CHAPTER 1

Introduction

Sometimes what people do not say is just as revealing as what they do say. We have experienced this over the past few years. Friends, family and fellow academics have asked us what our book will be about and we have answered 'How politicians manipulate statistics'. What never came next is revealing. No one expressed amazement, saying words to the effect: 'What an odd topic – surely you don't suppose that politicians manipulate statistics? You're not going to find anything to write about.' Everyone seemed to assume that we would find plenty of stuff to occupy our attention. They all presumed that politicians do not invariably tell the truth, the whole truth and nothing but the truth. Public opinion polls back up this belief. For the past forty years in Britain, where most of these conversations took place, Ipsos has been conducting annual surveys into trustworthiness. The 2023 survey revealed that the two least trusted professions were 'politicians generally' and 'government ministers'. Less than 10 per cent of the population trusted politicians – a new record low.[1] Surveys conducted across the world also report that people are becoming increasingly distrustful of politicians and governments.[2]

If people tend to generally distrust what politicians say, then it is likely that they will distrust what they say about statistics. There has been long-term distrust of statistics, as expressed by the well-known saying that dates

[1] IPSOS (2023), 'Trust in politicians reaches its lowest score in 40 years', www.ipsos.com/en-uk/ipsos-trust-in-professions-veracity-index-2023. The survey uses the Ipsos Veracity Index.

[2] The Edelman Trust Barometer for 2024 surveyed twenty-eight countries and reported that in seventeen of those, more people distrusted their governments than trusted them. The most trusted governments tended to be non-democratic: Edelman Barometer (2024), '2024 Edelman Trust Barometer: Global Report', www.edelman.com/sites/g/files/aatuss191/files/2024-01/2024%20Edelman%20Trust%20Barometer%20Global%20Report_0.pdf, p. 42. Two years previously the Edelman Barometer had reported that 76 per cent of all the respondents surveyed in twenty-seven countries indicated that they worried about fake news: Edelman Barometer (2022), '2022 Edelman Trust Barometer: Global Report', www.edelman.com/sites/g/files/aatuss191/files/2022-01/2022%20Edelman%20Trust%20Barometer%20FINAL_Jan25.pdf, p. 8. See also: Esteban Ortiz-Ospina and Max Roser (2016), *Trust*, Our World in Data, https://ourworldindata.org/trust.

from the latter part of the late nineteenth century: 'There are lies, damned lies and statistics'. In his excellent book *Damned Lies and Statistics*, Joel Best writes that some people have attributed the saying to Mark Twain and others to Benjamin Disraeli.[3] Convincing evidence about the saying's originator does not exist.[4] However, its fame and easily understood humour taps into a widespread suspicion that statisticians and the politicians who quote them can make numbers mean whatever they want them to mean. This suspicion of statistics is certainly not confined to the nineteenth century. Following its publication in 1954, Darrell Huff's book *How to Lie with Statistics* was for a good number of years the best-selling work on statistics. In his introductory chapter, Huff described his book as 'a sort of primer in ways of using statistics to deceive'.[5]

Consequently, there are reasons to think that people might become extra suspicious when they hear a politician spouting numbers, especially if those numbers do not match their own experiences or their long-held beliefs. To mention again the conversations about our book's topic, no one said to us: 'No, no, no! Think what you like about politicians, but when it comes to numbers, politicians are pretty straightforward.' Once more, the absence of words spoke loudly. And we should say here, in the opening part of our introduction, that we have not struggled to find examples of politicians manipulating statistics. In fact, we have had to be selective, choosing which examples to analyse in detail and excluding, for reasons of space, a far greater number of possibilities.

Rhetoric of Numbers

There is one thing to emphasise right at the start. This may be a book about numbers and statistics, but it is not a mathematical book. There are no complicated formulae or pages full of numbers and technical symbols that only those with mathematical training can understand. Our medium is words, and words constitute much of our topic. That might seem strange given that we are writing about statistics, but statistics are much more than numbers: they need words.

