In November 2010, those of us who had bought the hardback edition of *The Spirit Level* were given the opportunity to spend a further £10.99 on a revised paperback which promised ‘a new chapter responding to their critics’. More of a counter-attack than a defence, this new chapter sadly addressed very few of the substantive points made in this book. Nevertheless, as a crystallisation of Wilkinson and Pickett’s response to criticism in the previous six months, it remains of interest and deserves comment.

The main themes of the new postscript are that (a) *The Spirit Level* is a summary of well-established peer-reviewed evidence which is considered uncontroversial in academic circles, and (b) despite not being politically motivated themselves, Wilkinson and Pickett have been the victims of a co-ordinated attack by right-wing ideologues.

Neither of these claims stands up against the facts.

**Consensus?**

A key plank in Wilkinson and Pickett’s defence is the notion that they are merely informing the general public about issues that have long since been agreed upon by the academic
community. Since most people will never read any of the studies in the field, this has been largely successful as a public relations exercise, but it is a gross distortion.

It also represents something of a U-turn for the two epidemiologists. Wilkinson and Pickett’s sudden insistence that they are reflecting the scientific consensus is at odds with the way they promoted their book when it was first published. In an interview with the couple in March 2009, a journalist from the *Guardian* reported that: “For a while, Wilkinson and Pickett wondered if the correlations were too good to be true. The links were so strong, they almost couldn’t believe no one had spotted them before.”¹ This could just about be excused as shoddy journalism were it not for Wilkinson and Pickett’s eagerness to take the credit for what they described as their “discoveries” in *The Spirit Level* itself. The book’s preface leaves the reader in little doubt that these discoveries are genuinely new and exciting, hence the comparisons with Joseph Lister and Louis Pasteur. “The reason why the picture we present has not been put together until now is probably that much of the data has only become available in recent years,” they write, adding that “it could only have been a matter of time before someone came up with findings like ours.”²

The fact of the matter, as discussed in Chapter 1 of this book, is that there has been a large amount of research into health and inequality spanning three decades. Richard Wilkinson has been a key figure in this field, but his views do not represent the consensus. Not could they, since there is emphatically no consensus. The only honest way to describe the state of the literature on health and inequality is to say that it is mixed and conflicting. Researchers are broadly divided into three groups. There are those, like Wilkinson, who believe that there is a solid correlation between inequality and health and that this represents a causal link. There are those who believe there is a statistical correlation but that it is not causal, and there are those who believe there is no link at all.
Only the first of these positions is reflected in *The Spirit Level*, and the reader is given the false impression that academics have firmly established that inequality leads to poor health. Wilkinson and Pickett accuse their critics of not being familiar with the “extensive research literature”, but it is precisely because we *are* familiar with it that we know how grievously it is misrepresented in *The Spirit Level*. In the new postscript, Wilkinson and Pickett say “there are around 200 papers in peer-reviewed academic journals testing the relationship between income inequality and health”.\(^3\) ‘Testing’ is the key word. There is no hint of how many of these studies have not found a relationship, nor of how many found a statistical relationship but concluded that it was not causal.

Their source for the ‘200 studies’ claim is, as so often in the book, one of their own papers. This article, from 2006, assessed 169 results from 155 studies on inequality and health (plus some other studies related to violence). By Wilkinson and Picket’s own reckoning, 88 of these were supportive of their theory (including six of their own studies) while 81 were either unsupportive or provided conflicting results.\(^4\)

Wilkinson and Pickett stress that many peer-reviewed articles have offered at least partial support to their theories about income inequality. This is true—at least in the area of health—just as it is true that there are many peer-reviewed articles that beg to differ. Hence the long-running academic debate about inequality which *The Spirit Level* has done much to inflame but little to resolve. This debate has already been discussed in Chapter 1, but it might be useful to quote from some other researchers in the field:

All along, however, critical questions were being asked about the quality and interpretation of the data. In an early exchange, serious criticisms of the selection of countries, the quality of the data, and the lack of control for confounding in [Wilkinson’s] *BMJ* paper of 1992 were only half countered. Although many aspects of this debate are still unresolved, it has recently
become clear that the findings of that paper were an artifact of the selection of countries.\textsuperscript{5}

\textit{—British Medical Journal} editorial, 2002

This paper extends previous studies by examining long time series for 12 of the world’s richest countries rather than one or two. Our findings are consistent with those of Deaton and Paxson (2001) and Lynch et al. (2004b), not with those of Wilkinson (1989, 1996) or Sen (1999). In our preferred specifications we find only small and statistically insignificant relationships between income inequality and mortality. This holds true regardless of whether we measure mortality using life expectancy at birth, infant mortality, homicide, or suicide.\textsuperscript{6}

\textit{—Leigh & Jencks, 2007}

The study found limited evidence of an association between income inequality and worse self-rated health in Britain, which was greatest among those with the lowest individual income levels. As regions with the highest income inequality were also the most urban, these findings may be attributable to characteristics of cities rather than income inequality. The variation in this association with the choice of income inequality measure also highlights the difficulty of studying income distributions using summary measures of income inequality.\textsuperscript{7}

\textit{—Weich et al., 2002}

Estimates of the effect of income on health (the absolute income hypothesis) are likely to be biased. Tests of the relative income hypothesis are contaminated by the non-linearity of the individual health income relationship any association between income distribution and population health could be entirely due to it, rather than to any direct effect of relative income on individual health.... However, whilst Rodgers (1979) found that income distribution had a significant negative association with life expectancy in almost all of his regression, we have found that the association is sometimes positive and sometimes negative and is never statistically significant.... The findings should however be a further warning against using aggregate level studies as evidence for the relative deprivation hypothesis.\textsuperscript{8}

\textit{—Gravelle, 2000}

Across Canadian health regions, health status in populations was a function of absolute income but not relative income.\textsuperscript{9}

\textit{—Vafaei et al., 2010}

It can be firmly concluded, however, that there is insufficient evidence supporting Wilkinson’s hypothesis once individual’s income and its differential
impact are taken into account... There are substantial international variations in self-reported health, but they are not linked to the degree of income inequality... Wilkinson’s argument regarding contextual influences was based on a statistical artifact. 10

—Jen et al., 2009

Those with a healthy scepticism will have noticed that I have only quoted studies that support one side of the debate. It’s a slippery and misleading trick and it is exactly what Wilkinson and Pickett do throughout The Spirit Level. The difference is that I made it clear from the outset that there are many conflicting studies. Readers of The Spirit Level would be hard-pressed to guess that there was any debate at all.

