand the underlying reasons that may explain such results are made by National or European politicians. And no one even tries to find out if the adoption of such specification limits in the first place makes sense as it stands (no sound statistical or scientific judgment has shown why it should be 60% as opposed to let´s say 50% or 70%). Another common strategy for coming up with "outstanding" good or bad performance in an indicator has to do with the way you stratify a given sample of data. For instance, an MP may state that "Portugal is the most competitive Ibero-American country in the EU with less than 40 million people", while another MP would say that "Portugal is the least competitive Ibero-American country in the EU with less than 40 million people". And both would in some sense "be right and not lying", since Portugal ends up being the only European Ibero-American country with less than 40 million people. If you stratify samples up to this point, you may reach a point where both MP may be right, although their observations end up being totally meaningless.

As a final remark, connected with all of the above examples and concerns about parliamentary data (mis)interpretation, it is important to take seriously how much societal improvement might be obtained if basic "data rights" were to become more protected from the now so common MP or media data misuses. Such data misuses may be conducted either on purpose or only due to the lack of an appropriate basic statistical training. There is much to be gained if some basic data analysis and fact-based thinking becomes more common among politicians, opinion makers, and journalists. That would result in better decision making and savings on the order of several billions of euros. To any

investment in this type of training efforts would correspond a tremendous societal net present value, as well as a very short-term payback time.

Achieving better results in Parliaments through quality and statistics

Given the kind of organizational culture that dominates Parliaments, as described above, one can find considerable room for more fact-based approaches, namely through the simple usage of some basic quality or exploratory data analysis tools. Here we provide a few examples of what that may represent. Quality and statistical thinking can therefore both help to understand and to improve the activities of Parliaments and MP, and this section tries to illustrate this second component.

Application of basic quality and statistical tools

An approach that I have adopted consists of comparing word counts over certain critical documents, to find out possible substantial semantic differences regarding some key issues, although many other more elaborate data analytics tools may also be employed for similar purposes. By doing so it is sometimes possible to uncover how different semantics are used to convey different messages, depending on the end users of such documents. Or how, as time goes by, reality checks lead to evolutions that make of what a Government actually does something quite different from what it expected to do and promised to do in the election campaigns. As a specific example of this type of comparison, we report here how two documents of a similar size (around 120 pages and 12,000 words each), and produced one just a couple of months after the other, do indeed have significant differences over some keywords counts, as evidenced by a simple statistical test comparing proportions (Figure 8). The first document corresponds to the Government Program, discussed in the Parliament, which broadly speaking states the priorities and main initiatives that a newly elected Government proposed for Portugal. It is mostly used for national purposes, at the beginning of a new legislature period of four years. A couple of months later, the same Government produced another document, mostly used for discussion and negotiations with the European Commission, also after discussion in the Parliament, the so-called Stability and Growth Program. This second document

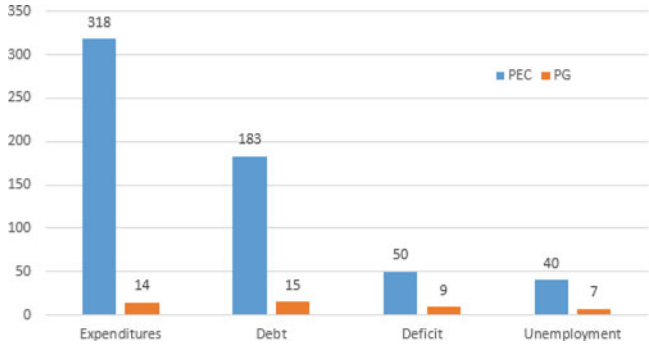

Figure 8. Word count comparison between the Government Program (PG, right columns) and the Stability and Growth Program (PEC, left columns).

is particularly relevant not for national but for international purposes, under the scope of the EU. Although both have similar scopes of analysis, dealing with the priorities and measures for Portugal in the forthcoming 3–4 years, basic text mining tools allow us to see that they seem to adopt quite different perspectives. By comparing through word counting certain keywords across both documents, we found out for instance that the Government Program does refer to public expenditure, debt, deficit, or unemployment much less often than the Stability and Growth Program, which was presented to the European Commission just a couple of months later. Apparently, this couple of months is enough for the Government running the country, after experiencing all the demands in place coming both from Portugal and the EU, to significantly change its perspectives, as reflected in these important strategic documents.

Nowadays, quite comprehensive text mining or analytic tools are available and can be applied to produce deeper studies conducted over different kinds of documents being written, discussed, and voted in Parliaments. It is desirable to use such tools to extract key features, main messages, similarities, and differences across them, this being also an interesting field for further applications to emerge, since the lives of Parliaments and MP are quite intensive in terms of the number of words being said or written daily.

In a similar two-way comparison mode, often there is much discussion regarding the so-called "left" vs. "right" dominant political ideologies and their impacts over societies. When one tries to find out such differences from a data perspective, sometimes they seem to be more of a myth than of a reality. For the sake of illustration, we can look at time longitudinal data,

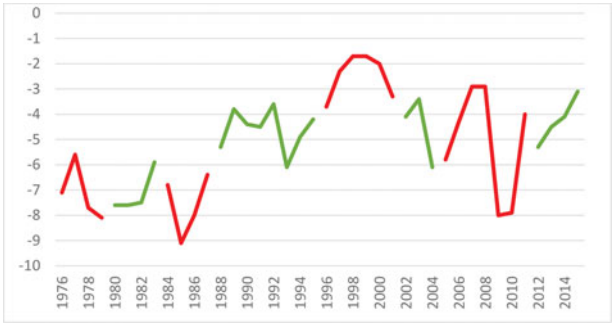

Figure 9. Time evolution for the Portuguese public deficit (as a % of gross domestic product), considering years with a left (dark) or right (light) majority in the Parliament.

regarding the evolution of the Portuguese public deficit (as a percentage of GDP), over a period of 40 years, ranging from 1976–2015. Then, we considered, for each of those years, whether there was mostly a majority of MP from left or right parties in the Portuguese Parliament (Figure 9). By looking at the corresponding boxplots and statistical tests, no significant differences of averages were found, although strong autocorrelation is present, possibly meaning that what really drives these values has more to do with other exogenous and structural factors, rather than with parliamentary compositions. This was also found to be the case for other macroeconomic indicators, where evident patterns could not be associated with the corresponding alternating Parliament compositions, as opposed to what one might expect just by listening to the common "left to right" exchanges of arguments and political opinions.

