Beware False Prophets

Equality, the Good Society and *The Spirit Level*

By Peter Saunders, edited by Natalie Evans

Policy Exchange

Beware False Prophets

Equality, the Good Society and *The Spirit Level*

Peter Saunders,
edited by Natalie Evans

Policy Exchange

Policy Exchange is an independent think tank whose mission is to develop and promote new policy ideas which will foster a free society based on strong communities, personal freedom, limited government, national self-confidence and an enterprise culture. Registered charity no: 1096300.

Policy Exchange is committed to an evidence-based approach to policy development. We work in partnership with academics and other experts and commission major studies involving thorough empirical research of alternative policy outcomes. We believe that the policy experience of other countries offers important lessons for government in the UK. We also believe that government has much to learn from business and the voluntary sector.

Published by
Policy Exchange, Clutha House, 10 Storey's Gate, London SW1P 3AY
www.policyexchange.org.uk

ISBN: 978-1-906097-78-3

Printed by Heron, Dawson and Sawyer
Designed by SoapBox, www.soapboxcommunications.co.uk

Contents

About the Authors

Peter Saunders was until 1999, Professor of Sociology at the University of Sussex, where he is still Professor Emeritus. Since then he has been Research Manager at the Australian Institute of Family Studies in Melbourne, and Social Research Director at the Centre for Independent Studies in Sydney. He is now back in the UK working as an independent consultant, specialising in social policy. He has written a number of reports for Policy Exchange on poverty and family policy. His website address is www.petersaunders.org.uk

Natalie Evans is Deputy Director at Policy Exchange, responsible for the strategic direction and output of the research team.

Executive Summary

> *The ultimately possible attitudes towards life are irreconcilable, and hence their struggle can never be brought to a final conclusion... Science is organised in the service of self-clarification and knowledge of interrelated facts. It is not the gift or grace of seers and prophets dispensing sacred values and revelations... As science does not, who is to answer the question, 'What shall we do, and how shall we arrange our lives?'...Only a prophet or saviour can give the answers. If there is no such man, then you will certainly not compel him to appear on this earth by having thousands of professors, as privileged hirelings of the state, attempt as petty prophets to take over his role.*
>
> (Science as a vocation, reprinted in H.Gerth and C. Wright Mills, From Max Weber, Routledge 1948, pp.152-3).

In a book published last year, called *The Spirit Level*, Richard Wilkinson and Kate Pickett argued that income inequality harms not only the poorest people at the bottom end of the income distribution, but almost everybody in society, no matter how prosperous they are. They backed up their claim with statistical evidence apparently showing that more unequal countries (and within the USA, more unequal states) suffer from higher crime rates, worse infant mortality, greater obesity, poorer education standards, lower average life expectancy, less social mobility, and much else besides. The authors concluded that we would all benefit from a more egalitarian distribution of income.

Their argument has major implications for public policy. For centuries, political philosophers have argued about what (if anything) should be done about unequal shares. On the one hand, it seems right to redistribute resources from people who have plenty to those who have little, but on the other hand, it seems wrong to take resources away from people who have worked hard or taken risks simply to

make others more equal. *The Spirit Level* seems to offer a resolution of this ethical dilemma, for it claims that rich and poor alike stand to gain from income redistribution. Once a society has achieved a modest level of prosperity, Wilkinson and Pickett suggest that people's wellbeing depends on sharing resources, so redistribution turns out to be in everybody's best interests.

Not surprisingly, this message has received an enthusiastic reception from politicians and pundits on the left who believe *The Spirit Level* offers a rational, evidence-based justification for the radical egalitarianism to which they have long been emotionally committed. However, careful evaluation and analysis shows that very little of Wilkinson and Pickett's statistical evidence actually stands up, and their causal argument is full of holes.

They base their claims on two sets of statistics: international data on 23 of the world's richest countries, plus data on the 50 US states. On most of the indicators they examine, income inequality is found to correlate with social problems in both data sets. It is not just that more equal countries perform better than less equal countries, but that more equal states within the USA perform better than the less equal states. Any critique of Wilkinson and Pickett's analysis must therefore account for both their international findings and their US state data.

In this report, Wilkinson and Pickett's empirical claims are critically re-examined using (a) their own data on 23 countries, (b) more up-to-date statistics on a larger sample of 44 countries, and (c) data on the US states. Very few of their empirical claims survive intact.

Of 20 statistical claims examined, 14 are shown to be spurious or invalid, and in only one case (the association internationally between infant mortality and income inequality) does the evidence unambiguously support Wilkinson and Pickett's hypothesis. Contrary to their claims, income inequality does not explain international homicide rates, childhood conflict, women's status, foreign aid donations, life expectancy, adult obesity, childhood obesity, literacy and numeracy, patents, or social mobility rates. Nor

does it explain variations among US states in homicide, infant mortality or imprisonment rates.

The statistical analysis in *The Spirit Level* is heavily flawed. There are many instances where graphs are presented in which just one or two extreme cases are used to support unwarranted generalisations. For example, the claim that there is an association between a country's homicide rate and its level of income inequality depends entirely on the high murder rate in the USA (which probably has more to do with its gun control laws than its income distribution). Across the other 22 countries, there is no association between income distribution and murder rates. Similarly, the claim that average life expectancy is linked to income inequality rests entirely on the longevity of people in Japan (which probably has something to do with their diet, genes or a mixture of the two). Take Japan out of the analysis, and the apparent association with income inequality again collapses.

> "Contrary to their claims, income inequality does not explain international homicide rates, childhood conflict, women's status, foreign aid donations, life expectancy, adult obesity, childhood obesity, literacy and numeracy, patents, or social mobility rates"

In other instances, the authors' claims rest illegitimately on the influence of specific 'clusters' of countries or states. For example, the (more equal) Scandinavian nations routinely appear at one end of many of their graphs, and the (less equal) Anglo nations often appear at the other. But these differences probably reflect a deeper divergence between Nordic and Anglo cultures, for when we look beyond these clusters and search for evidence that might link inequality to social outcomes in other countries, we search in vain. The argument that women's interests are better served in more equal countries, for example, rests entirely on the fact that women do better in Scandinavia. But outside Scandinavia, women fare no better in egalitarian countries like Japan and Austria than in inegalitarian ones like Portugal or Singapore.

Wilkinson and Pickett are loathe to allow history and culture into their analysis, yet it is clear that it is cultural patterns that are often generating their findings. Sweden and Japan, for example, have the income distributions they have because of the kinds of societies they are. They are not cohesive societies because their incomes are equally distributed; their incomes are equally distributed because they evolved as remarkably cohesive societies. To explain why, we have to look to their histories and at factors like social homogeneity and closure, but *The Spirit Level* resists any such analysis.

The authors of *The Spirit Level* are very selective in their choice of evidence, and the book ignores an array of social indicators which are worse in more equal countries. Suicide rates, HIV infection rates, alcohol consumption and divorce rates are all higher in more equal countries, and fertility is lower, but these trends go unanalysed. The book also ignores trends over time which show, contrary to their hypothesis, that the countries where income inequality grew fastest over the last 30 years are those where infant mortality rates and average life expectancy actually improved the most.

The book also ignores explanatory factors which might undermine its core hypothesis. In the US, for example, the proportion of African-Americans in a state is often a much stronger predictor of social outcomes than the level of income inequality, but Wilkinson and Pickett never take ethnic composition into account in their models. When we do this, the association with income distribution often disappears: state homicide rates, infant mortality rates, average life expectancy and imprisonment rates all reflect ethnic composition, not income inequality.

This report shows that *The Spirit Level* has little claim to validity. Its evidence is weak, the analysis is superficial and the theory is unsupported. The book's growing influence threatens to contaminate an important area of political debate with wonky statistics and spurious correlations. The case for radical income redistribution is no more compelling now than it was before this book was published.

1. Inequality, politics and social science

Income inequality in Britain is higher than in most other EU countries. It has also been increasing over time. Both of these facts are incontrovertible. The question to be addressed in this report is whether this matters, and if so, why?

In a much-publicised book called *The Spirit Level*, Richard Wilkinson and Kate Pickett have recently argued that Britain's relatively high level of income inequality is hugely damaging – not just for the poorest people at the bottom end of the income distribution, but for almost everybody, regardless of how prosperous they are.[1] They back up this dramatic claim with statistical evidence which purports to show that in more unequal countries (and within the USA, in more unequal states), life is worse for almost everybody: crime rates are higher, infant mortality is higher, obesity is higher, education standards are lower, average life expectancy is lower, social mobility is lower, and so on. They conclude that we would all benefit from a more egalitarian distribution of income.

This report evaluates Wilkinson and Pickett's claim, looking both at their evidence and at the explanations they offer for why inequality might have the effects they attribute to it. We shall see that very little of their statistical evidence stands up, and that their causal argument is full of holes.