In their new postscript and in response to an article I co-wrote for the Wall Street Journal, Wilkinson and Pickett cited a 2009 study of self-reported health in the British Medical Journal which, they say, “leave[s] little room for doubt as to the veracity of these relationships [and] shows unequivocally that inequality is related to significantly higher mortality rates.” 11 With so many studies to chose from, it is reasonable to expect Wilkinson and Pickett to cite one which strongly supports their position. But while the BMJ study is more supportive than most, it can hardly be called unequivocal. It begins by noting that:

Empirical studies have attempted to link income inequality with poor health, but recent systematic reviews have failed to reach a consensus because of mixed findings.

And concludes:

The results suggest a modest adverse effect of income inequality on health, although the population impact might be larger if the association is truly causal... The findings need to be interpreted with caution given the heterogeneity between studies.
It says much how about how weak the alleged ‘consensus’ is that the study Wilkinson and Pickett use as killer proof that inequality causes poor health did not find a strong relationship and acknowledged that the “modest” association was weak enough to imply a lack of causality. If this is “unequivocal” evidence, the reader is entitled to ask what the rest looks like.

Other researchers who have reviewed the evidence have not been so generous. For example:

The undeniable absence of a strong or consistent relationship between inequality and health stands in stark contrast to previous claims.... Contrary to the claims of previous researchers, there is no strong empirical support for the contention that inequality is a determinant of population health, let alone one of the most important determinants.\(^\text{12}\)

This article reviews 98 aggregate and multilevel studies examining the associations between income inequality and health. Overall, there seems to be little support for the idea that income inequality is a major, generalizable determinant of population health differences within or between rich countries.\(^\text{13}\)

Much of the literature, both theoretical and empirical, needs to be treated skeptically, if only because of the low quality of much of the data on income inequality. Although there are many remaining puzzles, I conclude that there is no direct link from income inequality to mortality; individuals are no more likely to die or to report that they are in poor health if they live in places with a more unequal distribution of income.\(^\text{14}\)

The last quoted paragraph comes from a review of the literature conducted by Angus Deaton of Princeton University, one of the world’s most respected economists, whose summary of the evidence has twice as many citations in the scientific literature as Wilkinson and Pickett’s 2006 paper. Despite this, the postscript to *The Spirit Level* finds Wilkinson and Pickett stating that “it is now extremely difficult to argue credibly that these relationships don’t exist. Indeed, those who do so are almost always those who are making political attacks rather than any kind of academic criticism.”\(^\text{15}\) This statement is entirely insupportable. In 2009,
The Oxford Handbook of Economic Inequality evaluated the evidence for the inequality/health hypothesis and concluded:

The preponderance of evidence suggests that the relationship between income inequality and health is either non-existent or too fragile to show up in a robustly estimated panel specification. The best cross-national studies now uniformly fail to find a statistically reliable relationship between economic inequality and longevity.\(^16\)

Having to resort to appealing to authority is regrettable, but since Wilkinson and Pickett are so eager to bill themselves as “epidemiologists with decades of experience in analysing the social determinants of ill health”, it needs to be said that each chapter of The Oxford Handbook of Economic Inequality is written by a team of distinguished professors who are regarded as international experts in their field. The implication that the work of these eminent scholars is “ill-founded and politically motivated criticism” is risible. Unlike Wilkinson and Pickett, none of these academics have formed any political pressure groups and do not have a long history of demanding radical wealth redistribution.

As Sanandaji et al., have noted, the idea that Wilkinson and Pickett took their message directly to the public only after winning the debate in academia is one of The Spirit Level’s most enduring myths:

The general public—the target audience for The Spirit Level—cannot be expected to be aware of the state of research in the field. Wilkinson and Pickett exploit the trust of their readers and give them the impression that their claims represent consensus science, when the opposite is closer to the truth.\(^17\)

The Spirit Level’s endemic misrepresentation of the academic literature is made no less worrisome by its authors’ apparent inability to distinguish between a study which agrees with their hypothesis and one which merely mentions the word ‘inequality’. In response to criticism from Sanandaji et al. that
Epilogue

their book focused on their own work while ignoring heavyweigh academic, Wilkinson and Pickett wrote:

Other ‘heavyweight’ economists, including Nobel laureate, have also written about the significance of inequality for wellbeing and human capital formation. 18

As proof, they cited a study by James Heckman, winner of the Nobel Prize for Economic Sciences. Heckman is the co-author of a study titled ‘The Economics and Psychology of Inequality and Human Development’ 19 but nothing in that paper—or in any of his work—implies support for Wilkinson and Pickett’s inequality hypothesis. When Sanandaji asked Heckman how he felt about having his study cited by the two social epidemiologists, he said bluntly: “This is a misrepresentation of my work.” 20 As Sanandaji explains:

Note Wilkinson and Pickett’s choice of words. They write that Heckman has “written” about inequality and health, which is of course technically true. But what they don’t tell the readers is that while he has indeed written about these variables, he has not found any evidence supporting the claims of Wilkinson and Pickett. 21

Whether it be contemporary academics like James Heckman and Robert Putnam or—almost unbelievably—outspoken opponents of socialism such as Alexis de Tocqueville, Wilkinson and Pickett routinely cite the work of other scholars in a context which suggests that they agree with their hypothesis. Throughout The Spirit Level, they refer to research that links stress and poverty to certain conditions as proof that inequality causes those same conditions. But the material effects of poverty are entirely different from the ‘psychosocial’ effects of inequality, and The Spirit Level conspicuously fails to demonstrate that income inequality is a major cause of stress. It is all well and good citing studies which show, for example, that some people overeat when stressed, or that obesity is more common amongst lower socioeconomic groups, but these valid observations in no
way prove that inequality causes obesity. Without evidence that inequality causes the stress that leads to poor health outcomes, *The Spirit Level* is a series on non-sequiturs.