Basic quality tools, such as flowcharts or cause-and-effect diagrams, can also help to improve the conceptual clarity of laws being produced in Parliaments. As an example, when time came for discussing and making changes to the law addressing academic careers at public higher education institutions in Portugal, I decided to build a very simple one page visual scheme for better understanding all the possible situations under consideration, rather than relying only on the complex set of paragraphs and sentences spreading across many different pages and paragraphs of text in the law draft presented for discussion and analysis. After a while, it was rewarding to realize that most MP belonging to the education committee, where the discussion took place, were focusing arguments and suggesting changes around my simple one page flowchart summary of the contents, rather than over the somewhat tedious and rather confusing original text, full of articles, clauses, and sub-clauses. This also helped to assure that no forgotten situations or inconsistencies between different articles of the law remained in the final version. Curiously enough, some years later questions about the exact interpretation we did have in mind at the time were raised by national courts, and once again going back to the original flowchart provided a much easier and precise answer, as opposed to looking at the text that ended up being published as the final law approved by the Parliament.

I have no doubt that laws would become better, easier to understand, and able to avoid intrinsic contradictions or infinite loops around certain clauses if Parliaments were to adopt, whenever appropriate, the good habit of coming up with and publishing simple visual tools for explaining the situations and underlying logic that the text is supposed to follow. In doing so, amongst other advantages, one would also be able to improve legislation quality by: (i) screening and removing potential inconsistencies between articles or paragraphs; (ii) avoiding the possible existence of infinite loops, with clause remissions to a number of other clauses that may end up going back to a never ending labyrinth; (iii) supporting a much easier, well-focused, and constructive discussion about alternatives and changes amongst the different MP involved in the discussion; and (iv) making communication easier, leading also to laws that media and average citizens can understand, with their essentials being captured and presented by a simplified visual representation of contents.

In an analogous context, the use of a simple fish-bone diagram or application of the 5 consecutive whys principle can make a substantive difference between remaining at a shallow discussion or going after root cause analysis and getting into the real problems of any given society. As an illustration of what this may represent, under the context of a Parliament, we can start with a look over perception data collected from citizen surveys that point out that Portugal stands in a very low position (26^{th}) when one compares the overall life satisfaction average values obtained across the 28 EU countries. Rather than remaining at the tip of the iceberg, further exploration of one branch of a cause and effect diagram, combined with the 5 whys, does show that part of the underlying reasons for that to be the case have to do with high unemployment rates (above 10% for a number of years). This results from a lack of significant economic growth (less than 2.5% GDP

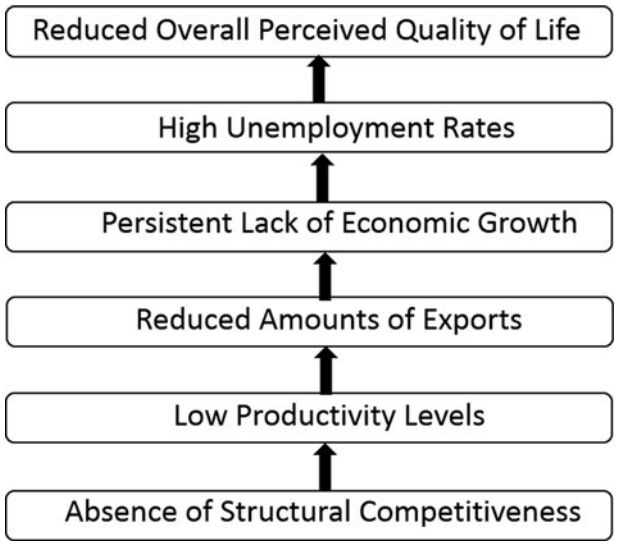

Figure 10. Going from the surface to the root cause analysis for some of the major societal challenges facing Portugal.

growth since 2001), which is due mainly to reduced amounts of exports taking place. Partially due to reduced productivity levels, discussions and measures were necessary to reinforce the national structural competitiveness (Figure 10). Once we reached this point, it was possible to set up a working group, within the scope of the Parliament Committee for Economic Affairs, to address ways of reinforcing sustainable, structural and resilient competitiveness in Portugal. This would not have been possible if the discussion had remained at the first level, without enough consecutive why´s being considered. Such a deeper view also allowed me to make what I believe to be the first time that a conceptual triple integral was presented in a parliamentary plenary session. It was aimed at explaining that such a competitiveness model does rely mostly on the cumulative effects over time, people, organizations, and territories of an integrand function that corresponds to the product of quality, innovation, and entrepreneurship, as being the major driving force for building a future of increased well-being in Portugal.

Fact-based analysis and the collapse of a major bank

An interesting example of parliamentary activity, to which I dedicated almost six full time months of my experience as MP, corresponds to the role for which I was elected, as rapporteur for a special Parliamentary Inquiry Commission. This commission was set up to study the circumstances that led to the bankruptcy, back in 2014, of BES, a private bank (the third largest one in the country, with two million customers and 800 branches), after 145 years of being in business. Its bankruptcy, among costs for the other banks, required a public investment of about 4 billion euros, to protect customers and lead to the creation of a new bank out of the proceeds.