The Spirit Level has attracted many enthusiastic plaudits among left-wing commentators since its publication in 2009. The reason is that it appears to offer a 'scientific' validation of their ideological commitment to income levelling. Because they like its message,

1 Richard Wilkinson and Kate Pickett, *The Spirit Level* Allen Lane, 2009

these commentators have been disinclined to delve too deeply into its evidence or its methods of analysis. Yet as soon as the book is subjected to even a fairly cursory examination, it becomes obvious that it is deeply flawed. Given the growing influence that this book is having on policy debates in Britain and overseas, it is important that these weaknesses should be exposed.

The weaknesses of *The Spirit Level* do not, of course, mean that the case for income redistribution collapses. There have always been powerful ethical arguments for and against greater equalisation of incomes, and this impassioned debate will doubtless continue. But social science cannot resolve this ethical argument.

Income inequality in Britain

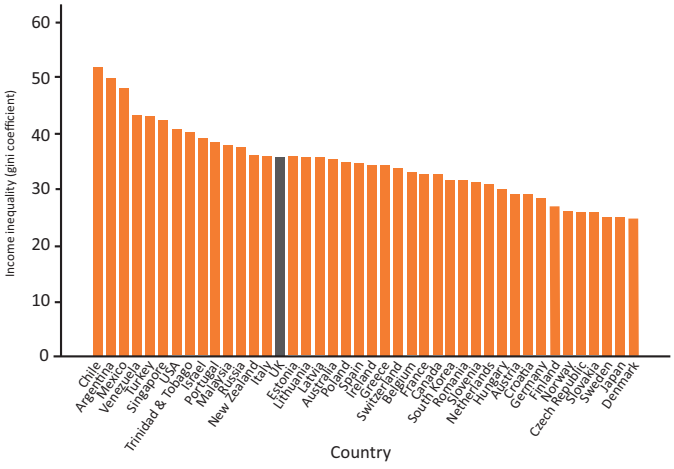
Economists measure income inequality in a variety of ways. The most common is a statistic called a gini coefficient which tells us how much the distribution of incomes deviates from perfect equality. The higher the gini coefficient, the more unequal the distribution. In a society where one person earned all the income and everyone else received nothing, the *gini coefficient* would be 1. In a society where everybody received exactly the same income, the gini coefficient would be zero. Ranking countries by the gini coefficient therefore allows us to see which are more, and which are less, unequal.

Figure 1 ranks 42 of the richest countries in the world according to their degree of income inequality as measured by the gini coefficient.[2] Britain comes 15th, with around the same level of inequality as Australia, New Zealand, Italy and the Baltic states. We are a lot less unequal than the Latin American countries, which cluster at the far left-hand side of the chart, and we are somewhat less unequal than the Russians and the Americans. But most EU nations come below us in the rankings, and the Scandinavian countries, which cluster at the far right-hand end of the chart, are clearly

2 The chart excludes rich countries, like Luxembourg, with a population under 1 million, as well as Saudi Arabia, Libya, Botswana and Gabon.

much less unequal than we are. On this measure, Denmark is the most equal country in the world, with Japan a close second and Sweden third (although when it comes to the distribution of wealth, rather than the distribution of income, this pattern looks rather different, with the UK significantly more egalitarian than Sweden).[3]

Figure 1: Income inequality in the world's richest nations[4]



Not only does Britain appear to be a relatively unequal country when comparing our pattern of income distribution with that of other European nations, but (like many of these other countries) we have also become significantly more unequal over the last 50 years.

Figure 2 plots our income inequality trend since 1961. It measures inequality using both the gini coefficient (the solid line) and the ratio of the income received by people near the top of the distribution to that received by those near the bottom (the dotted line).[5] It is clear from the graph that it makes little difference which measure we use, for the pattern in both is very similar.

3 According to the National Equalities Panel Report, *An Anatomy of Economic Inequality in the UK* (Government Equalities Office, 2010, Table 2.1) the gini index for wealth inequality in the UK is 66, compared with 68 in Finland, 80 in Germany and 89 in Sweden. The report suggests that this 'surprising' result may reflect the relatively weak development of private pensions in countries where the welfare state is more all-encompassing. The analysis in *The Spirit Level* focuses on income inequality, rather than wealth inequality, although the logic of the book's argument should apply just as much to the latter as to the former.

4 National gini coefficients taken from UN *Human Development Report* 2009 http://hdr.undp.org/en/reports/global/hdr2009/

5 More specifically,. we compare the incomes of people who are nine-tenths the way up the distribution with the incomes of those who are only one-tenth the way up. This 90/10 ratio is another common measure of inequality. Again, the data come from the 2009 *Human Development Report*. In Figure 3, both the gini coefficient and the 90/10 ratio are calculated on 'equivalised incomes' (i.e. taking account of household size and composition) and are before housing costs are met.

Figure 2: Income inequality in Britain since 1961

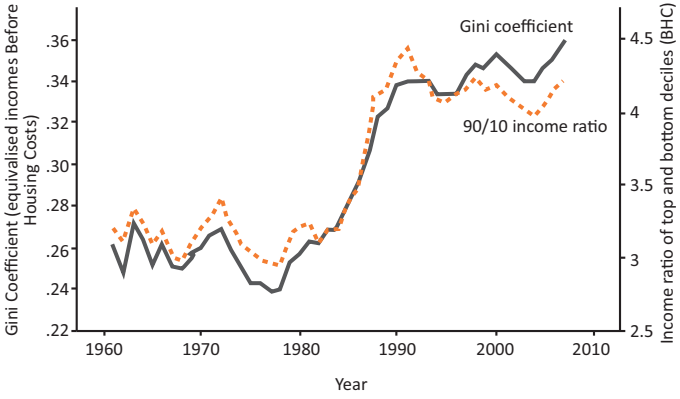


Figure 2 shows that income inequality in Britain did not vary much in the 1960s, and it actually dropped in the 1970s (mainly because the value of state pensions rose significantly in real terms). However, inequality then rose steeply during the 1980s, partly because taxes on high earnings were cut, partly because more women entered the labour force as part-time workers, and partly because the wages of highly-skilled workers relative to those with few skills were driven up by technological change and the opening up of world markets (something that increased income inequality in most advanced economies).

Since the early 1990s, the trend has flattened out again, and oscillations have been more modest. The recession of the early nineties compressed wages relative to welfare benefits, and the Labour government's anti-poverty programme boosted the incomes of working families and those on benefits.[6] Nevertheless, inequality is today higher than it has ever been since records began, and even after 13 years of a Labour government, there was no reversal of the substantial increase in inequality that occurred during the 1980s.

6 Most of this period is analysed in detail by Mike Brewer, Alastair Muriel and Liam Wren-Lewis, *Accounting for changes in inequality since 1968* (2009 paper prepared for the government's Equality and Human Rights Commission and available from the Institute of Fiscal Studies website, http://www.ifs.org.uk/publications/4699)

Is inequality unjust?

More than any other single issue, economic inequality has for generations functioned like a litmus test of political ideology:

- The majority on the left believe in equalising incomes and wealth. Few left-wingers think income differences should be flattened completely, but they do think it is wrong that anybody should receive a lot more than anybody else. They therefore tend to favour tax and welfare policies which aim, not only to improve the living standards of those at the bottom, but also to reduce the prosperity enjoyed by those at the top. Seen in this light, the fact that Britain's income distribution is more unequal than that of most other western European countries, coupled with the evidence that it has increased significantly over the last 30 years, is a serious cause for concern.
- For those on the political right, concern about the way incomes are distributed is more muted. Few right-wingers disapprove in principle of 'progressive' taxation and the provision of state welfare benefits, but they worry that redistributive policies like these can destroy work incentives, and they believe it is right that people who work hard and exploit their talents should enjoy the material rewards that come with success. From this perspective, what really matters is not equality of outcomes, but equality of opportunity.[7] Provided there are no major barriers to people competing for material rewards, it is not 'unfair' if some end up with more than others.

Neither of these positions seems obviously 'wrong', yet they are logically incompatible. The reason the issue of equality has polarised political debate so sharply for so long is precisely that it revolves around a clash between two sets of basic, moral principles, both of which seem intuitively correct and desirable to many people, even though they each undermine the other.

7 I have recently examined the evidence on equality of opportunity in Britain in Peter Saunders, *Social Mobility Myths*, London, Civitas, 2010

On the one hand, it does seem wrong and unfair when we see football stars, bank executives and business tycoons earning more money than they know what to do with while the unemployed, pensioners and single parents struggle to pay their rent and heating bills.

On the other hand, it also seems wrong and unfair to take money away from people who have worked hard to give it to those who show little inclination to work; or to take away the profits of those who have risked their life savings to bring a new invention to market in order to compensate those who have risked nothing.