In some cases, the studies cited say the exact opposite of what Wilkinson and Pickett claim. As discussed earlier in this book, they attempt to explain the higher rate of suicide in more equal countries as a trade-off for a lower homicide rate. The problem with this is two-fold: less equal countries don’t have higher homicide rates, and the countries studied in *The Spirit Level* show no evidence of an inverse relationship between homicide and suicide.

Responding to this on their website, Wilkinson and Pickett wrote: “In fact, there are several pieces of research which show that homicide rates are inversely related to suicide.” But the first study they cite as supporting evidence states quite clearly:

Our analysis indicates, overall, the correlation between homicide and suicide rates across all nations is very weak and statistically insignificant.\(^{22}\)

The shard of truth here is that homicide tends to be more common in very poor countries, while suicide tends to be more common in rich countries. But, as shown in Chapter 4 of this book, there is no correlation between homicide and suicide amongst the rich countries studied in *The Spirit Level*. And that, of course, is the relevant comparison group when discussing Wilkinson and Pickett’s hypothesis.

Either Wilkinson and Pickett are relying on readers not checking their references or they genuinely believe that any study that mentions the word ‘inequality’ in any context is supportive of their case. This was highlighted again when Kate Pickett was interviewed on BBC Radio 4’s *More or Less* programme. It would be hard to find a less politically motivated radio show that *More or Less* – a programme dedicated to discussing the use and abuse of statistics in the modern media. Wisely deciding against passing judgement on such a voluminous topic in a half-hour magazine show, presenter Tim
Harford opted for an interview with Pickett which, in its quiet way, was as devastating as anything written about *The Spirit Level* in 2010.

In this excerpt, Pickett uses the usual ‘consensus’ defence, before being asked about a study she and Wilkinson reference in *The Spirit Level* to support their claim that “researchers at Harvard University showed that women’s status was linked to state-level income inequality.”

KP: We wrote a book that’s intended to be a synthesis of a very vast body of research. Not only our own, but those of other people... There is a consistent and robust and large body of evidence showing the same relationship.

TH: That’s an interesting point that you make. Often, in response to critics, you have referred not to your own book, not to your own data, but to other published research. I’d really like to focus on the research that’s presented in your book. It’s very easy to say ‘there are 50 papers, there are 200 papers, that support our research’ but we don’t really know how you’ve selected those papers.

KP: We actually have completed a systematic review of all of the studies of income inequality and health, and we reference that in our book. We do examine things systematically and certainly—when we are doing our own research, publishing in peer-reviewed journals—we have to be aware of all the literature in the field. But that doesn’t mean that every paper in the field has good methods, comes to the right conclusion, studies the right thing.

TH: I absolutely agree. One of the papers that you refer to in support of your argument on women’s empowerment and women’s status which was published in 1999 by Kawachi and some other authors, you claim supports your findings on women’s status and income inequality. I’ve looked at their abstract. It doesn’t seem to attack that question at all. It’s simply on another subject—a somewhat related subject but not on the subject of income inequality.

KP: They’ve definitely published and we may have inadvertently put the wrong reference into that document [laughs nervously]. But Kawachi and Kennedy have certainly published finding a relationship between income inequality and women’s status. The paper is ‘Women’s Status and the Health of Women and Men: a view from the States’ and it was published in *Social Science and Medicine* in 1999.
TH: That’s the one I’m looking at.23

The only claim in The Spirit Level that has generated anything approaching “a very vast body of research” is that related to health and inequality. Since their book was published, Wilkinson and Pickett have admitted that the correlation between life expectancy and inequality disappears when different measures of inequality are used. They have also said that “we accept that the inequality/health relationship is one of the weaker associations demonstrated in The Spirit Level.”24

The best that can be said of the health/inequality hypothesis is that it remains unresolved and the scatterplot presented on page 82 of The Spirit Level is unlikely to change that. Richard Wilkinson published a similar scatterplot in the BMJ in 1992 and the peer-reviewed literature shows that he was accused of cherry-picking and data-mining at the time. It is no great surprise that he has received similar criticism now that he has filled an entire book with the same type of evidence.

But while there is an ongoing controversy amongst academics regarding the question of inequality and health, the bulk of The Spirit Level involves theories which have little or no support in the scientific literature. Wilkinson admitted as much in an interview with the magazine International Socialism:

“There are about 200 papers on health and inequality in lots of different settings, probably 40 or 50 looking at violence in relation to inequality, and very few looking at any of the other things in relation to inequality. In a way, the new work in the book is all these other variables—teenage births, mental illness, prison populations and so on—and the major contribution is bringing all of that into a picture that had previously been just health and violence.”25

What, then, is left of the idea that The Spirit Level is a “synthesis of a very vast body of research”? Wilkinson himself concedes that “very few” studies have looked at anything other than health in relation to inequality. Although Wilkinson and Pickett
now portray themselves as standing on the shoulders of giants, in almost every important respect they stand alone.

Conspiracy?

Having hastily reinvented themselves as bearers of a consensus, rather than the pioneers they had been in 2009, it is a simple matter for Wilkinson and Pickett to portray those who have put their claims to the test as deniers, right-wing extremists and paid lackeys of industry. It is an impressive trick for a long-standing member of the Socialist Health Association to write a book which concludes with a rousing political call-to-arms while forming two left-wing pressure groups and penning articles in the Guardian about how “broken Britain is Thatcher’s bitter legacy” to accuse other people of being “politically motivated”. This unlikely defence has, however, been remarkably successful.

Wilkinson and Pickett’s first response to the criticisms made in Peter Saunder’s Beware False Prophets was from page one of the manual of knee-jerk student politics. They called him a racist and described his publishers at the Policy Exchange, the manifestly moderate centre-right think tank, as being from the “far-right”. This can hardly be described as a slip of the tongue, since Wilkinson has repeated the slur whilst touring his book in Canada (“then the attacks started coming from the far-right”). Wilkinson can hardly be unaware that the term “far-right” is used almost exclusively to describe neo-Nazis and fascists. That he immediately resorted to using such language about a fellow Emeritus Professor and a former colleague at the University of Sussex, was an early sign that the debate about The Spirit Level was going to be ugly.