At the beginning, almost everyone considered that filing a good final report, supported by different parties, was going to be an "impossible mission", due to the complexity of what was at stake and the time and resources available for conducting the work. There was a wide variety of ideological positions coming from the 24 MP belonging to this commission, chosen by the several parties represented in the Parliament, with different views about the financial world and the banking systems in general.

However, with strong dialog, some hard work, good project management and, most of all, an evidence-based approach (facts do not belong to any parties or do they have political ideologies attached, neither do they belong to any preconceived school of thought about banking activities), we produced a final report and conclusions that were recognized as being of high quality and received wide approval from all the MP that stood in the commission. That happened regardless of their parties and against most of the prior odds one might have about how the final report would be perceived by such a large variety of MP.

For preparing and writing the report, several tools or basic principles were quite valuable. These included: (i) appropriate "mass balances" that I had learned from Chemical Engineering classes of industrial stoichiometry (money could not disappear from my spreadsheets, in spite of all the movements performed across the world); (ii) order of magnitude simplifications and applications of the Pareto rule (to focus on the critical issues, and not so much on the details, in this case meaning that I had to get used to the idea that cash-flows of less than 100 million euros represented details, since they corresponded to less than 0.6% of the overall size of the problem, estimated in the end as being of around 18 billion euros); and (iii) significant data compression efforts (a final report of 416 pages, with 5.3 MB was produced from 55 auditions lasting almost 300 hr and corresponding to around 8,000 pages of written statements, as well as 50 GB of compiled documentation, therefore representing just 0.01% of the total information collected).

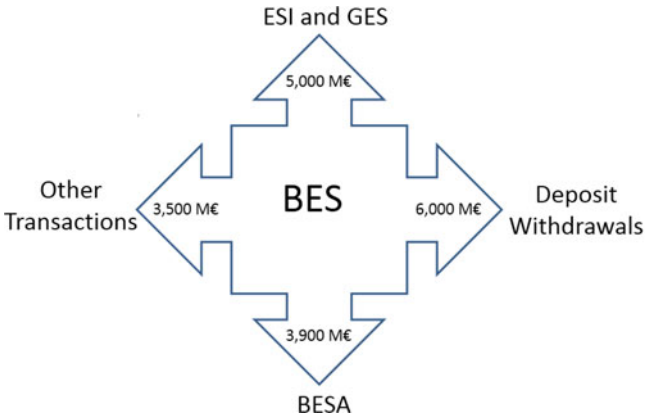

Figure 11. Simple visual representation of the main sources of problems and allocations of resources that resulted in the bankruptcy and resolution of a major bank in Portugal (BES), with values corresponding to millions of euros.

To handle such a demanding task, always looking for information quality (Kenett and Shmueli 2017), a main source of inspiration came from some of the well-known statistical learning paradigms. This came from George Box, who stated that continuous iterations between inductive and deductive steps (Box 2000) are critical for learning purposes and knowledge creation. Continuous creation and testing of hypotheses did take place at the same time that I always made an effort to adopt simple visual ways for presenting and representing information, knowledge or main conclusions. I used figures such as the one presented here (Figure 11), that shows the "perfect storm" the bank went through, leading to a drainage of resources on the order of 18.4 billion euros. This corresponds to half the amount of its deposits, resulting in the bankruptcy of the original bank. Such a simple visual explanation of what happened occupies just half a page, but to get here months of data-based screening, evaluation and inspiration were needed. Its size of 40 KB is equivalent to only 0.8 ppm of the total data made available to the inquiry parliamentary commission, but it illustrates clearly the following main categories of atypical money that did flow out of the bank: (i) 5 billion euros for supporting the non-financial part and the top of this large conglomerate group of companies, GES, with over 300 organizations operating in 50 countries and having 30,000 employees (Espírito Santo International [ESI] and GES); (ii) 3.9 billion euros related with the bank activities in Angola (Banco Espírito Santo Angola [BESA]); (iii) 3.5 billion euros regarding overvalued assets and some last minute transactions, just before the bank was forced to a resolution by the national central bank (other operations); and (iv) 6 billion euros of deposit withdrawals in the last month of BES operations (deposit withdrawals).

In the end, this systematic adoption of quality management and statistical thinking paradigms paid off. It comprised a well-defined approach to data compression, data-based learning, feature extraction, continuous hypothesis formulation, refinement or testing, and some "designed experiments", that included different rounds of questions made across auditions. All these were applied on an almost daily basis, together with quite rigorous quality planning, control, and improvement, complemented with strict project management practices. Risk analysis was also applied, taking into consideration the most important critical factors that needed to be kept under control, namely through traceability measures aimed at reducing the probability of having information leaks to take place while the report was being prepared. As opposed to what had always been the case in similar circumstances, it was possible to assure that the report conclusions and contents were presented and discussed first hand in the Parliament, without any sort of previous press related leaks of information. This systemic oriented work and mindset, taken very seriously from the first and to the last minute, allowed us to separate and remove several types of random or induced noises from the analysis. This allowed for extracting useful information from opinions and facts for coming up with knowledge, learnings and conclusions. That allowed us to explain what happened, and also for the commission to generate a strong set of 70 recommendations, aimed at improving the quality and levels of trust associated with the Portuguese banking system.

After all these efforts, and as a very rewarding outcome, not only was the report quality widely recognized by media and opinion makers, but all the facts identified were unanimously approved in the commission. Furthermore, there were no MP votes against the 70 recommendations made, and in the final overall voting of the report amongst the committee members, it received 13 yes votes, 1 abstention, and 1 no vote (from a single MP who stated that this had to do mainly with the fact that his party does believe that all banks should be state owned, and not so much with disagreements regarding report contents). There was a consensual voting pattern that is unusual under the context of the Portuguese Parliament, and even more so given the subject matter studied and discussed. I believe that this was

made possible mostly due to the data-based approach that was followed across this large-scale parliamentary project.