In the real world of politics, we constantly fudge the line between these two core ethical principles, yet the tension is always present. Philosophers and theologians have wrestled for centuries with the ethical dilemmas thrown up by this clash of fundamental principles, but nobody yet has come up with a clear and compelling reason for favouring one principle over the other.

“It does seem wrong and unfair when we see football stars, bank executives and business tycoons earning more money than they know what to do with while the unemployed, pensioners and single parents struggle to pay their rent and heating bills”

For a time in the 1970s, the left thought that John Rawls had succeeded in making a compelling case for egalitarianism when he proposed that we should think of ourselves in an ‘original position’ in which we have to agree on ethical principles of social organisation without knowing what position in society each of us will occupy. Rawls said a ‘just distribution’ is the one we would all accept while we were operating behind this ‘veil of ignorance.’ He was under no doubt that, in these conditions, we would agree to share resources equally. The one exception to this was ‘the difference principle’ – we would accept unequal distribution if it could be shown that it favours the least well-off (e.g. by incentivising economic activity from which the poor can expect to benefit). For Rawls, inequality is illegitimate unless it can be shown to help the poor.[8]

8 John Rawls, *A Theory of Justice* Oxford University Press, 1972

But no sooner had Rawls established this argument for equality than Robert Nozick offered an equally compelling refutation.[9] He likened Rawls's 'original position' to the situation of a group of students being asked to agree on the distribution of examination grades before starting their course. Having no way of knowing how well they are likely to perform, Nozick accepts that they would probably all agree to share the same marks. But in reality, they do not have to make such decisions in ignorance of their own vices and virtues. Some work hard and revise while others are lazy, and this would make it grossly unfair to insist they should all be graded the same. Nozick therefore proposed that we should gauge a just distribution simply by asking whether people have established a legitimate right to what they have. If they have worked for what they've got, or if they have received it from somebody else as a result of a voluntary gift or exchange, then they are entitled to keep it, end of story.

Philosophers like Rawls and Nozick have helped clarify our thinking about inequality, but they have clearly not resolved the ethical dilemma at the heart of the issue. In the end, we are still left wrestling with our own consciences. If we privilege the needy, we undermine the deserving. If we recognise just deserts, the needy go unheeded.

But if philosophy cannot help, what about science? Ever since Auguste Comte developed his blueprint for a 'positive science of society' in France in the 1820s, there have been visionaries who believe that social statistics could resolve the great ethical problems which the philosophers have failed to determine for us.

Comte (and later positivists like Emile Durkheim) believed that a new positive science of society, which he called 'sociology', would be able to identify the causes of social malaise, in the same way as medical science can discover the causes of diseases of the human body. Just as doctors draw on medical science to diagnose pathologies and prescribe remedies, so sociological experts (Comte saw

9 Robert Nozick, *Anarchy, State and Utopia* Oxford, Basil Blackwell, 1974

them as a new 'priesthood') should be able to draw on their knowledge of social causation to determine what is going wrong in a society, and how to put it right. Doctors do not need to consult moral philosophers to know that 'health' is good and 'illness' is bad. So too, social scientists should be able to judge which social policies are beneficial and which are harmful without having to get embroiled in interminable debates about ethics. Whatever promotes the harmonious functioning of society is good; whatever undermines it is bad; and social science should be able to distinguish the two.[10]

In the event, of course, sociology (and the other 'social sciences') has proved remarkably inept at providing us with a medical kitbag for righting social ills. Society has turned out to be a much more complicated 'organism' than Comte had imagined, our ability to measure social phenomena accurately has been much more limited than he had envisaged, the sociological 'priesthood' has been influenced by its own prejudices as much as by the force of evidence, and our understanding of social causation has repeatedly been undermined by individuals choosing to act in ways that social scientists find unpredictable and often irrational. For all these reasons, the idea that social science might be able to come up with some core statistics that would force us to accept one vision of the 'good society' over another was widely attacked in the 1960s, and Comte's dream has largely been abandoned and forgotten since then.

Until now, that is.

The spirit level

In 2009, two British epidemiologists published a book which claimed to show that income inequality is the cause of many of our most pressing social problems. Their statistics apparently show that inequality undermines trust and community cohesion. It creates

10 On Comte, see Raymond Aron, *Main Currents in Sociological Thought* volume 1, Basic Books, 1967. Also Emile Durkheim, *The Rules of Sociological Method* Free Press, 1964

mental health problems, encourages drug abuse and undermines physical wellbeing. It reduces literacy and numeracy levels, increases the teenage birth rate and promotes violence and law-breaking. And what is more, these negative effects of living in an unequal society impact on everybody, not just those at the bottom, so all of us would benefit if inequality were reduced or eliminated.

The book was *The Spirit Level*, and its authors were Richard Wilkinson and Kate Pickett. They based their claims on two sets of statistics. One consisted of international data on 23 of the world's richest countries. The other was made up of data collected from the 50 US states. What is particularly impressive about the book is that, on almost all of the indicators that they examine, income inequality is found to correlate with social problems in *both* data sets. It is not just that more equal countries (like Sweden and Denmark) perform better on all these indicators than less equal countries (like the UK and the USA), but that more equal states within the USA (e.g. New Hampshire and Vermont) similarly perform better than the less equal states (such as Mississippi or Louisiana). On the face of it, income inequality looks the most plausible common cause across these two very different data sets.

Why should inequality be so damaging? Wilkinson and Pickett argue it is because living in rigid hierarchies is bad for human beings. Human beings lived for most of their existence in small hunter-gatherer bands which the authors believe were broadly egalitarian, and this means we are hard-wired for social equality. Put us in modern, competitive hierarchies, and at almost every point in the pecking order, we start to fret about self-esteem and we get nagging doubts about our self-worth. These stresses are then expressed in mental disorders, health pathologies and increased aggression towards others. Restoring us to our natural, egalitarian state will reduce these stresses and promote wellbeing: 'A less unequal society causes dramatically lower rates of ill-health and social problems because it provides us with a better-fitting shoe.'[11]

11 *The Spirit Level* p.213

The authors claim that we all stand to benefit from equalising incomes – even those at the top of the income distribution would gain, despite having to give up some of their money. This is because, beyond a certain point, increased wealth does not promote increased happiness or wellbeing. The authors accept that economic growth enhances people's quality of life in 'poor countries', but they insist that it has 'largely finished its work' in richer ones.[12] Once we have passed a certain threshold (which appears to be an average income somewhere between US$10,000 and US$25,000 per person per year), more money does not bring much more happiness or life quality.[13]

According to Wilkinson and Pickett, once a country reaches this level of prosperity, the way to improve people's lives is not by increasing the size of GDP, but by sharing out income more equitably. More equality, they say, would create better health outcomes for everybody (not just those at the bottom), and would generate a more cohesive and happier society from which we would all benefit: 'The vast majority of the population is harmed by greater inequality... the effects of inequality are not confined just to the least well-off.'[14]

It is this calculus of universal benefit that allows the authors to propose their empirical solution to the ethical dilemma that has divided the left and right for two centuries or more. They think their evidence proves that the left has been correct all along in what it says about inequality and that the right has been wrong. Because inequality is bad for all of us, it is in all our interests that it be decisively reduced: 'We need to create more equal societies able to meet our real social needs.'[15] There is no need to keep arguing about this, for the statistical evidence makes the case. The political debate is over, and the left won: 'The advantage of the growing body of evidence of the harm inflected by inequality is that it turns what were purely personal intuitions into publically demonstrable facts. This will substantially increase the confidence of those who have always shared these values and encourage them to take action.'[16]

12 *The Spirit Level* p.5

13 Wilkinson and Pickett do not spell out where exactly this income threshold lies, but reviewing international data on life expectancy and reported happiness, they provide graphs (Figures 1.1 and 1.2) indicating that diminishing returns kick in at around $10,000. By $25,000, the graphs 'flatten off' (*The Spirit Level* p.8).

14 *The Spirit Level* p.176

15 *The Spirit Level* p.231

16 *The Spirit Level* p.247

Not surprisingly, left-wing intellectuals and commentators have embraced *The Spirit Level*, welcoming it as the long-awaited empirical proof of their ideological assumptions. Praising the book's 'inarguable battery of evidence,' *The Guardian* noted: 'We know there is something wrong and this book goes a long way to explain what and why.'[17] *The Independent* thought the book's evidence was 'compelling and shocking' and argued that 'all free marketeers should be made to memorise it from cover to cover.'[18] In *The New Statesman*, former Labour Deputy Leader, Roy Hattersley, said the book 'demonstrates the scientific truth of the assertion that social democrats have made for a hundred years... that all of us, irrespective of income, have much to gain from the creation of a more equal society.'[19]

And at a packed Policy Exchange seminar, the social affairs journalist, Polly Toynbee, likened *The Spirit Level*'s principal author, Richard Wilkinson, to Charles Darwin, and referred to his 'discovery' of the deleterious effects of inequality as a 'Eureka moment' in the development of human thought.[20]

It is clear from comments like these that *The Spirit Level* is more than just an academic book. It is a manifesto. Its apparent 'scientific' backing for a core, traditional element of left-wing ideology is being used to spearhead a new political movement aimed at putting radical income redistribution back at the heart of the political agenda.