The sort of statistics that we are dealing with are official statistics, and it is not possible to have official statistics that are purely numerical. To be

[3] Joel Best (2001), *Damned Lies and Statistics*, Berkeley: University of California Press.
[4] Anyone interested should definitely look at the website established by Peter M. Lee. More than twenty possible sources for the saying are listed: Peter M. Lee (2012), 'Lies, damned lies, and statistics', www.york.ac.uk/depts/maths/histstat/lies.htm.
[5] Darrell Huff (1982/1954), *How to Lie with Statistics*, New York: W.W. Norton, p. 11.

both meaningfully official and meaningfully statistical, the numbers must be linked to words. The official statistician will attempt to calculate social phenomena such as the level of unemployment, the rate of crime and the gross national product. These social phenomena – the things to be measured – are identified by words. Statisticians must always label the numbers that they calculate.

There may be different ways of making the calculations and these different calculations may produce different numbers, often with the same label attached. There again, statisticians, economists and politicians may dispute what is the most appropriate number for each label. Does unemployment currently stand at 5.5 per cent or at 9.0 per cent? If we measure it this way it is 5.5, but if we measure it that way it is 9.0. So, which is the better way of calculating the rate of unemployment? Although this type of dispute is a numerical one, it cannot be conducted entirely in numbers for the disputants will phrase their arguments in words: 'Your measure does not count those people who have stopped looking for work' to be countered by 'But yours includes those who do not want to look for work'. Such disputes are often about the links between numbers and their verbal labels. In Chapter 4, we examine how Donald Trump, not the most sophisticated mathematician in the world, entered into such a debate about the rate of unemployment in the United States.

Modern life is saturated in numbers. In the United Kingdom, we can expect to encounter them when we watch the news, read a newspaper or surf current affairs on the internet. According to one survey, almost three-quarters of respondents reported seeing statistics on the news several times a week.[6] One study, which for a month examined broadcast and online news platforms in the UK, calculated that 22 per cent of all news items contained at least one statistical reference.[7] Often journalists use statistics to give their stories 'colour'; in consequence, they tend not to present their statistics in any detail or depth.[8] There are some topics that are rarely presented with statistics – such as stories about celebrities. Stories about terrorist attacks are unlikely to be accompanied by comparative statistics, especially those which might indicate that the risk of being killed in a

[6] Sarah Butt, Benjamin Swannell and Alisa Pathania (2022), *Public Confidence in Official Statistics 2021*, National Centre for Social Research, https://natcen.ac.uk/our-research/research/public-confidence-in-official-statistics/.

[7] Stephen Cushion, Justin Lewis and Robert Callaghan (2017), 'Data journalism, impartiality and statistical claims', *Journalism Practice*, 11, 1198–215.

[8] Brendan T. Lawson (2022), 'Re-imagining the quantitative–qualitative relationship through "colouring" and "anchoring"', *Journalism*, 23, 1736–50.

terrorist attack is likely to be considerably lower than the risk of being killed in a road accident.[9] Had such comparative figures been presented, the dramatic 'colour' of the terror story would have become that bit drabber.

When journalists and politicians use numbers, they generally give those numbers rhetorical meaning. We will be examining Trump and his views on US unemployment figures. When he quoted the official figures before he became president, he left no doubt that he thought the figures far, far too low to be trusted; and he suggested that the 'real' figures were far, far higher (much higher than our imaginary 9.0 per cent illustration). By his choice of words, he gave political and emotional meaning to the numbers.

Giving rhetorical meaning to numbers is certainly not peculiar to Trump or to his particular view of the world. When the news breaks about a new treatment for a disease such as cancer, the reporters may cite the success rate for the new treatment. The statistic on its own would not convey much because the audience will not know the rates for other treatments. So, the reporters will indicate whether this new number should be understood as being a low or high success rate. Some researchers call these verbal guides to cited numbers 'quantification rhetoric'.[10] David Spiegelhalter, the distinguished British statistician, emphasises that we have to speak on behalf of our numbers, for the numbers cannot speak for themselves.[11]

Some numbers are used as if they are words that contain their own rhetorical signals. Campaigners for a cause which they believe has been ignored by the public will frequently cite numbers to indicate that the problem is widespread. Again, the number on its own is not sufficient to accomplish their purpose; they must indicate that the number represents a figure that is far too high for comfort.[12] Often precise numbers or statistics are not available because the problem has been ignored. Then campaigners might use intentionally imprecise numbers, expressed in words not

[9] Cushion et al. (2017).