After the completion of this particularly intensive process, I did come up with "twelve commandments" for playing the role of rapporteur under the scope of a Parliamentary Inquiry Commission. Some of these commandments are strongly connected with quality and statistical-based principles. Here are three: (i) facts are indeed your best arguments, since they do not have attached any ideological biases, and are, by definition, nonpartisans; (ii) do not leave for tomorrow what you can report accurately today; and (iii) do not leave for tomorrow the hypothesis that you can build or test today.

Overall, this peculiar and unique experience, given its special nature and dimension, has helped me to reinforce, by an additional empirical real-life experience, that with quality and statistical-based mindsets and approaches even in parliaments the almost impossible can indeed be made possible.

Macroquality issues handled with statistical tools: The European quality scoreboard

The role played as MP has also reinforced my strong beliefs that statistics and quality management can play a significant role in addressing particularly complex societal challenges, and this inspired me now to get involved and lead some projects in this area, as will be shown in this section. Indeed, applications of quality and statistical approaches can be quite helpful across a large variety of scales, both in terms of time and space (Saraiva 2015). There are many interesting opportunities that can be addressed at what I have coined as being "macroquality" challenges (Figure 12). To address these challenges, we define problems and come up with solutions or conclusions at the country or international levels and over time scales that may take years to complete. To complement the more common usages of quality engineering to products, processes or organizations, there is therefore room for dealing with quality at the level of municipalities, regions, countries, or even the world. As we move to a "Big Data" world, full of large, complex, and unstructured problems (Dibenedetto, Hoerl, and Snee 2014), as well as to the so-called VUCA (standing for Volatility, Uncertainty, Complexity, and Ambiguity) environment, there are plenty of opportunities for quality professionals to

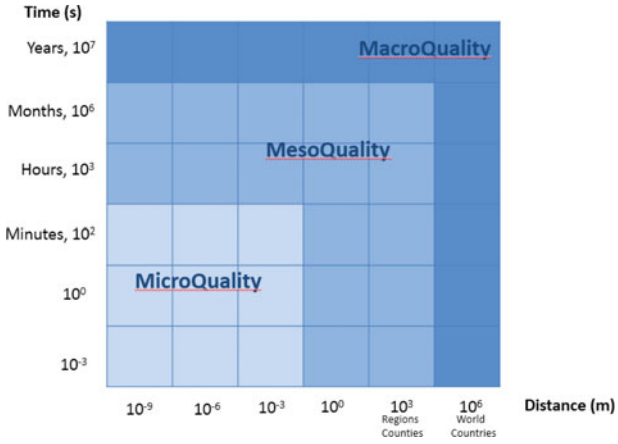

Figure 12. A multiscale taxonomy of quality related opportunities, challenges, and approaches.

get involved in a wide variety of such "macroquality" challenges and problems. Just as there has always been room for dealing with both "microeconomics" and "macroeconomics", multiple scales can be adopted for dealing with quality problems, going all the way from "nanoquality" to the less common but extremely relevant and impactful scale of "macroquality" (Saraiva 2015).

In the past, I have pioneered and conducted a number of projects in this field of macroquality, including namely: (i) statistical analysis and modeling of ISO 9001 diffusion across regions and countries (Saraiva and Duarte 2003); (ii) adaptation and application of regional barometers to measure, from an integrated perspective, the levels of quality achieved at given territories (Orey and Saraiva 2016); (iii) usage of structural equation models to understand employee satisfaction at the level of particular organizations, but also leading to national aggregated results (www.onrh.org); and (iv) application of structural equation models to define perspectives and priorities for the definition and implementation of a national quality policy (Saraiva et al. 2010).

The intensive experience lived in the Parliament, between 2009 and 2015, has reinforced my interests in these types of analysis, particularly reflected also now in the World State of Quality (WSQ) project, that I have been developing since 2016. After I realized that international views and rankings are currently performed over many domains, such as Competitiveness, Innovation, or Entrepreneurship, but were still missing in terms of overall evaluation and comparisons of performance regarding quality. As a first stage for this project, we have defined and implemented our