The principal instrument for this campaign is 'The Equality Trust', a not-for-profit organisation set up by the authors and other activists in 2009 with money from the Joseph Rowntree Charitable Trust. The formal aim of the Trust is 'to educate and campaign on the benefits of a more equal society.'[21] Its web site encourages people to establish 'Equality Group' branches, makes leaflets, posters and animated video available to those seeking to spread the message, and advises on setting up worker cooperatives and liaising with trade unions. At the 2010 General Election, the

17 Lynsey Hanley, 'The way we live now' *The Guardian* 14 March 2009

18 Yasmin Alibhai-Brown, 'In an unequal society, we all suffer' *The Independent* 23 March 2009

19 Roy Hattersley, 'Last among equals' *New Statesman* 26 March 2009

20 Speech to Policy Exchange seminar 'The future of the left', 18 March 2010

21 www.equalitytrust.org.uk

web site also ran an 'Equality Pledge' for supporters to get parliamentary candidates to endorse (430 candidates had signed it by election day), and it spawned a second organisation, *One Society*, with the specific task of influencing the election campaign.

As the wheels of this bandwagon gather pace, it is obviously crucial that the evidence on which this whole edifice has been built should be carefully analysed.[22] If it were true that inequality harms all of us, then it is difficult to see how anybody of good faith could remain loyal to core conservative principles like support for free enterprise, low taxes and the ideals of self-reliance. The case for using state power to bring about a radical redistribution of income and wealth would be unanswerable. But the book's core claims have not been seriously tested.[23] The book has been widely accepted without subjecting its statistics to rigorous scrutiny. So the simple yet crucial question we have to ask is: Is any of this true?

A note on methodology

In this report, I aim to replicate as far as possible the main findings reported in *The Spirit Level* in order to evaluate the claims the authors make.

In the case of the *international comparisons*, the authors have posted their data on their Equality Trust web site, so I have downloaded their data from there. However, their statistics are limited, for they exclude many countries which could be included, and they omit any indicator which does not vary with income inequality in the way the authors want. Some of their statistics are also out of date. I have therefore constructed my own data set, which I use in parallel with theirs when I need to take the analysis further than theirs allows.

To develop my version of this international data set, I have followed the same procedure that Wilkinson and Pickett outline in

22 Sadly, we cannot expect the academic establishment in Britain to do it, for they are almost universally wedded to the ideology that has spawned the book. On the political bias of sociology professors, see A. H. Halsey, *A History of Sociology in Britain* Oxford University Press 2004, Tables 8.5 and 8.6.

23 On 17 May 2010, as this report was undergoing its final edit, Christopher Snowdon's *The Spirit Level Delusion* was published by The Democracy Institute. On the web site that accompanies his book (http://spiritleveldelusion.blogspot.com/) Snowdon claims Wilkinson and Pickett 'rely on questionable data and misleading or obsolete statistics. When their graphs do reflect reality, the authors turn a blind eye to more plausible explanations and ignore evidence that would contradict their theory.' Snowdon's book therefore offers an exception to my comment here, and his critique appears broadly consistent with my own evaluation.

the Appendix of their book. Like them, I started by identifying the 50 richest countries in the world. Although some of these countries are much richer than others (they range from Norway at $53,433 per capita to Venezuela at $12,156), they have all achieved a level of GDP per head which should be sufficient, according to Wilkinson and Pickett, to secure an adequate level of individual and social wellbeing.[24]

Although Wilkinson and Pickett initially started out with a list of 50 countries, they ended up with a sample of only 23:

- This is partly because they deleted all countries with a population of less than 3 million, because they say they did not want to include 'tax havens.' However, this population cut-off looks unnecessarily severe, and it could safely be reduced to 1 million without picking up places like Monaco and the Cayman Islands. Adopting a 1 million population threshold has allowed me to reinstate six countries which they dropped unnecessarily (see Table 1).
- They also say they dropped any country where they could not find data on income distribution. This led to a further 21 countries being dropped. However, the 2004 UN *Human Development Report* (which is the source they say they used) contains income distribution data (both gini coefficients and percentile ratios) on all but two of the countries in the richest 50 (data are missing only on Libya and Saudi Arabia). A vast swathe of countries therefore appears to have been omitted from their sample even though the required data were available. The explanation for this is unclear.[25]

After reinstating some of the smaller countries and including all those where income data are available, I end up with a sample of 44 countries.[26] This is almost twice as many as Wilkinson and Pickett covered.

24 Remember that Wilkinson and Pickett also started out by selecting the 50 richest countries, which suggests they think they are all appropriate cases for testing their thesis. Moreover, Figures 1.1 and 1.2 in *The Spirit Level* plot life expectancy and happiness respectively against national income per head, and in both of these graphs, countries (like Venezuela and Turkey) which are the poorest in my sample appear above the crucial income threshold point in the graph.

25 Wilkinson and Pickett say only: 'We excluded countries without comparable data on income inequality' (p.275). But the data they say they could not find are provided in section 14 of the 2004 *Human Development Report*, pp.188-191. Note that for my sample, I use the 2009 report

26 I also retain Saudi Arabia and Libya for analyses not including income distribution. Four countries in the richest 50 were excluded because they have populations under 1 million. The Chinese autonomous regions of Hong Kong and Macau were also excluded as they are not sovereign countries and they were replaced by the next two countries on the GDP rankings.

Table 1: The expanded international sample of countries (in order of prosperity)[27]

Country	Population	GDP per capita	Included in *Spirit Level*	Added (pop > 1m and <3m)	Added (gini data)
Norway	4860100	53433	Yes	X	X
Singapore	4987600	49704	Yes	X	X
USA	308752000	45592	Yes	X	X
Ireland	4459300	44613	Yes	X	X
Switzerland	7779200	40658	Yes	X	X
Netherlands	16592550	38694	Yes	X	X
Austria	8372930	37370	Yes	X	X
Sweden	9340682	36712	Yes	X	X
Denmark	5534738	36130	Yes	X	X
Canada	34013000	35812	Yes	X	X
UK	62041708	35130	Yes	X	X
Belgium	10827519	34935	Yes	X	X
Australia	22166000	34923	Yes	X	X
Finland	5356300	34526	Yes	X	X
Germany	81757600	34401	Yes	X	X
France	65447374	33674	Yes	X	X
Japan	127430000	33632	Yes	X	X
Spain	45989016	31560	Yes	X	X
Italy	60275846	30353	Yes	X	X
Greece	11306183	28517	Yes	X	X
New Zealand	4357700	27336	Yes	X	X
Slovenia	2053750	26753	No	Yes	X
Israel	7509000	26315	Yes	X	X
South Korea	49773145	24801	No	No	Yes
Czech Republic	10512397	24144	No	No	Yes
Trinidad & Tobago	1339000	23507	No	Yes	X
Saudi Arabia*	25721000	22935	No	No	No

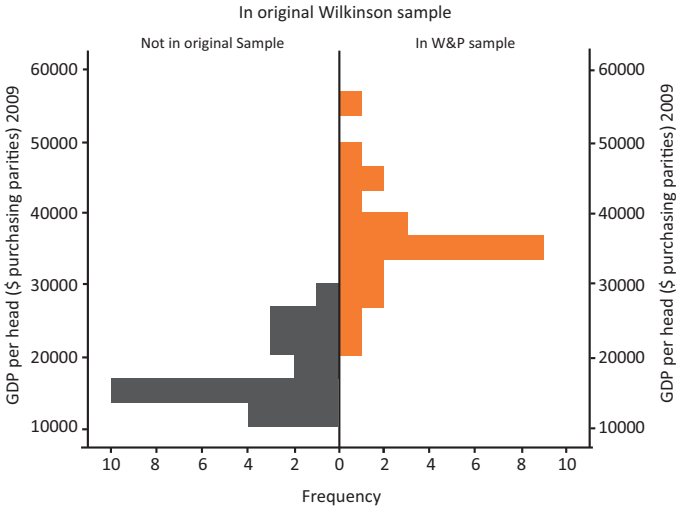
Country	Population	GDP per capita	Included in *Spirit Level*	Added (pop > 1m and <3m)	Added (gini data)
Portugal	10636888	22765	Yes	X	X
Estonia	1340021	20361	No	Yes	X
Slovakia	5421937	20076	No	No	Yes
Hungary	10013628	18755	No	No	Yes
Lithuania	3329227	17575	No	No	Yes
Latvia	2248400	16377	No	Yes	Yes
Croatia	4435056	16027	No	No	Yes
Poland	38100700	15987	No	No	Yes
Gabon	1475000	15167	No	Yes	Yes
Russia	141927297	14690	No	No	Yes
Libya*	6420000	14364	No	No	No
Mexico	107550697	14104	No	No	Yes
Chile	17038000	13880	No	No	Yes
Botswana	1950000	13604	No	Yes	Yes
Malaysia	28306700	13518	No	No	Yes
Argentina	40134425	13238	No	No	Yes
Turkey	72561312	12955	No	No	Yes
Romania	21466174	12369	No	No	Yes
Venezuela	28676000	12156	No	No	Yes

*Libya and Saudi Arabia are excluded from the main sample but may be included in analyses not requiring income distribution statistics (see note 26)

It is clear from Table 1 that most of the countries omitted by Wilkinson and Pickett are less prosperous than those they included, although 5 of the 23 that I have reinstated are richer than Portugal, which they included in their original sample. Figure 3 summarises the prosperity of the countries in the original Wilkinson and Pickett sample (on the right of the chart) as compared with those I have added (on the left).