[10] The concept of 'quantification rhetoric' describes how numbers are presented with verbal indicators suggesting whether the numbers should be understood as high or low, good or bad, worrying or comforting, etc.: Jonathan Potter, Margaret Wetherell and Andrew Chitty (1991), 'Quantification rhetoric – cancer on television', *Discourse & Society*, 2, 333–65. For announcements of miracle medical cures and why their exaggerated statistics should be distrusted, see: Tim Harford (2021), *How to Make the World Add Up*, London: Bridge Street Press. See also: Daniel Libertz (2018), 'Framed for lying: statistics as in/artistic proof', *Res Rhetorica*, 5(4), www.resrhetorica.com/index.php/RR/article/view/289/177.

[11] David Spiegelhalter (2020), *The Art of Statistics*, Harmondsworth: Penguin.

[12] Joel Best (2001), 'Promoting bad statistics', *Society*, 38(3), 10–15.

numerals, that rhetorically convey hugeness – i.e. 'thousands and thousands' or 'millions' or 'billions'.[13] These sorts of semi-magical, round numbers are often used by those who wish to play what Teun van Dijk calls the 'numbers game' in relation to immigration: 'thousands upon thousands are pouring in every day'.[14]

We will be looking at examples of politicians using numbers rhetorically throughout the book, but using numbers rhetorically does not necessarily mean that the numbers are being manipulated. In Chapter 10 we examine in detail a British government minister of health who used rhetoric to suggest that he had successfully organised a large number of tests for Covid-19. It is not the minister's use of rhetoric that makes the example so suitable for being discussed here, but that the numbers, which he was claiming to be so large, had been manipulated to appear larger than they were. This illustrates a general point. If you want to understand how politicians manipulate numbers for their own purposes, then you have to look at what they are doing with their numbers rhetorically; and what they are doing may lead you towards recognising some of the ways and means of manipulating numbers and audiences.

Yet, to understand how and when manipulation occurs, you must go a bit further than noting the rhetorical uses of numbers. That is why we start our book with two historical chapters, which chart the rise of our two key words: 'statistics' and 'manipulation'. Both words emerged in the late eighteenth and early nineteenth centuries. And both have had more than a single meaning. By following the words through time, we can compare statistical manipulation today with the past. We will be seeing how the early originators of numerical statistics were social campaigners. They used their numbers rhetorically, even creating entirely new ways to depict numbers graphically to achieve maximum effect. This does not mean that they manipulated their numbers, certainly not in the ways that have developed in recent years. The history of the very phrase 'statistical manipulation' contains a story in itself.

[13] Michael Billig (2021), 'Uses of precise numbers and semi-magical round numbers in political discourse about Covid-19: examples from the government of the United Kingdom', *Discourse & Society*, 32, 542–58.

[14] Teun A. van Dijk (2018), 'Discourse and migration', in Ricard Zapata-Barrero and Evren Yalaz (eds.), *Qualitative Research in European Migration Studies*, OAPEN https://library.oapen.org/handle/20.500.12657/29754?show=full. For analyses of the use of numbers in immigration rhetoric, see: Charlotte Taylor (2021), 'Metaphors of migration over time', *Discourse & Society*, 32, 463–81; Simon Goodman and Steve Kirkwood, (2019), 'Political and media discourses about integrating refugees in the UK', *European Journal of Social Psychology*, 49, 1456–70.

Manipulating Public Statistics and Public Statisticians

Recently, there have been a number of well-written, insightful books about public statistics, written by authors such as Joel Best, Tim Harford, Brendan Lawson, Reimund Mink, Georgina Sturge and David Spiegelhalter.[15] These fine books are addressed to general readers as well as specialists. They are public-spirited works, because their authors aim to correct misconceptions about statistics and to educate their readers to spot poor, misleading data. We shall be drawing on these works in the course of our book.

One might wonder whether there is a need for yet another book about the uses and misuses of public statistics. However, the topic is broad: it covers different countries, different policies, different types of government and different ways of misusing data. The authors whom we have just mentioned have their own slants and special interests. Although they might sometimes use similar examples and draw insights from each other, none of their books is redundant – each makes an original contribution. If you want to understand the sociology of statistics, turn to Best; if your interest is in the imprecision of statistics, then Spiegelhalter is your statistician; should you want to know how particular numbers can become publicly important for a time before fading away, then you must read Lawson on the lives of numbers; Mink will tell you about the organisation of public statistics around the world; and Sturge will give you insightful anecdotes about British politicians and the poor data that they sometimes use, whether by choice or necessity. Joel Best demonstrates how seemingly inexhaustible the topic is. Having published his *Damned Lies and Statistics*, he found he had a whole new range of stories to tell and examples to offer: hence *More Damned Lies and Statistics*.[16] And still the politicians of the world provide new material for the analysts to write about and the public to distrust.