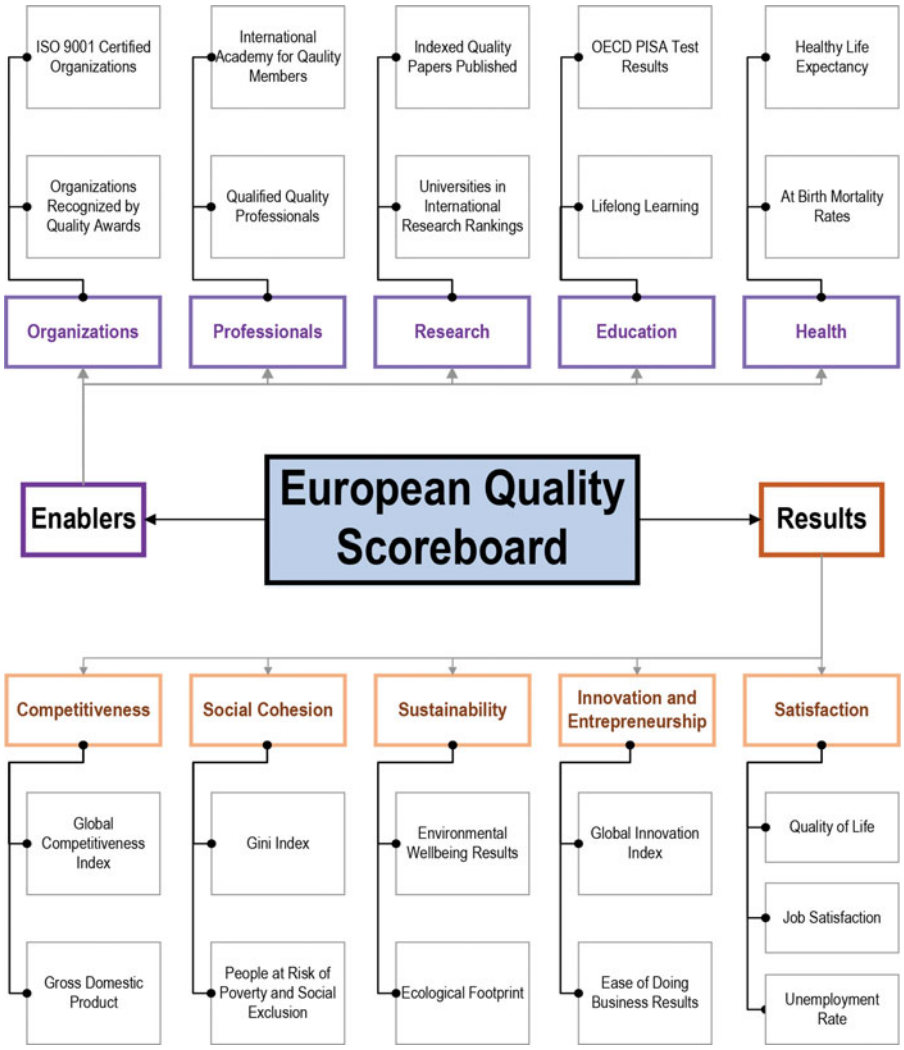

Figure 13. European Quality Scoreboard underlying structure of dimensions and indicators.

framework to cover the 28 EU countries (European Quality Scoreboard-EQS), and are now expanding results to additional countries.

After data collection for each of these 28 countries, corresponding to a set of 10 dimensions and 21 quality-related indicators (Figure 13), and their statistical treatment, we produced the first conclusions deriving from this project.

The complete sets of results, including the final country quality profiles and rankings, are provided in the available WSQ report (Saraiva et al. 2017). Thus, here we will focus our attention on some of the conclusions derived from the multivariate statistical analysis that was performed over this "macroquality" data set, comprising 28×21 values. In particular, we would like to underline the following topics and results achieved: (i) when four groups of seven countries are considered, taking into account the overall final scores (OEQS) obtained, we can easily see the diversity of geographical performance that is obtained (Figure 14); (ii) such a variety of situations is curious and important to understand the quality realities in the EU geography, where no single country outperforms or underperforms the others over all the indicators considered, each of them presenting relative strengths and weaknesses; (iii) first applying Principal Component Analysis (PCA) to the data set, we can see that one principal component explains approximately 45.5% of the total data variation, meaning that there is a clear and strong correlation backbone linking a large portion of the indicators; (iv) all countries seem to fit under a common underlying PCA model, with the possible exception of Luxembourg, which falls somewhat outside the corresponding zone of confidence (Figure 15); (v) several types of clustering techniques were also conducted over the same sets of data, aimed at identifying relative

	Country	EQS Score
1	Finland	7.85
2	Austria	7.97
3	Sweden	8.33
4	Netherlands	8.45
5	Denmark	9.05
6	United Kingdom	10.20
7	Luxembourg	10.86
8	Germany	10.90
9	Ireland	11.18
10	Slovenia	11.45
11	Czech Republic	12.76
12	Belgium	12.99
13	France	13.10
14	Spain	14.24
15	Portugal	14.40
16	Estonia	14.52
17	Malta	14.65
18	Italy	15.65
19	Slovakia	17.00
20	Poland	17.67
21	Hungary	17.72
22	Cyprus	17.83
23	Romania	18.03
24	Lithuania	18.61
25	Latvia	19.39
26	Croatia	19.65
27	Greece	19.96
28	Bulgaria	21.97

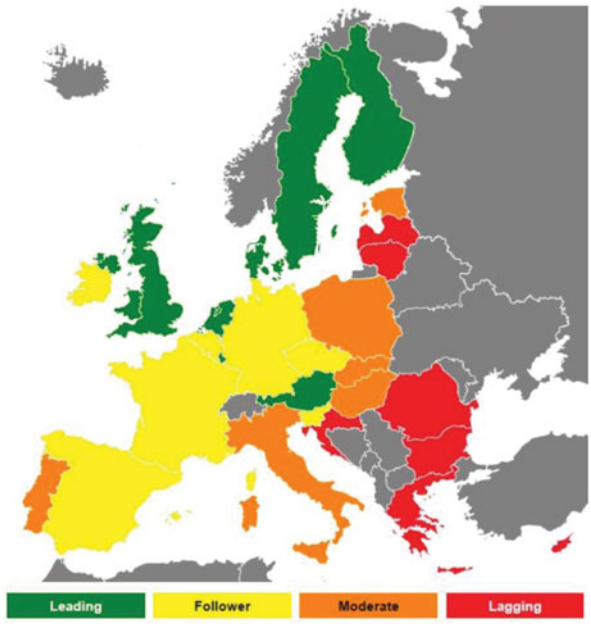

Figure 14. Country performances according to European Quality Scoreboard 2016 scores.