27 Population data and GDP per head in $ purchasing power parities from UN *Human Development Report* 2009

Figure 3: GDP per head of countries included and excluded by Wilkinson and Pickett



Most of the countries that have been reintroduced into my expanded sample are based in regions of the world outside Western Europe and North America. The sample on which *The Spirit Level* is based is culturally quite narrow, comprising 16 western European countries, 2 from North America, and 5 from Asia, Australasia and the Middle East. My expanded sample adds 11 new countries from the former Soviet bloc, 3 more countries from Asia, 2 from Africa, 3 from South America, and 2 from Central America and the Caribbean. This expanded and much more culturally diverse sample provides an opportunity to test more thoroughly the authors' claim that it is income inequality, and not culture or history, that explains the patterns they identify.

In the case of the *US state comparisons*, Wilkinson and Pickett's data set was not available for downloading at the time of writing. I have therefore built a replica data set which matches their sources as far as possible, updating where newer statistics are available. Where my sources vary from theirs, I note this in the text. As with the international data, I have also added more indicators to my version, and this will allow us to examine the influence on social outcomes of factors like ethnic composition, which they left out.

Wilkinson and Pickett measure their key explanatory variable, income inequality, in a different way when they look at the US states than in their international comparisons:

- For the US states they use the gini coefficient;
- For their international comparisons, they use a ratio of the income of the lowest 20% as compared with that of the highest 20%.

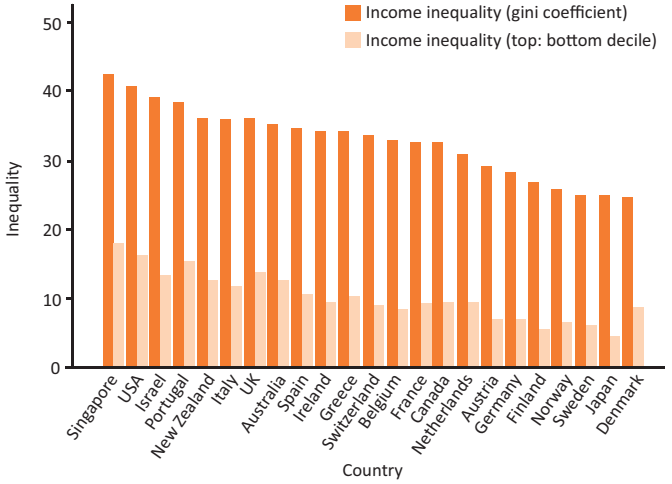
They never explain why they use different measures for each of these two samples.[28] The gini coefficient is, as they say themselves, 'the most common measure,' it is 'more sophisticated' than the simple ratio measure they adopt for their international comparisons, and it is available in the UN data tables.[29] Wilkinson and Pickett say it makes little difference which measure of income inequality they use – 'the choice of measures rarely has a significant effect on results'[30] – and we can see from Figure 4 that this is broadly true (although Britain comes out rather more unequal in international comparisons using the ratio measure than on the gini coefficient). But for the sake of consistency, I shall use the gini coefficient to analyse my international data (as well as the US state data). When re-analysing their international data set, however, I shall be using their 80/20 income ratio measure, for they do not provide gini coefficients in their spreadsheet.

28 Nor do they explain why they chose the 80/20 ratio measure for the international data when 90/10 or 50/10 are much more usual.

29 *The Spirit Level* pp.17-18

30 *The Spirit Level* p.18

Figure 4: Income inequality in the countries included in *The Spirit Level* measured by two different statistics



For the analysis itself, I shall follow Wilkinson and Pickett in using simple correlation and regression procedures.[31] Like them, I shall produce a series of graphs (called 'scatterplots') in which income inequality is plotted on the horizontal (x) axis, and whichever dependent variable we are interested in (life expectancy, homicide rate, obesity, or whatever) is plotted on the vertical (y) axis. We shall gauge the strength of association between two variables by fitting a straight trend line (a 'least squares' or 'regression' line) to each graph so that it minimises the total distance between all the points and the line.

Wilkinson and Pickett use Pearson's correlation coefficient (r) as a summary statistic to express the strength of association between variables. I shall generally use a related statistic called 'the coefficient of determination', or R^2.[32] The coefficient of determination is calculated by squaring the correlation coefficient, and it tells us

31 For an explanation of correlation and least squares regression, see Alan Buckingham and Peter Saunders, *The Survey Methods Workbook* Polity Press, 2004

32 When analysing the international data I shall use the 'Adjusted R^2' which corrects for a tendency for 'goodness of fit' statistics to be slightly exaggerated when calculated for samples. For the analysis of the US states (a whole population, rather than a sample) I shall use the unadjusted R^2

what proportion of variance in the dependent variable is accounted for by the independent variable. A significance test, called an *F test*, will be used to determine whether any association we find between variables is likely to have occurred by chance. Like Wilkinson and Pickett, we shall only accept as 'significant' patterns which are unlikely to arise by chance more than 5 times in 100 samples (designated as $p<0.05$).

In one important respect, our use of regression modelling will go beyond Wilkinson and Pickett. They content themselves with plotting simple graphs depicting the association between two variables, A and B. It is, however, common practice in social science to look out for the effects of third variables. Third variables may disguise an association between A and B, or they may generate an apparent association where in reality there is none. Controlling for third variables can, therefore, be crucial when interpreting data, and we shall encounter several examples in this report (particularly in Chapter III, where we look at the US states) where Wilkinson and Pickett claim to have found an association between inequality and some social outcome when in fact the association is being driven by a different variable which they have altogether overlooked.

Regression – whether bivariate or multivariate – is a powerful statistical procedure, but it does require certain conditions to be met. In *The Spirit Level*, Wilkinson and Pickett are alarmingly cavalier in their approach, but we shall be more cautious. In particular, we shall need to look out for 'outliers' (extreme cases which can distort the slope of a trend line), and check whether it is appropriate to fit a straight trend line to our scatterplots (testing the 'linearity' assumption). We shall also need to inspect how the values of our dependent variables are distributed at different values of the independent variable, for regression analysis entails certain assumptions about normality and equality of variance which appear to be violated in some of the scatterplots in *The Spirit Level* (I shall explain this problem in more detail when we encounter specific instances of it).

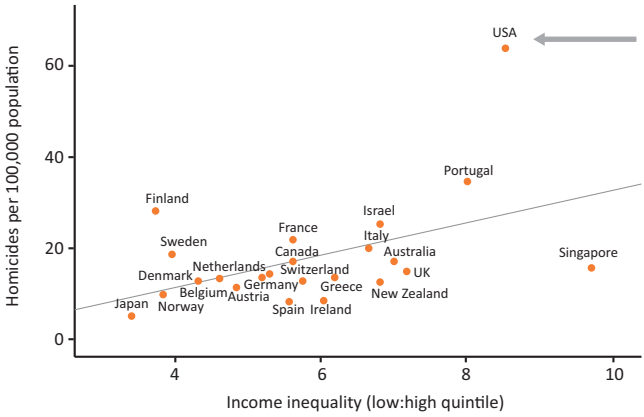
2. Are less equal societies more dysfunctional societies?

Let us start by looking at Wilkinson and Pickett's international comparisons. Like all the best investigations, we can begin with murder.

2.1 Homicide rates

In *The Spirit Level*, Wilkinson and Pickett show an apparent association between the level of income inequality in a country and its homicide rate. Using their data, I reproduce their graph as Figure 5a.