Our book has its own particular emphases. Our primary emphasis is on manipulation, and that is why we have an early chapter on the history of the word. Generally, we are interested in language and this lies at the root of our interest in the rhetorical usage of numbers. In the past, we have studied how politicians perform actions with words. If their actions are not

[15] Best, *Damned Lies and Statistics*, 2001; Harford (2021); Brendan T. Lawson (2023), *The Life of a Number*, Bristol: Bristol University Press; Reimund Mink (2023), *Official Statistics – a Plaything of Politics?* Cham: Springer; Spiegelhalter (2020); Georgina Sturge (2022), *Bad Data*, London: Bridge Street Press.

[16] Joel Best (2012), *More Damned Lies and Statistics*, Berkeley: University of California Press.

straightforward, as is often the case, then you must look very closely at what they are saying in order to spot the gaps between their words and apparent actions.[17] As a result, when we analyse how politicians manipulate statistics, we will be paying attention to some very small aspects of language. To manipulate numbers deviously, politicians and others need to use language subtly.[18]

In addition, we will be stressing that statistical manipulation is not a simple action, something that many theorists of manipulation have tended to overlook. Government ministers are not typically skilled statisticians, so if they want particular numbers for political reasons, they will have to get their statisticians to produce those numbers for them. There can be a double-barrelled process of manipulation, as politicians manipulate statisticians to manipulate the statistics. This is a recurring theme throughout our analytic chapters.

We have different sorts of analytic chapters. There are three biographical chapters, each of which looks at the way a single politician has manipulated statistics over the course of their career. Each of our chosen three has held, still holds or dreams of holding senior office in the future, and each comes from a different country: Donald Trump from the United States, Gérald Darmanin from France and Boris Johnson from the United Kingdom. There are similarities and differences between the three in their ways of misusing and degrading statistics.

We also have a chapter which looks at incidents of statistical manipulation in four very different countries: the autocracies of Stalin's Russia and contemporary China; and democracies in Argentina and Greece. We relate these specific incidents to what we offer as a schematic guide to how politicians manipulate statistics. This guide points to three stages, two of which typically occur away from public sight. We stress that incidents of statistical manipulation do not necessarily go through all the three stages, nor necessarily in the 'right' order – not even the incidents that we discuss

[17] Our previous work on political language includes: Michael Billig and Cristina Marinho (2017), *The Politics and Rhetoric of Commemoration*, London: Bloomsbury; Michael Billig and Cristina Marinho (2019), 'Literal and metaphorical silences in rhetoric: examples from the celebration of the 1974 Revolution in the Portuguese parliament', in Amy Jo Murray and Kevin Durrheim (eds.), *Qualitative Studies of Silence*, Cambridge: Cambridge University Press, pp. 21–37.

[18] We will be offering examples throughout the book. We have written about the importance of social scientists using examples: Michael Billig (2019), *More Examples, Less Theory*, Cambridge: Cambridge University Press; Michael Billig and Cristina Marinho (2022), 'Using examples to misrepresent the world', in Jeanne Fahnestock and Randy A. Harris (eds.), *Routledge Handbook of Persuasive Language*, New York: Routledge, pp. 113–28.

in the chapter. That is why the schema is offered only as a guide, not as a theory.

Then there are chapters that reflect the second distinctive concern of our book. We are not just interested in examining how the politicians might manipulate statistics and statisticians, but also how such manipulation might be exposed and opposed, at least in democracies. Our interest is not primarily theoretical. We are not saying 'If only society could be organised in a better way, then we would be able to trust what our politicians tell us about numbers'. Our interest in combatting manipulation is empirical. We look at the regulatory authorities which have been legally established in Britain and France to oversee the standards of official statistics. We devote a chapter to each authority, examining how it was established and how it operates. Most importantly, we examine the strategies that each authority uses in its mission to counter bad official statistics. We note something surprising – at least it was something that surprised us. Both authorities combat statistical manipulation by virtually never using the word 'manipulation' (whether the French or English word) in their official documents.