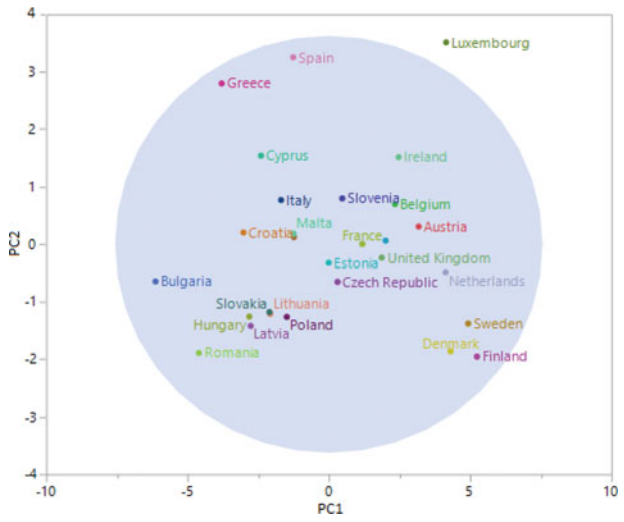

Figure 15. PCA analysis: scores plot for the first two PCs, with 95% confidence level perimeter.

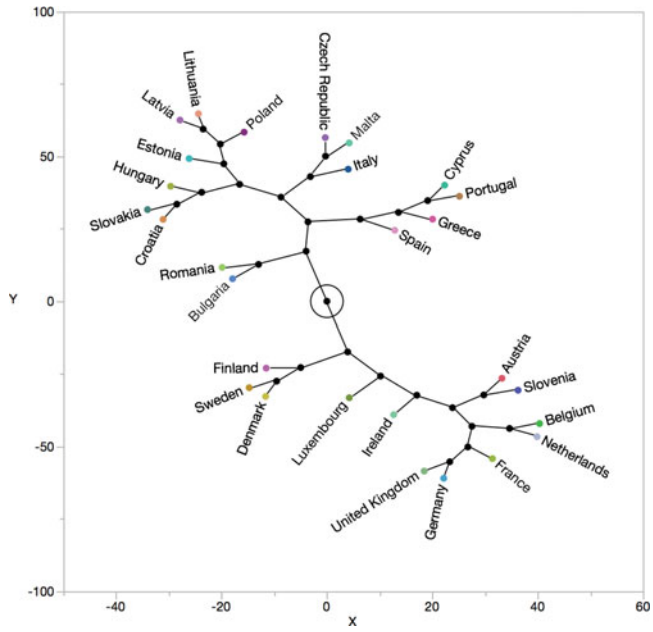

Figure 16. Cluster constellation plot with European Quality Scoreboard positions for the 28 European Union countries.

similarities and closeness (or lack of them) across the 28 EU countries, according to the positions they occupy in the multidimensional space that corresponds to the 21 EQS indicators considered; (vi) as a first result, portrayed here under the format of the corresponding constellation plot (Figure 16), one can see from an hierarchical clustering perspective that in the last agglomeration there are two major groups of countries, one corresponding to 12 countries (that include all of our EQS Leading countries, and 5 of the top 7 Follower countries), and the other with the remaining 16 countries (corresponding to all of the EQS Lagging and Moderate countries, as well as 2 of the Follower countries); (vii) a more detailed geographical analysis also points to the apparent existence of closely connected countries, having similar quality profiles according to the set of 21 indicators considered, pointing towards a diversity of quality cultures, journeys or "ways" for its interpretation and implementation; (viii) such is the case for the "Scandinavian quality way" (Finland, Sweden and Denmark), the "Southern Europe quality way" (Spain, Greece, Portugal and Cyprus) or the "Eastern

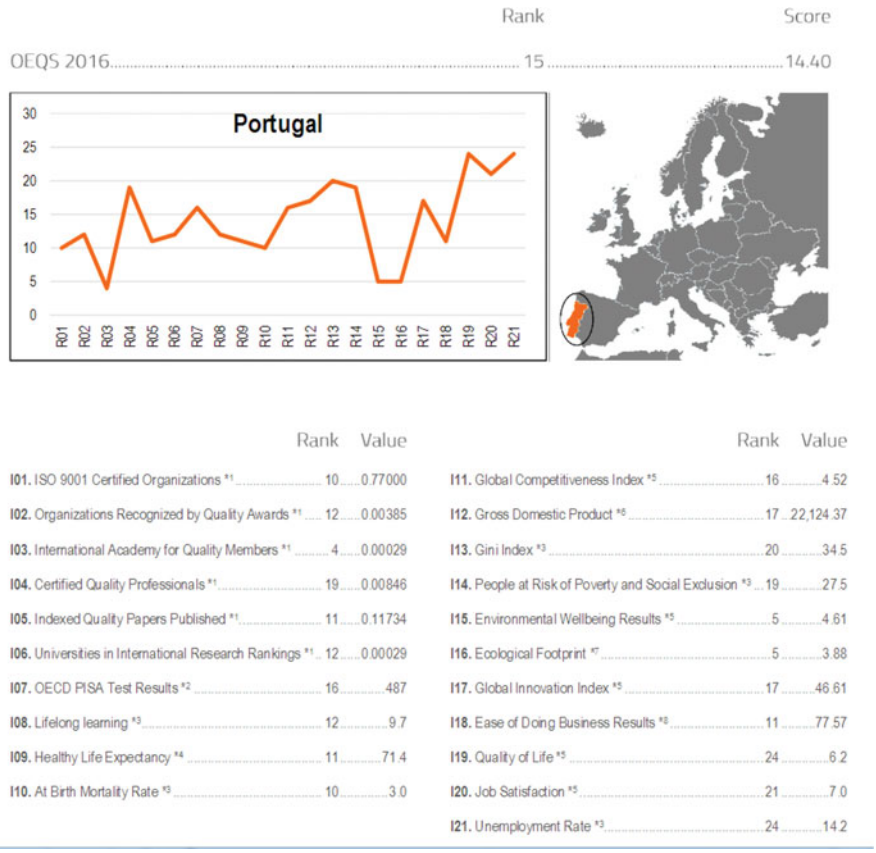

Figure 17. Summary of quality profile and 2016 European Quality Scoreboard results for Portugal.