It was also a sign that Wilkinson and Pickett would cast their net far and wide in seeking to disparage their opponents. In the new postscript, they write about “the bans on smoking in public places (implemented in Scotland, parts of the USA and
Canada, Rome, Ireland, and England); which in each case have been followed by declines in death rates and have saved thousands of lives.”

This requires a little background information. In recent years, a number of studies have been published purporting to show a large drop in the heart attack rate in the aftermath of a smoking ban. In Scotland, for example, it was claimed that the rate of acute coronary syndrome fell by 17% following the implementation of smoke-free legislation. Oddly, however, the study was based on extrapolations from a selection of hospitals, rather than the admissions records for all Scottish hospitals, which were freely available. When the full figures from the NHS were examined, it became clear that there had not been a drop of 17%, or anything like it.

Today, several years after the ban came into force, it is quite apparent that the smoking ban had no effect on the rate of acute coronary syndrome in Scotland. A number of other studies have also claimed to find a drop in heart attacks following the enactment of smoke-free legislation, but whenever hospital admissions data have been publicly available there has, without exception, been no indication of a significant decline. A recent study—the largest ever conducted on the subject—found that “large short-term increases in myocardial infarction incidence following a smoking ban are as common as the large decreases reported in the published literature”. The disproportionate number of studies finding a decline in numbers is, the authors suggested, the result of publication bias and retrospective data-mining.26

I was one of a number of journalists to write about the Scottish ‘heart miracle’ and similar studies elsewhere. I was not alone. When the Scottish hospital records were released in 2007, the BBC reported it with the headline ‘The facts get in the way of a good story’.27 The Times included it in its end-of-the-year ‘Worst Junk Stats of 2007’ feature.28
If this seems wildly off-topic, it is. Wilkinson and Pickett’s reason for going off on this tangent is to mark me down as some sort of tobacco industry lobbyist just for having written about such issues. They are wise enough not to risk libel by making the allegation explicit, but the implication is allowed to hang in the air.

Upon this thread of innuendo, Wilkinson and Pickett construct an elaborate fantasy involving two unassuming and impartial social scientists under siege from industry-funded “merchants of doubt” who are trying to “give the impression that crucial areas of science affecting public policy are controversial, long after the implications of the science were quite clear.” (Why the tobacco industry would want to discredit *The Spirit Level*, of all books, can only be guessed at. One would think they had bigger fish to fry, but conspiracy theorists are able to overlook such logical conundrums.)

Wilkinson and Pickett’s combination of paranoia and self-aggrandisement falters for the simple reason that critics of *The Spirit Level* are not “free market fundamentalists” and they are certainly not all right-wing. The left-wing journalist Gerry Hassan has written about what he calls “the Fantasyland of *The Spirit Level*”:

Yet, it is almost impossible to compare these countries on equality; they are very different in their cultures, values and histories. Wilkinson and Pickett claim that ‘more equal societies almost always do better’—a universalist, sweeping statement—which cannot be substantiated by most of their data.... Part of the success of *The Spirit Level* is liberal guilt, part the retreat of the left, part wish-fulfilment and projection.29

John Goldthorpe, Emeritus Professor of Sociology at Oxford University, said: “As I read through the book, I have to say that my reaction was one of increasing dismay.”30 Also a left-winger, Goldthorpe’s review of *The Spirit Level* can hardly be attributed to “free market fundamentalism.”
Wilkinson and Pickett [WP] have no time for nicely balanced judgements. They believe that the evidence they present shows beyond doubt that more equal societies ‘do better’, and they are also confident that they have the right explanation for why this is so... Their case is by no means so securely established as they try to make out... it has been called into question by other leading figures in the field—a fact that WP might have more fully acknowledged... WP’s inadequate, one-dimensional understanding of social stratification leads to major problems in their account of how the contextual effect is produced.\textsuperscript{31}

John Kay, Professor of Economics at London Business School, prefaced his review of *The Spirit Level* by saying that he was “sympathetic to its basic stance.” Nevertheless, he found it difficult to take the book’s methodology and conclusions seriously when he reviewed it in the *Financial Times*:

A larger source of irritation is the authors’ apparent belief that the application of regression methods to economic and social statistics is as novel to social science as it apparently is to medicine. The evidence presented in the book is mostly a series of scatter diagrams, with a regression line drawn through them. No data is provided on the estimated equations, or on relevant statistical tests. If you remove the bold lines from the diagram, the pattern of points mostly looks random, and the data dominated by a few outliers.

... An obvious conclusion is that there are many societies which perform well in terms of their own criteria. America, Sweden and Japan are just different from each other. Their achievements are not really commensurable. But Wilkinson and Pickett are not content with this relativist position.\textsuperscript{32}

Andrew Leigh describes himself as “about as anti-inequality an economist as you’ll find”. Formerly a Professor of Economics at the Australian National University, and now an Australian Labour Party politician, Leigh said of his own research into equality: “I had begun the project secretly hoping to find that inequality was bad, and wound up reluctantly reporting no such thing.” When asked his opinion of *The Spirit Level*, he wrote that “John Kay’s view in the *FT* comes closest to my own.”\textsuperscript{33}

“He didn’t read the book thoroughly, obviously,” was Kate Pickett’s response when told about Kay’s review. Another person
who apparently didn’t read it properly was Christian Bjørnskov, Professor of Economics at the University of Aarhus, who reviewed it in *Population and Development Review*:

The bottom line is that this is a well-written, stimulating polemic. It nevertheless suffers from the same problems as one-trick ponies: if the one trick does not impress you, the show is a failure. Wilkinson and Pickett’s trick simply does not hold up to empirical scrutiny. When assessing this book as a contribution to the debate on the “right” level of income differences in modern society, it is a highly interesting, sympathetic attempt at addressing some of the important problems of Western societies. Yet, when assessing this book from a scientific point of view, one is forced to conclude that it is a failure.  