Figure 5a: Wilkinson and Pickett's plot of inequality against homicide rates[33]



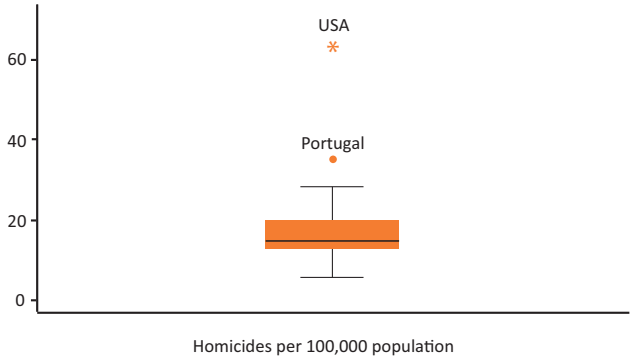
33 Based on Fig 10.2 of *The Spirit Level*, recreated from the international data set downloaded from The Equality Trust web site.

They report a moderately strong correlation of 0.47, which translates into a R^2 of 0.22. This means that inequality explains 22% of the variance in the homicide rates of different countries. This association is found to be statistically significant (p=0.025).

But look at the scatter of the countries on the vertical (y) axis in Figure 5a. Most of them seem to have homicide rates which are compressed in a range between about 10 and 20 murders per 100,000 population. The glaring exception is the USA (flagged by an arrow), with its homicide rate of over 60 per 100,000. Judging by this graph, we might suspect that the USA is a unique case, and that its exceptionally high homicide rate is being caused by factors which are specific to that one country alone (the laxity of gun control laws is an obvious possible explanation).

There is a simple test we can run to detect what statisticians call 'outliers' in any distribution of data. It is called a 'boxplot', and it provides a visual representation of how cases are distributed on any given variable. In Figure 6, I reproduce the boxplot for Wilkinson and Pickett's homicide data.

Figure 6: Boxplot of Wilkinson and Pickett's international homicide data, showing USA as an 'extreme outlier'



There is no need to go into the details of how to interpret a boxplot, other than to note that 'outliers' are identified by a circle, and 'extreme outliers' are identified by an asterisk.[34] We can see from this example that Portugal is an 'outlier', and the USA is an 'extreme outlier' when it comes to murder rates.

Outliers and extreme outliers can cause serious problems in regression analysis, particularly when we are dealing (as here) with a relatively small number of cases, for just one or two extreme points can skew an entire graph. The point of constructing a graph like the one in Figure 5a is to see if there is an association across *all countries* between inequality and the murder rate, but the trend line here is clearly being pulled upwards by just one extreme case – the USA – whose exceptionally high murder rate might have nothing to do with its level of income inequality.

Throughout *The Spirit Level*, Wilkinson and Pickett routinely ignore this problem of outliers (we shall see that they discuss outliers only once, and that is in discussion of a graph where outliers do not skew their results). In their analysis of homicide rates, they argue that murders are more common in more unequal countries, but for this to be true, they should have asked whether this association still holds among the other 22 countries in their sample if the USA is taken out of the picture.

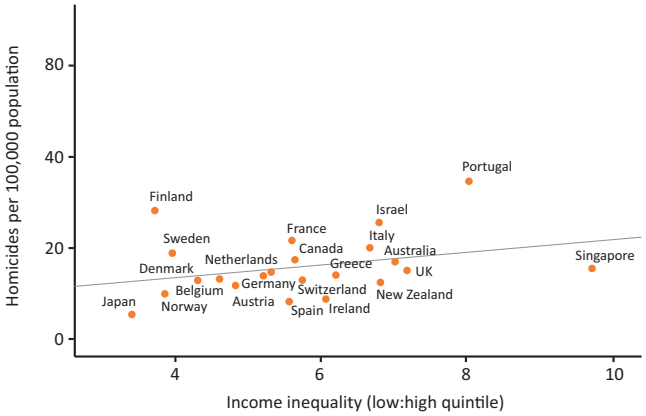
Figure 5b shows that it does not. Here we have the same plot as in Figure 5a, but without the USA. Note that one of the most equal countries in the sample (Finland) has one of the highest murder rates, and one of the most unequal (Singapore) has one of the lowest murder rates. Similarly, Australia has fewer murders than Sweden, and the UK has fewer than Denmark, yet the Anglophone countries are more unequal than the Scandinavian nations. There does not appear to be any clear association between income inequality and homicide.

Comparison of the slope of the line in Figure 5b against the slope in Figure 5a shows the effect that removing just one outlying

34 The values (i.e. the number of homicides per 100,000 population) are listed on the left-hand vertical axis. The grey-shaded box represents half of the countries in the sample, ranging from the country at the 25th percentile to the country at the 75th percentile. The heavy black line through the box is the median case (the fact that it is near the bottom of the box tells us that these data are positively skewed). The lines running vertically above and below the box, known as 'whiskers', extend between the lowest and highest values which are not considered to be 'outliers.' Outliers are defined as cases which are more than 1.5 box-lengths above the 75th percentile or below the 25th percentile, and 'extreme outliers' are more than 3 box-lengths away.

case can have on a graph like this. Indeed, once the USA is omitted, it is not really appropriate to fit a regression line at all, for there is no longer any association to determine. The R^2 value has fallen to just 0.10, and the relationship fails to achieve statistical significance (p=0.159).[35]

Figure 5b: Wilkinson and Pickett's plot of inequality against homicide rates, excluding the USA



It is clear from all this that there is no association between income inequality and homicide rates, despite Wilkinson and Pickett's confident claim that there is.[36]

Conclusion: There is no evidence of a significant association between the level of income inequality in a country and its homicide rate.

2.2 Conflict in childhood

Wilkinson and Pickett follow their analysis of homicide rates with a graph of 'children's experience of conflict.' This is based on an index which they construct based on children's reports that (a) they have

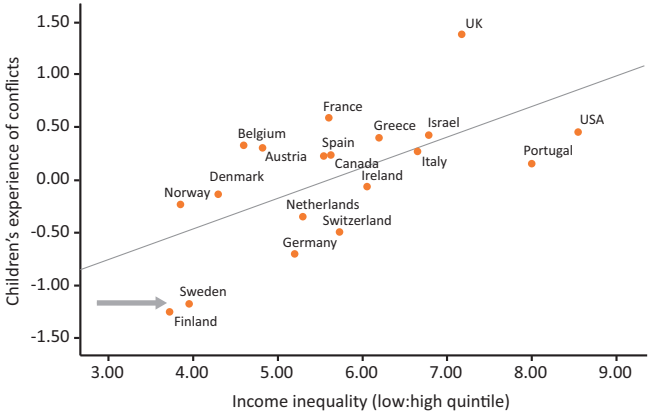
35 Given that Portugal is also identified as an outlier in Figure 6, there is a case for omitting it too. If this is done, the association between inequality and homicide in the remaining 21 countries falls even further, with an R^2 of just 0.02 and a significance level of p=0.533.

36 The homicide statistics in Wilkinson and Pickett's data set are actually quite old, dating from 1999-2000. If we substitute more recent figures (available from the World Health Organisation web site: http://www.who.int/whosis/en/index.html), Portugal no longer appears as an outlier, but the USA remains an extreme outlier. Re-running the same regression model using these more recent data (and omitting the USA) again gives an insignificant result (R^2=0.03 and p=0.428).

been involved in fights, (b) they have been victims of bullying, and (c) their peers are 'not kind and helpful.'

We might ask whether this index makes much substantive sense as a measure of conflict (is having 'unhelpful friends' really the same as being bullied?). But leaving such concerns aside, the authors say they find an association between the average scores of each country on this index and their level of income inequality. The relevant graph is reconstructed from Wilkinson and Pickett's data as Figure 7a. The association they find appears reasonably strong (R^2=0.35) and is statistically significant (p=0.004).

Figure 7a: Wilkinson and Pickett's plot of the association between inequality and children's experience of conflict[37]



A boxplot indicates that the UK is an outlier on this measure, but even if Britain is removed from the analysis, the association with inequality still holds.[38] This time, therefore, their graph looks legitimate.

But there is another problem, and it is one that crops up time and again in the pages of *The Spirit Level*.

37 Based on *The Spirit Level*, fig 10.4 and recreated from the international data set downloaded from The Equality Trust web site.

38 Without the UK, the adjusted R^2 = 0.305, F=8.445, sig= 0.010.

The Scandinavian nations (in particular in this case, Sweden and Finland, marked on Figure 7a by an arrow) appear to have relatively low levels of childhood conflict, while the Anglophone nations (in particular, the USA and the UK) exhibit relatively high levels. On this measure at least, it is clear that British children are more involved in conflict and violence than Scandinavian children. The question, however, is whether this is because the Anglo countries are more unequal than the Scandinavian countries, or because of cultural differences between them. We shall see later in this report that the Anglo countries are more individualistic cultures, while the Scandinavian countries have been more culturally homogenous. We need somehow to take account of these cultural and historical differences before we can isolate the effect (if any) of variations in income inequality.