Europe quality way" (Croatia, Slovakia, Hungary, Estonia, Latvia, Lithuania, and Poland); and (ix) given its special role as an indicator of overall quality performance, when we considered perceived Quality of Life as a dependent variable, and performed several types of multiple regression analysis, using the other variables as independent variables, 60% of the response variability ended up being captured by only two predictors, one having to do with the country´s wealth generation capabilities (such as the Global Competitiveness Index), and the other one with dispersion of wealth distribution (such as the Gini Coefficient). This last result is particularly curious also from a basic quality and statistical thinking point of view: to achieve high overall perceived macroquality, we must take into account not only that one has to be able to come up with high average wealth performances, but also to achieve the smallest possible dispersion of wealth distributions across the corresponding populations.

At a countrywide level, we were also able to come up with 28 national quality profiles (Figure 17), thus leading to the identification of strengths and weaknesses (typically in the extreme quartiles of ranking positions achieved by any particular country), that may help in the definition of appropriate customized targets, priorities, and policies for further promotion of quality in the different nations studied.

Many more details can be provided, as we try to move to include non-EU countries in future work, but we hope that this sample of evidence-based results is able to illustrate the potential of proper statistical analysis conducted over macroquality related data. This is an area where we believe that many other interesting quality engineering opportunities and challenges can be found.

Some final thoughts and conclusions

As many of us did learn from George Box, there is a lot to expect from statistical and quality thinking as we face several huge societal challenges in this data intensive 21st century (e.g., Industry 4.0, Internet of Things, Big Data, Machine Learning). Even more so, since we live in an ever stronger knowledge environment: "the quality movement will undergo healthy changes over time and may even be called by different names; however, insofar as it is a catalyst to the generation of new knowledge, it is here to stay" (Box 2000).

For that to be the case, in the communities of quality or statistical thinking and practice it is important to communicate in simple terms, especially when addressing the quite relevant world of politics and politicians. At first it may be better not to mention most of the technical jargon that lies behind what is being done. We may get along better if we rather focus on evidence-based outcomes, presented under user friendly formats, aimed at helping to achieve significant improvements at multiple levels. If we are able to do so, increased recognition for our contributions will be obtained, and more statisticians or quality professionals will also be asked to lead efforts related with new societal realities, rather than relying for that purpose just on other "new professions" (such as data science and data scientists). Simple visual tools conveying powerful messages are particularly helpful under this type of context, that being also part of the reason why traditional Shewhart Statistical Process Control charts have survived now for almost 100 years (they convey important messages about process behavior in a format that is easy to follow, explain, and understand).

At another level, it may also be important to create more harmonized and larger consensus about ways to solve problems under structured approaches, rather than adding methodologies one after the other, with different names but having many things in common. Statisticians and quality professionals need also to get better prepared for dealing with trends in the quality world, which is going increasingly to: (i) deal with adaptation and flexibility; (ii) become mandatory for survival and results driven; and (iii) be strongly related with innovation, data analysis, products, and operations management (Sampaio and Saraiva 2016).

If one takes these issues into consideration, then bridges with Parliaments, politics, and politicians will also become easier to establish. After having won some initial challenges, interactions between quality or statistical thinking and Parliaments may then result in further and deeper learning opportunities. Just as a simple example of the kind of pedagogic efforts that may need to be conducted, while in the Parliament I often remembered some of the lessons learnt from the Deming funnel experiment. It conveys a simple but yet powerful way to understand the adverse effects of tampering with a process, when excessive changes or adjustments are made without a proper understanding or characterization of the underlying variation. Therefore, such adjustments just make things even worse. If more and more politicians happen to know some of these concepts, this may end up having strong implications for societal evolution and improvement. Namely, this may lead to removing the common situation that occurs every time there is a new President, Prime-Minister, or Parliament. The new government immediately introduces changes over the changes made by the previous politicians in charge. This is done without any sound analysis of previous results, trends, or even confirming if enough time went by for enabling such results evaluations to be made over systems with large time-dependent dynamics (such as education, social security, healthcare, or taxes). Nevertheless, new policies are forced into place. In many occasions, from a statistical thinking paradigm, Parliaments and countries by doing so are therefore paying the price of useless additional societal variation due to the lack of stable policies. That is the result of the "natural" tendency to reverse previous decisions and manage decisions under a constant wish to change just for the sake establishing ownership of a policy instead of improving a policy's efficiency or efficacy. Sounder understandings of variation on the side of politicians might prevent or at least reduce the large intensity of permanent over-adjustments that often take place in Parliaments and Governments, one election after the other. Appreciating variation could also provide a good example of how statistical thinking and training may help change the way politics and policies are discussed, conducted, and managed. It would therefore be worthwhile thinking about how to build pedagogical tools and initiatives to reinforce training in this area for MP and other politicians.

Winston Churchill once said, "Democracy is the worst form of government, except for all the others", and Parliaments are, of course, a central piece with regards to how societies live under democratic paradigms. They carry out very relevant and useful missions, as my rewarding experience as MP has also shown. That experience was long enough to realize also how Parliaments and politics can be better understood and sometimes improved by adopting quality, evidence-based, and statistical thinking approaches.

For that to become a reality, I have no question, drawing from my own journey, that having MP with a strong professional background in quality or statistics is something worthwhile pursuing and considering, for mutual benefits. If that starts to be the case more often, I do believe that this will reinforce the contributions of

quality and statistics to the progress of our societies and in the world. I hope that this article illustrated, through a sample of specific examples and thoughts, that this is indeed the case. My examples were presented as a humble tribute to all that I learned and will keep learning from Stu Hunter and other inspiring top-level statisticians across the planet.