Robert Putnam, author of *Bowling Alone* (see Chapter 3) has also expressed his discomfort with *The Spirit Level*. When asked his view of their work by journalist Shane Leavy, Putnam replied:

I have a mixed view about *The Spirit Level*. On the one hand, I believe that inequality is bad for society in many ways, just as that book argues. On the other hand, Pickett and Williamson’s [sic] work has been heavily (and I believe correctly) criticized as methodologically flawed. (For example, they don’t really show that the relationship between inequality and other bad things is causal, though they assume it is.) I hope that they (or others) will pursue that basic hypothesis in ways that are more scientifically persuasive.

These criticisms, and others like them, are manifestly not politically motivated. While there was no shortage of positive reviews from journalists, particularly on the left (*The Guardian*, *The Independent*, *New Statesman*, *Socialist Review* all provided rave reviews), many respected academics from both left and right have expressed serious concerns.

It suits Wilkinson and Pickett’s narrative to portray critics as being professional ‘merchants of doubt’ from the ‘far-right’. It helps to marginalise those who find fault with the book while deterring their natural supporters from reading the critiques. It is, however, another fiction.
The Spirit Level relies on the premise that countries are fundamentally the same, with income inequality being the main variable that distinguishes them. Wilkinson and Pickett effectively disregard other variables such as absolute income, culture, history, demography, ethnicity, geography, law, politics and climate. Throughout *The Spirit Level*, it is taken for granted that such factors have little or no bearing on their findings and so there is no attempt to adjust the figures for confounding factors, nor to seriously discuss alternative explanations.

In the new postscript, Wilkinson and Pickett group all these other variables together and dismiss them as “cultural differences” which, they say, have a negligible effect on their findings. To illustrate this, they say that Portugal and Spain perform very differently despite being culturally similar, while Japan and Sweden perform similarly despite being culturally different. This is simply not true. In most of the graphs, Portugal is actually closer to Spain than Japan is to Sweden.36

More telling would be a comparison between Japan (the most equal country) and Hong Kong or Singapore (the least equal countries). Despite the huge disparities in income inequality, these three societies perform much the same across nearly all criteria (imprisonment being the main exception). The obvious explanation is that these Asian societies are culturally similar and that culture, not inequality, is the main determinant.

Ignoring other variables and confounding factors would be a flaw in any study but when entire countries are under examination, this flaw becomes overwhelming. Tim Harford asked Pickett about their failure to consider other variables on *More or Less*. Her response was revealing. She and Wilkinson did not “believe” that factors other than inequality have an effect on a country’s performance, so they did not go to the trouble of studying them.
TH: All of your studies are what are called bivariate analyses. In other words, they’re all income inequality plotted against some other variable. Now, my understanding of best practice in social sciences is that you would always control for other variables. You would include 2, 3, 4, 5, 6 other variables and...

KP: Well, you wouldn’t do that arbitrarily. You would do that if you believed those variables were potential alternative explanations of the relationship you’re looking at.

TH: So, if I understand your statement correctly, you didn’t include any multiple variable analyses because you just think that actually none of these variables are of interest—none of them are potential alternative explanations and you can just do the simple income inequality versus x analyses?

KP: That’s right, but of course, again, other researchers have conducted studies that do control for more, where, as well as examining the effect of income inequality at the level of the whole society, people include individual’s own levels of income or levels of education in those analyses and, again, those bear out our findings in relation to health.

TH: We come again to...you’re basically rowing back from your analysis and saying...

KP: No. Indeed I’m not...

TH: ”Don’t look at our analysis, look at these other people because they support us.”

KP: We believe that to control for individual income is actually over-controlling, so we would not consider that best practice.\(^{37}\)

Wilkinson and Pickett may not believe that individual income explains any of the differences between the countries they study, but while this is taken for granted in The Spirit Level, it is not unreasonable to take the view that social outcomes in Portugal, for example, would improve if its national income was the same as Norway’s (which would require a threefold increase in wealth).
Pickett is, however, correct in saying that other researchers have controlled for other variables. Shibuya et al., for example, controlled for income in their study of inequality in Japan and concluded:

After adjustment, individual income was more strongly associated with self-rated health than income inequality.\textsuperscript{38}

Fiscella and Franks controlled for income in their study of inequality in the USA and found:

In this nationally representative American sample, family income, but not community income inequality, independently predicts mortality. Previously reported ecological associations between income inequality and mortality may reflect confounding between individual family income and mortality.\textsuperscript{39}

Absolute income is a crucial confounding factor in studies of income inequality. Much of the debate about inequality and health revolves around the question of whether we can truly disentangle the effects of inequality from the effects of low income. Wilkinson and Pickett completely overlook this issue, and they never remark on the important observation that the poorest countries in their list (Portugal, Greece and New Zealand) all happen to be ‘less equal’. Nor do they comment on the fact that the perennially underachieving US states of Alabama, Louisiana and Mississippi also happen to be amongst the very poorest.

From the outset, income is assumed to have no role to play in \textit{The Spirit Level}. Having announced that economic growth has “largely finished its work”, Wilkinson and Pickett simply assume that further wealth would not benefit the citizens of the countries they study (another glaring ecological fallacy, incidentally). It is assumed that absolute income has no effect because—as they show on page 12—life expectancy is no longer correlated with national income. But they do not test every criteria against income. If they did, they would find that several key outcomes are much more closely correlated with income
than with inequality. This is true even of their cherished survey about trust, as Figures 10.1 and 10.2 show.
Having breezily dismissed income as a third variable, Wilkinson and Pickett turn a blind eye to all other explanations for a country’s performance. Indeed, the only examples of them mentioning real-world differences occur when the ‘more equal’ countries fail to live up to their billing of ‘almost always’ doing better. For example, Wilkinson and Pickett are eager to explain Finland’s high homicide rate by pointing to its high level of gun-ownership while the USA’s high homicide rate is blamed squarely on inequality. When Japan’s foreign aid contributions turn out to be “lower than expected”, they attribute it to the country’s “withdrawal from the international stage following the Second World War”. Britain’s “higher than expected” foreign aid spending, meanwhile, is explained by its “historical, colonial ties to many developing countries.” All this may be true but Wilkinson and Pickett only seem awake to cultural and historical differences when it suits their argument.