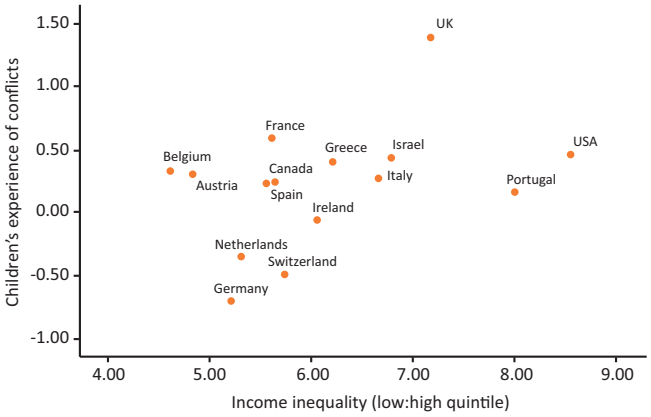
“The Scandinavian nations appear to have relatively low levels of childhood conflict, while the Anglophone nations exhibit relatively high levels”

The only way to be confident that inequality really is the explanation is to see whether the association between inequality and conflict exists across other countries as well. In other words, take out the Scandinavians and/or the Anglo nations and see whether the association still holds. If the Wilkinson and Pickett hypothesis is true – that childhood conflict rises as inequality intensifies – then we should also find evidence for this when we focus on countries like Germany, Canada, Greece, France and Israel.

But when we take out the Scandinavians, the apparent association between inequality and childhood conflict collapses. Even with the outlying UK restored to the analysis, Figure 7b reveals that there is no significant association without the influence of the Nordic countries. Egalitarian Belgium performs almost as badly as inegalitarian America; highly-equal Austria performs worse than highly-unequal Portugal; and so on. The apparently strong correlation coefficient reported in *The Spirit Level* has vanished.[39]

39 Adjusted R^2= 0.08, F= 4.453, p= 0.157

Figure 7b: The association between inequality and children's experience of conflict, without the Scandinavian countries



This is a pattern we shall encounter repeatedly as we interrogate Wilkinson and Pickett's 'findings'. There is no doubt that, on many of their preferred indicators, the Scandinavians perform better than the Anglophone countries. But we shall see in Chapter IV that the explanation for this almost certainly lies in deep-seated historical and cultural differences between them, rather than in their differing levels of income inequality. If inequality really were the explanation for these differences, it should have an effect on all the other countries as well. But in this and in many other graphs reported in *The Spirit Level*, it does not.

Conclusion: Wilkinson and Pickett's data show no association between inequality and childhood conflict. The statistics only show that Scandinavia has low levels of conflict.

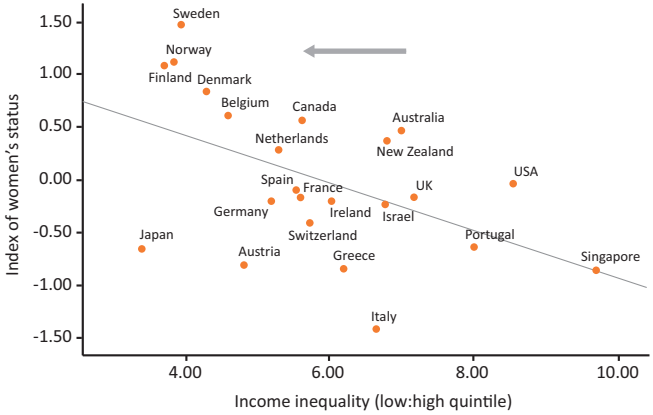
2.3 Women's status and national generosity

Two more examples of this same problem occur in the chapter in *The Spirit Level* dealing with 'community life and social relations' where

the authors suggest that more equal countries treat women better and are more generous in their aid to poor countries.

To measure women's status, Wilkinson and Pickett construct another of their indexes. This time, they measure women's status by the number of female politicians, the number of women in the labour force, and their assessment of the 'social and economic autonomy' enjoyed by women. They find a modest but significant association between countries' scores on this index and income inequality (their graph is recreated as Figure 8a).[40]

Figure 8a: Wilkinson and Pickett's plot of inequality against women's status[41]

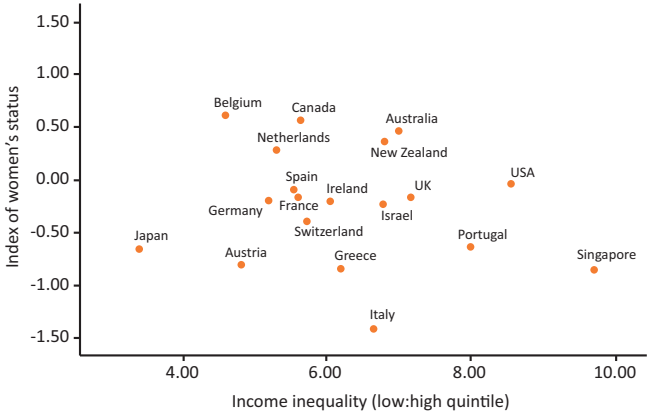


Not surprisingly, given their comparatively high women's employment rates and their parties' use of political gender quotas, the Nordic states stand out on this index. Looking at Figure 8a, we see all four of them clustering in the top left-hand quadrant of the graph (identified by the arrow). Once again, this clustering raises the suspicion that it is the Nordic states alone that are generating Wilkinson and Pickett's 'finding,' and this is confirmed when we run the same graph again, but this time without the Scandinavians (Figure 8b).

40 For some reason the corresponding graph in the book omits Singapore, so the model fit statistics I get are slightly different – and more supportive of the authors – than those reported in the book. With Singapore included, the adjusted R^2 = 0.211, F= 6.877, p = 0.016; without Singapore, R^2 = 0.153, F = 4.803, sig = 0.400.

41 Based on *The Spirit Level*, fig 4.5, and recreated from the international data set downloaded from The Equality Trust web site.

Figure 8b: Wilkinson and Pickett's plot of inequality against women's status, excluding the Scandinavian countries



In Figure 8b there is no association. The R^2 is close to zero and it falls a long way short of statistical significance.[42] Women fare just as badly in egalitarian countries like Japan and Austria as they do in inegalitarian ones like Portugal and Singapore. There is, quite simply, no relationship between income distribution and women's status.

The Spirit Level also claims that 'more equal countries are also more generous to poorer countries.'[43] What the authors mean by this is that the *governments* of more equal countries are more generous (i.e. they give more foreign aid than other governments do). It seems that in their minds, the generosity of a government can be equated with the generosity of a country. They want to show that equal societies are more caring and warm-hearted, and they do it by estimating the compassion of people from the spending patterns of their governments.

The evidence they muster for their proposition is presented in Figure 9a. The association between inequality and the size of

42 Adjusted R^2 = -0.030, F = 0.469, p = 0.503

43 *The Spirit Level*, p.60

foreign aid budgets is reasonably strong and clearly significant.[44] But as with previous graphs, it rests entirely on the contribution of the Scandinavian countries (marked by the arrow).

Figure 9a: Wilkinson and Pickett's plot of inequality against government foreign aid[45]

This time, it is Norway, Sweden and Denmark that cluster (together with the Netherlands) at the top of the y axis. With their foreign aid spending running at around 0.09% of national income, these four are way above all the other countries in the graph, most of which are concentrated in a band around 0.03 to 0.05 per cent.

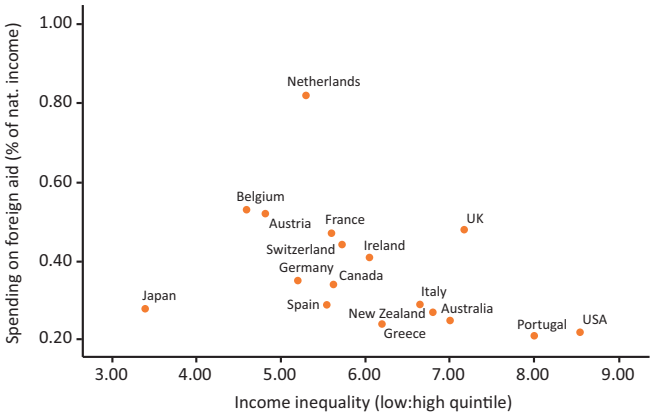
It is clear just by looking at this graph that the effect is being generated by the Scandinavians alone. Take out the four Nordic nations (Figure 9b) and there is no significant association with inequality.[46] Japan – the most equal country in this sample – has about the same per capita foreign aid budget as the USA – the least equal – and the rest are scattered randomly between them.

44 Adjusted $R^2 = 0.340$, $F = 11.320$, $p = 0.003$

45 Based on *The Spirit Level*, fig 4.6, and recreated from the international data set downloaded from The Equality Trust web site.