Therefore, the community of quality engineers and readers of this journal can provide in several ways powerful contributions to reinforce, at multiple levels, quality democracy. We live in a world where it is becoming increasingly evident how important it is for quality democracy not to be taken for granted. Since it rather needs to be constantly preserved, promoted, nurtured, and continuously improved.

About the author

Pedro Manuel Saraiva received his Diploma from the University of Coimbra, Portugal (1987) and then a Ph.D. degree from MIT (1993), both in Chemical Engineering. He has been Assistant (1993–2001), Associate (2001–2010), and Full (2010-) Professor at the Department of Chemical Engineering at the University of Coimbra. Member of the Portuguese National Parliament (2009–2012 and 2012–2015). Pro-Rector of the University of Coimbra (2003–2004 and 2005–2006). Vice-Rector of the University of Coimbra (2007–2009) in charge of Quality, Innovation, Entrepreneurhsip, and Technology Transfer. Member of the Board of BIOCANT-Biotechnology Science Park (2005–2009). Advisor of the President of Portugal for Higher Education (2006–2009). President of the Regional Agency for the Development of the Centro of Portugal (2004–2005 and 2012–2014). First receiver of the Feigenbaum Award (1998) by ASQ (American Society for Quality). Associate (2010–2015) and then Full Member (since 2015) of the International Academy for Quality (IAQ) where he is also the Chair of the Think Tank for Quality in Education (since 2016). Main research activities and interests in the fields of Process Systems Engineering, Data Analysis, Innovation, Entrepreneurship and Quality Management, with several books and over 100 indexed papers published. Has supervised around 10 Ph.D. students that have already completed their thesis. Currently belongs to the Process Chemometrics Laboratory of the Process Systems Engineering research group at the Department of Chemical Engineering, University of Coimbra, Portugal.

Acknowledgments

I want to thank the organizers, scientific committee, and participants of the Fifth Stu Hunter Research Conference and particularly to Ronald Does, for first suggesting my participation in it as a speaker. I also acknowledge the help, support, additional important comments, or incentives provided by Marco Reis, the article discussants for tremendous insights and encouragement (Bo Bergman, Geoff Vining, and Ronald Does), Ron Kenett, and Bradley Jones (who also invited me to be make a similar presentation at the 2017 European Discovery Summit, where additional insights were also received and allowed to further improve the contents of this article).

ORCID

Pedro Manuel Saraiva http://orcid.org/0000-0002-4465-4597

References

Box, G., and N. Draper. 1998. *Evolutionary operation – a statistical method for process improvement*. New York: Wiley.

Box, G. 2000. *Box on quality and discovery, with design, control and robustness*, ed. G. Tiao, S. Bisgaard, D. Hill, D. Pena and S. Stigler. New York: Wiley.

Box, G., W. Hunter, and J. Hunter. 2005. *Statistics for experimenters – an introduction to design, data analysis and model building*. New York: Wiley.

Conti, T. 2003. *A strategic view of organizational stakeholders,* chapter in *Quality into the 21st Century – perspectives on quality and competitiveness for sustainable performance*, ed. T. Conti, Y. Kondo, G. Watson. Milwaukee, WI: ASQ Quality Press.

Dibenedetto, A., R. Hoerl, and D. Snee. 2014. Quality Progress. *Jun* 2014:50–53.

European Commission. 2016. European Innovation Scoreboard (http://ec.europa.eu/growth/industry/innovation/facts-figures/scoreboards_pt)

Hoerl, R., and R. Snee. 2012. *Statistical thinking – improving business performance*. 2nd Ed. New York: John Wiley & Sons.

Kenett, R., and G. Shmueli. 2016. From quality to information quality in official statistics. *Journal of Official Statistics* 32 (4):1–19.

Kenett, R., and G. Shmueli. 2017. *Information quality – the potential of data and analytics to generate knowledge*. Chichester, West Sussex: Wiley.

Nolan, T., R. Perla, and L. Provost. 2016. Understanding variation: Correctly assessing variation is fundamental to sound decisions. *Quality Progress* Nov. 2016:28–37.

OECD. 2016. PISA 2015 Results. Paris: OECD Publishing.

Orey, J., and P. Saraiva. 2016. Regional excellence measurement. Poster presented at the 60th EOQ (European Organization for Quality) Congress, Helsinki, Finland.

Sampaio, P., and P. Saraiva. 2016. *Quality in the 21st century: Refreshing perspectives from ASQ Feigenbaum medal winners*. Switzerland: Springer.

Saraiva, P., and B. Duarte. 2003. ISO 9000: Some statistical results for a worldwide phenomenon. *Total Quality Management & Business Excellence* 14 (10):1169–78.

Saraiva, P. 2015. Multiscale Quality. Oral presentation made at the World Quality Forum, promoted by the International Academy for Quality, Budapest, Hungary.

Saraiva, P., J. Orey, P. Sampaio, M. Reis, C. Cardoso, J. Pinheiro, and L. Tomé. 2010. O Futuro da Qualidade em Portugal (The Future of Quality in Portugal). Book in Portuguese published by Associação Portuguesa para a Qualidade (Portuguese Association for Quality), Lisbon.

Saraiva, P., P. Sampaio, C. Cubo, M. Reis, and J. Orey. 2017. *World state of quality: 2016 european quality scoreboard*. Portugal: University of Minho and University of Coimbra. (http://wsq.dps.uminho.pt/)

Tukey, J. 1977. *Exploratory data analysis*. Boston, MA: Addison-Wesley.

Zonnenshain, A., E. Naveh, and A. Halevy. 1998. A survey of nonquality to a Ntion´s economy: The Israeli Experience. *Quality Progress* October 1998:93–97.