In the biographical chapters on Darmanin and Johnson, we concentrate on both politicians' clashes with their respective national statistical agencies. We closely observe how these politicians used their official powers to ignore the recommendations of their agencies and how they both tried to devalue statistics. When the agencies confront powerful politicians they are hampered by a lack of power. They can only recommend, not enforce, good statistical practice. Yet, the agencies tended to come out on top in their confrontations with Darmanin and Johnson, and we examine how and why this happened.

The final analytic chapter digs down deeply into a single incident that occurred in Britain during Johnson's time as prime minister. The incident did not directly involve Johnson. We see in detail how the authority's strategy was initially to be diplomatic, but when the minister ignored its advice, it switched to using more direct language. This brought immediate results, but without a total victory. A manipulative politician can indeed be a slippery customer.

As far as we are aware, the actions of the British and French statistical regulators have not previously been examined, certainly not in the depth that we have studied them. Just as the biographical chapters tend to reveal villains, so the later chapters on the authorities throw up some rather surprising, bureaucratic heroes. We discuss the implications of this in the final chapter, and we briefly say why it is so difficult, perhaps virtually impossible, to establish similar authorities in the United States.

Lastly a word about our writing. We have tried to avoid using technical language, whether that of statistics or of linguistics. Nor do we use the technical language of social psychology, the academic discipline in which we both trained. Social psychology is probably the discipline that has most studied the topic of persuasion; and it has produced shelf-loads of specialist words. Instead of reaching out for technical terms, many of which would be unsuitable for our purposes, we have tried to explain the processes of statistical manipulation and counter-manipulation as simply as we can. Although psychologists might look for, and fail to find, their own terminology in our writing, we consider this to be a psychological work. Each time we examine an example of political manipulation, we come back to the motives of the political manipulator and what they are seeking to gain by their manipulations. We try to look at their actions, words and motives as directly as we can, rather than through the intervening lens of a favourite theory.[19]

We believe that statistical manipulation is a worrying feature of the modern world and one that potentially corrupts democracy. In this book, we occasionally use irony. This is not because we want to get a cheap laugh or because we think our topic is full of fun. Quite the reverse. In our view, political leaders who are so confident of themselves that they have few scruples about inventing their own numbers pose a dangerous problem. It is because such leaders take themselves and their invented untruths seriously that they deserve to be mocked.

[19] Marie Jahoda, the great social psychologist of a previous era, advocated that psychologists should try to look at the world directly, even naively. See: Marie Jahoda (1989), 'Why a non-reductionist social psychology is almost too difficult to be tackled, but too fascinating to be left alone', *British Journal of Social Psychology*, 28, 71–8. See: Billig (2020), chapter 8, for an appreciation of Jahoda. See also: Alexandra Rutherford, Rhoda Unger and Frances Cherry (2011), 'Reclaiming SPSSI's sociological past: Marie Jahoda and the immersion tradition in social psychology', *Journal of Social Issues*, 67, 42–58.

CHAPTER 2

The Rise of Statistics

Our main topic – politicians manipulating statistics – is very much part of present times. Official statistics have grown in importance during the past fifty years, especially in political life. Today politicians are constantly ready to quote statistics as rhetorical weapons to boost their own case and to weaken the arguments of their opponents. Opinion polls show that the public's trust in politics has declined significantly and that politicians are now widely suspected of manipulating the statistics that they cite. Thus, our problem seems to be very much a modern one.

In later chapters we will be showing some high-profile politicians manipulating statistics to their own advantage. But before this, a bit of history might give some perspective. In this chapter we look back to the early days when statistics, as we understand them today, were not yet called 'statistics'. We look backwards for two basic reasons. The first is a general point: most explorations in the social sciences benefit from having a historical dimension. If our thinking remains locked within our own times, then we run the risk of imagining that the world has always been the same; or conversely that our own times are absolutely without precedent.

A historical perspective will help us to pinpoint why the political manipulation of statistics has developed as it has. We will be making a point that might seem obvious once it has been made: namely, when governments acquire the power to control statistics, they will try to use that power. In this chapter we look backwards to the first half of the nineteenth century, when public statistics grew dramatically but politicians had little power over their day-to-day production, analysis and dissemination.

There is a second reason for looking backwards. It is to meet some fascinating individuals. We will encounter the Belgian polymath Adolphe Quetelet, a key figure in the early social sciences, and William Farr, the notable but often forgotten founder of medical statistics in the United Kingdom. It is easy to assume that statistics provides a very masculine way of perceiving the social world. However, the most famous woman in