They can hardly be unaware that other variables have been shown to explain differences between countries far more convincingly than inequality. In their 2006 review of the literature, they identified twenty-one studies which “started off with supportive findings but then lost them as a result of the various control variables.” Income is one of those variables, but other recognised confounders include spending on health care, which has been found to explain the correlation between inequality and infant mortality:

The association of higher income inequality and higher infant mortality disappears when we control for health care expenditure. Our results indicate that the correlation between infant mortality and income inequality arises as income inequality is high in countries where public investment in health care is low.

Although income inequality was positively associated with low birth weight and infant mortality, the association with infant mortality disappeared with the addition of sociodemographic covariates.
Levels of education have also been shown to explain correlations with inequality:

Multiple regression analysis of the 50 US states and District of Columbia for 1989-90 indicates that the relation between income inequality and age adjusted mortality is due to differences in high school educational attainment: education absorbs the income inequality effect and is a more powerful predictor of variation in mortality among US states.\textsuperscript{43}

Race is another important variable which is never adequately addressed in \textit{The Spirit Level}. For example, one of the few studies looking at inequality and obesity acknowledged that:

Race is known to be significantly correlated with weight status, and is also associated with inequality... As race is a potential confounder of the relationship of interest, we stratify all analyses by race as well as sex.\textsuperscript{44}

Quite rightly so. To any truth-seeking epidemiologist, controlling for a known confounder would always be “best practice”. The results of this study are worth repeating, since they are ignored in \textit{The Spirit Level}, in favour of Pickett’s own research:

We do not find a positive association between inequality and the likelihood of clinically relevant outcomes such as overweight and obesity. Indeed, the direction of association between inequality and weight status is generally negative among subgroups (though significant only for white women)... for non-Hispanic white women, living in a metropolitan area with greater income inequality is associated with lower BMI, lower odds for being overweight, and lower odds for being obese. [Italics in the original]

Race has been shown time and again to be a major confounder in studies of inequality, to the extent that this one variable explains the \textit{entire} correlation between inequality and poor health. This has been shown to be true in the USA:

In the results presented below, we show that, once we control for the fraction of the population that is black, there is no relationship in 1980 nor in 1990
between income inequality and mortality across either states or cities... That the estimated effects of income inequality are potentially confounded by the effects of race has been recognized since the first papers on the topic. Blacks have higher mortality rates than whites and, on average have lower incomes, so that in places with a substantial black population, both income inequality and mortality tend to be higher. 45

In Canada:

We replicate the finding that, net of the racial/ethnic composition of the population, the effects of income inequality are not significant.46

And in New Zealand:

There is no convincing evidence of an association of income inequality within New Zealand with adult mortality. Previous ecological analyses within New Zealand suggesting an association of income inequality with mortality were confounded by ethnicity at the individual level.47

The well-established importance of race as a confounding factor provided Wilkinson and Pickett with the excuse to land their lowest blow. In his book *Beware False Prophets*, Peter Saunders demonstrates that health and social outcomes are more closely correlated with the ethnic make-up of US states than with their levels of income inequality. For this, Wilkinson and Pickett accused him of a “seriously racist slur”. It was, they said, “racist because it implies the problem is inherently the people themselves rather than their socioeconomic position”.48

It implies nothing of the sort. If Wilkinson and Pickett think it is racist to say that there are a host of cultural and historical reasons why blacks tend to do worse than whites in the USA, then there are plenty of black community leaders and black politicians who are racist. No serious discussion of modern-day America can ignore the legacy of slavery and segregation, as well as the more subtle forms of ongoing discrimination which continue to hold African-Americans back. Black Americans have, on average, higher rates of obesity, higher
homicide rates and lower life expectancy. It should, therefore, be no surprise that states with large black populations tend to do worse under these criteria.

There is no doubt that racial inequality contributes to income inequality. Wilkinson and Pickett argue instead that income inequality is, at heart, the cause of racial inequality. Aside from being counterintuitive, this cannot be so because the correlation between race and health and social problems is stronger than the correlation with income inequality.

A significant clue lies in the pages of The Spirit Level itself. Wilkinson and Pickett’s discussion of mental health is a mass of contradictions. Having warned of the dangers of comparing apples and oranges, they proceed to do just that by cobbling together results from different studies which even they coyly admit are “not strictly comparable”. They attribute their failure to find a correlation between inequality and mental illness in the USA to the fact that mental illness does not have a social gradient, but this does not deter them from reporting a correlation between inequality and mental illness on an international level.

They then mention, almost in passing, that rates of mental illness are evenly distributed between different races. In light of their failure to find a correlation with mental illness in US states, this should have been a Eureka moment but, as Saunders writes:

[T]hey fail to draw the obvious conclusion from their failure to find a relationship with inequality, which is that they only get state-level correlations with income inequality when there are underlying correlations with race to generate them. [Italics in the original].

Since there is no relationship between race and mental health, they cannot find a relationship between inequality and mental health. But since there are relationships between race and many other criteria, they find correlations with inequality. But those correlations are statistical associations resulting from Wilkinson
and Pickett’s failure to adjust for race. They are not causal. Inequality is a symptom, not a cause.

Wilkinson and Pickett never adequately address the question of causality. There are many important confounders such as income, race, education and material deprivation which are correlated with inequality, but are not caused by inequality. Conversely, many social problems such as crime, drug abuse and gang formation do cause inequality because young people growing up in environments with gangs, drug abuse and high levels of crime are less likely to succeed in society. We can address those issues by fostering job creation or crime reduction in neighbourhoods with social problems. But, by Wilkinson and Pickett’s logic, inequality is the cause of these problems, which leads us to the improbable conclusion that societal malaise can be alleviated by reducing income in the surrounding neighbourhoods.

There is plenty of research—all of it ignored in The Spirit Level—showing that inequality does not have an independent effect on health and social problems once other variables have been controlled for. It should go without saying that countries differ from one another in many ways that have nothing to do with income inequality. That these differences will lead to different outcomes should be equally obvious. Wilkinson and Pickett justify their refusal to consider other variables in the postscript, saying “including factors that are unrelated to inequality, or to any particular problem, would simply create unnecessary ‘noise’ and be methodologically incorrect.” With this one sentence, every historical, cultural, religious, political, legal, geographical, climatic and demographic difference between whole societies is dismissed as ‘noise’. Again, they are assuming that these factors are “unrelated to inequality” without putting that assumption to the test. It is no wonder Wilkinson and Pickett fail to identify confounding factors. They were simply not looking for them.
The new postscript to *The Spirit Level* finds Wilkinson and Pickett accusing their critics of “selectively removing countries on the grounds that they were outliers.” Outliers do indeed play an important part in several of *The Spirit Level's* graphs. The correlation between inequality and homicide rests entirely on the USA being an extreme outlier. The correlation between inequality and obesity depends entirely on Japan and the USA being outliers (not to mention the exclusion of Singapore, Hong Kong and South Korea, all of which have similar rates of obesity to Japan). The correlation with trust depends entirely on the Nordic nations being outliers.

The significance of this should not need underlining. To take homicide as an example, there is no evidence of a relationship between inequality and homicide when 22 of the countries are studied. The 23rd country—the USA—has a much higher rate and pulls the regression line upwards dramatically. Using this distorted regression line as evidence that inequality causes murder means ignoring data from 22 countries in favour of data from just one. There are many reasons why the USA has a high murder rate, but if inequality was the root cause, we would expect to see it affecting the other countries. It doesn’t, and excluding the USA as an outlier demonstrates the lack of causation.

If we were presented with a graph showing low levels of participation in basketball in 22 countries but a much higher figure for the USA, few of us would conclude that there was a true causal relationship between inequality and basketball. Americans just play a lot of basketball. And yet, for several of *The Spirit Level's* graphs, outlying data of this type are used as proof of a causal relationship despite the great majority of the countries being totally unaffected by the supposed cause.

Wilkinson and Pickett feign ignorance about the importance of outliers. In their postscript, they portray testing
for outliers as an underhand trick to exclude unfavourable data. It is, of course, nothing of the kind. The point of testing for outliers is not to “selectively remove countries” and then present the result as the ‘real’ graph, but to see if the relationship holds up without the outlier being present. In *Beware False Prophets*, Peter Saunders explains how and why statisticians use box plots to identify outliers. He then shows, as I do in this book, that the trend line for homicide is being thrown out by a single extreme outlier.

It is fantastically implausible to think that Wilkinson and Pickett are not aware of the importance of outliers in statistics. In fact, we know that they are because when they find a reasonably strong statistical relationship (for rates of imprisonment) they write: “Even if the USA and Singapore are excluded as outliers, the relationship is robust among the remaining countries.”51 They make no such guarantee of their other graphs, for the simple reason that they are not robust enough to withstand this basic test.

One of the dangers of not recognising outliers is that your trend line will become skewed and no longer reflect reality. Wilkinson and Pickett focus on their trend line to such an extent that they forget what the actual data are telling them. In the last chapter of *The Spirit Level*, Wilkinson and Pickett claim that if Britain reduced income inequality to the same level as Sweden, Finland, Japan and Norway, its murder rate would fall by 75%. This prediction goes far beyond what the data show. (Even if the association was real, their correlation coefficient tells them that inequality accounts for less than half the difference, and yet they assume it accounts for 100% of the difference—a basic statistical howler.)

Worse still, they are basing their prediction entirely on their trend line, which tells them that Britain *should* have a much higher murder rate than it does. But that trend line has become hopelessly skewed by the USA. As Wilkinson and Pickett’s own graph shows, Britain actually has a lower murder rate than
Sweden and Finland and has a lower murder rate than the average of those four ‘more equal’ nations.

The irony of Wilkinson and Pickett accusing their critics of picking and choosing which countries to study will not be lost on readers of this book. Wilkinson was being criticised for his selective use of data long before *The Spirit Level* hit the shelves and his latest choice of countries raises the same questions. Their justification for confining their analysis to 23 countries is because “these countries are on the flat part of the curve at the top right in Figure 1.1 on p. 7, where life expectancy is no longer related to differences in Gross National Income.” Quite so, and it was that very graph which first alerted me to the fact that Wilkinson and Pickett had excluded several countries. South Korea, Hungary, Slovenia and the Czech Republic all appear on that graph as being as rich or richer than Portugal. They are all on the “flat part of the curve”. It was not me, but Wilkinson and Pickett, who arbitrarily decided that Portugal was ‘rich enough’ to merit inclusion. All I have done in this book is include countries of comparable or greater wealth than Portugal as shown in Wilkinson and Pickett’s own graph. Without a convincing justification for why places like the Czech Republic and South Korea cannot be considered “rich market societies”, we must ask the next question: why do these societies conspicuously fail to fit Wilkinson and Pickett’s theory? The United Nations classes these countries as being of ‘very high human development’, why doesn’t *The Spirit Level*?

Their insistence on never having “picked problems to suit our argument” is undermined by, for example, their focus on state aid at the expense of private philanthropy, or by their emphasis on imprisonment rather than crime. Their claim to “never pick and choose data points to suit our argument” is at odds with references 2 and 6 in *The Spirit Level* which show one year’s data being used for one graph and another year’s data being used for the next, even though the subject matter—life expectancy—is the same.
As for using “the same measures of inequality” (as they said they did in an article in Prospect magazine\textsuperscript{52}), they address this early in The Spirit Level, saying:

To avoid being accused of picking and choosing our measures, our approach in this book has been to take measures provided by official agencies rather than calculating our own.

This is no great claim to integrity. It would be very odd if they started developing their own bespoke measure of inequality. But if they really wished to “avoid being accused of picking and choosing” they would have used the same official measure throughout. In fact, they use no fewer than five different measures of inequality in The Spirit Level. Having correctly explained to the reader that the Gini coefficient is “the most common measure” which is “favoured by economists”, they proceed to ignore the Gini in favour of comparing the top and bottom 20% when making international comparisons. They then switch to the Gini coefficient when looking at US states and then use a completely different measure when comparing working hours (p. 229). They then adopt a measure which compares the bottom and top 10% (p. 240) and, finally, in their new edition, measure inequality in reference to the top 1% (p. 296).

The effect of this chopping and changing can be seen by comparing the graph on page 240 to the graph on page 296 (of the new edition). The first graph shows that inequality in the USA has fallen since its peak in the early 1990s; the second graph shows that inequality in the USA rose sharply in the 1990s and peaked in 2008. Wilkinson and Pickett’s aim in the postscript is to demonstrate a correlation between inequality and the financial crashes of 1929 and 2008. They write that “both crashes happened at the two peaks of inequality”. Either they have forgotten, or they are hoping the reader has forgotten, that they wrote in the previous chapter that inequality in the USA “peaked in the early 1990s”.

Epilogue
Epilogue

Whilst there is nothing wrong with using the share of wealth held by the top 1% as a measure of inequality, this is the only time it is used in *The Spirit Level*. This is perhaps unsurprising, since under this measure Norway and Denmark are less equal than the USA.\(^5\) It does, however, demonstrate how Wilkinson and Pickett switch reference points to suit whatever argument they are making at the time.

Conclusions

Wilkinson and Pickett’s new postscript will no doubt play well with readers who were convinced by the rest of *The Spirit Level*. Their insistence that they stick to “a strict set of criteria, applied with no departures or exceptions” has a hollow ring in light of their frequent switching between data sets. It also fails to address the point that the selection of countries was flawed and arbitrary from the outset, even if they had stuck rigorously to it thereafter.

A new edition of any book affords the author an opportunity to make corrections and include newly available data. Since *The Spirit Level* appeared in hardback, new statistics about trust, homicide, life expectancy, obesity and several other issues have been published by official agencies, most of which further weaken the correlations upon which Wilkinson and Pickett rely. No attempt has been made to incorporate these data in the new edition, nor have any criticisms been acknowledged, even on basic, irrefutable points of fact. Instead, they have used a reprint as an opportunity to create a contrived fairy-tale in which “free market fundamentalists” and “merchants of doubt” have conspired to make incontrovertible scientific evidence appear questionable.

This is the sheerest fiction. Questions have been raised about the bold conclusions of *The Spirit Level* because it is riddled with methodological errors, apparent cherry-picking, flawed reasoning and wishful thinking. Wilkinson and Pickett’s
misrepresentation of the academic literature was brazen in *The Spirit Level* and has become near-absolute with the appearance of the new postscript. Far from being the subject of a co-ordinated attack by nefarious vested interests—I happen to be in the bottom 20% of earners myself—their book has been criticised by everyone from Swedish economists, Irish psychologists and British sociologists—as well as numerous journalists, bloggers and reviewers—for the simple reason that they have read it. As a best-seller, it has transcended what Wilkinson calls the “left-wing ghetto”. And amongst its large readership have been many rational people whose jaws dropped a little more at the turn of every page.
10. Epilogue

(2) The Spirit Level, p. x
(3) The Spirit Level, p. 278 (second edition)
(7) Income inequality and self rated health in Britain’, S Weich, G Lewis, S P Jenkins, Journal of Epidemiology and Community Health, 2002
(9) ‘Relationships between income inequality and health: a study on rural and urban regions of Canada’, Vafaei et al., Rural and Remote Health, 10, 2010
(10) ‘Compositional and contextual approaches to the study of health behaviour and outcomes’ Jen et al., Health & Place, Vol. 15 (1), March 2009; pp. 198-203
(15) The Spirit Level, p. 279
Epilogue

(21) Ibid.
(23) More or Less, BBC Radio 4, 27.08.10
(27) http://news.bbc.co.uk/1/hi/magazine/7093356.stm
(29) ‘The Spirit Level lacks all balance to be correct’, The Scotsman, G. Hassan, 31.7.10
(30) ‘The Spirit Level: Britain’s new theory of everything?’, BBC Radio 4, 12.10.10
(33) ‘Look at the changes, not at the levels’, Andrew Leigh, http://previousleigh.wordpress.com/2009/12/14/look-at-the-changes-not-at-the-levels/ (14.12.09) (Leigh also notes that countries that see the biggest increases in inequality also see the biggest improvements in health.)
(36) Spain is closer to Portugal than Japan is to Sweden under the following criteria: imprisonment, obesity, women’s status, recycling, working hours, patents, ambition, foreign aid, child well-being, education and trust. The reverse only applies to teen births, drug use, homicide and (by the finest of margins) infant mortality and life expectancy.
(37) More or Less, BBC Radio 4, 27.08.10
(38) ‘Individual income, income distribution, and self rated health in Japan: cross sectional analysis of nationally representative sample’, K. Shibuya et al., *British Medical Journal*, 5.1.02
(41) ‘Income distribution, infant mortality, and health care expenditure’, Tilman Tacke, Robert J Waldmann, *CEIS Research Paper* No. 146, 9.3.09. (This finding was made by the same researcher—Waldmann—who first found the statistical correlation between inequality and infant mortality in the early 1990s and who was cited by Wilkinson at the time.)
(42) Primary care, infant mortality, and low birth weight in the states of the USA’; L. Shi et al., *Journal of Epidemiological & Community Health*, 58, 2004; pp. 374-80
(43) ‘Education, income inequality, and mortality: a multiple regression analysis’, Andreas Muller, *British Medical Journal*, 5.1.02; 324
(49) *Beware False Prophets*, Peter Saunders, Policy Exchange, 2010; p. 98
(50) *The Spirit Level*, p. 285
(51) *The Spirit Level*, p. 148-9
Epilogue

(53) ‘Long-run changes in the concentration of wealth’, H. Ohlsson, J. Roine and D. Waldenström, World Institute for Development Economics Research, Research paper No. 2006/103. The top 1% owns 21% in the UK, 25.4% in Norway, 27.2% in Denmark and 20.8% in the USA.