46 Adjusted $R^2 = 0.121$, $F = 3.198$, $p = 0.094$

Figure 9b: Wilkinson and Pickett's plot of inequality against government foreign aid, excluding Scandinavia



We can go further. If we really want to gauge the generosity of a country, we should look not at how much compulsorily-levied tax money its government gives away, but at how much of their own money individual citizens are willing to give away voluntarily.

Data are available on charitable giving for 11 of the countries in my expanded international data set. For these 11 countries, there is no association between income inequality and per capita charitable donations.[47] Wilkinson and Pickett's hypothesis does not stand up. It is also striking that the most generous country by far is the USA, with Britain and other Anglophone countries also performing creditably (Figure 10).[48] France and Germany, which are both much less unequal than the USA, are more than six times less generous when it comes to voluntary donations.

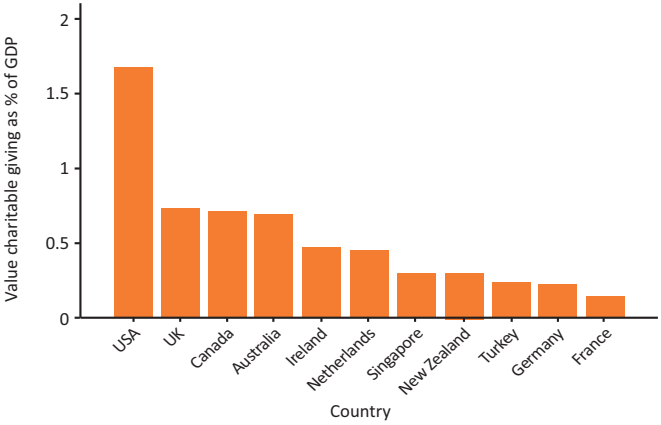
Statistics on charitable donations are unfortunately unavailable for many countries, and there are no figures for any of the Scandinavian nations. However, we can look at active membership

47 Adjusted R^2 = -0.042, p= 0.460

48 According to Hofstede's Individualism Index Value, the Anglophone countries are the most individualistic in the world (Geert Hodstade, *Culture's Consequences*, Sage, 2nd edn, 2001, p.215). The data on charitable giving suggest that generosity might be higher in more individualistic cultures, although on this small sample of 11 cases, the association falls just short of statistical significance (p = 0.088 with adjusted R^2 = 0.21).

of charities and humanitarian organisations as a proxy measure. This time, statistics are available for 26 of the countries in my expanded international data set, and they include three of the Scandinavian nations as well as egalitarian Japan.

Figure 10: Private charitable donations as a percentage of GDP (selected countries)[49]

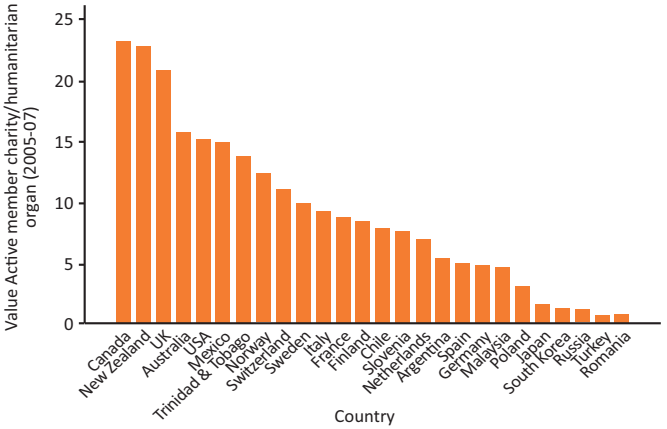


The pattern we get (Figure 11) is broadly similar to that found for charitable donations: the Anglophone countries are by far the most active. On this measure, the egalitarian Scandinavians come out as moderately active, but the egalitarian Japanese are more than 10 times less likely to get involved in a charity than the inegalitarian Kiwis and Brits.

To claim, as Wilkinson and Pickett do, that the most equal countries are the most generous is therefore untrue. The generosity of the people has nothing to do with how much their politicians spend, and when it comes to voluntary donations and voluntary activity, the Anglophone cultures appear to be the most generous in the world.

49 Source: 'International comparisons of charitable giving, November 2006' Charities Aid Foundation *Briefing Paper*, London, 2006

Figure 11: Active involvement in charities and humanitarian organisations (selected countries)[50]



Conclusion: Neither women's status, nor foreign aid, are affected by the degree of income inequality in a country. More unequal Anglophone countries appear the most generous.

2.4 Trust

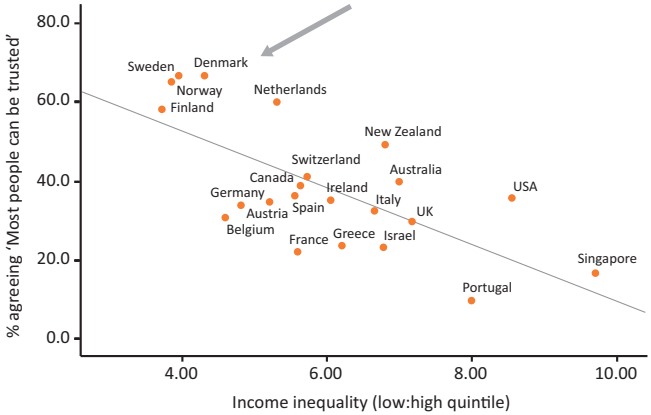
In his celebrated book, *Bowling Alone*, Robert Putnam argues that trust is a crucial component of 'social capital.'[51] The more people trust each other, the greater is the potential for them to co-operate, which means less time and money has to be spent encouraging or forcing them to pull together. Putnam suggests that social capital varies with the distribution of income and wealth because inequality undermines trust and destroys empathy.

Wilkinson and Pickett agree with this. They use data from the *World Values Survey* in which respondents in different countries are asked if they believe 'most people can be trusted,' and they compare the answers with the level of income inequality in each country. Their result is reproduced as Figure 12a.

50 Source: *World Values Survey* 2005-07 http://www.worldvaluessurvey.org/

51 Robert Putnam, *Bowling Alone: America's Declining Social Capital* Simon & Schuster, 2000

Figure 12a: Wilkinson and Pickett's plot of inequality against trust



There appears to be a strong negative association between the two variables (the adjusted R^2 is 0.442, suggesting that 44% of the variance in trust can be explained by income distribution, and the significance level is better than 0.001). As inequality rises, trust levels fall.

Clearly the four Nordic countries are once again influencing this finding, for together with the Netherlands, they form a distinct cluster in the top left quadrant of the graph (indicated by the arrow). But on this occasion, even if we take the Scandinavians out of the analysis, this association still stands up (Figure 12b). It is much weaker, to be sure (the adjusted R^2 falls dramatically to just 0.169, and the association is now only marginally significant – p=0.046). But there does still seem to be a weak pattern across the sample as a whole, and Wilkinson and Pickett suggest that other studies have similarly reported that more equal countries tend to be more trusting, so we should accept it as a valid result.

Figure 12b: Wilkinson and Pickett's plot of inequality against trust, excluding Scandinavia

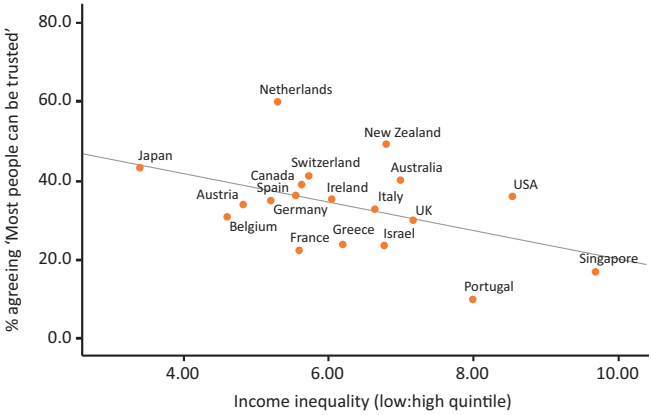
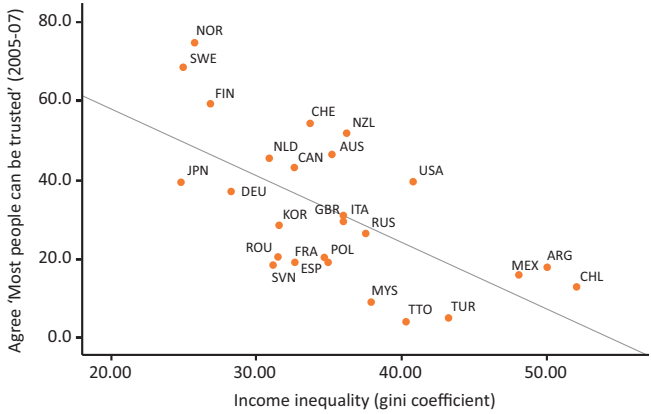


Figure 12c: The association between income inequality and trust in an expanded sample of 26 countries[52]



52 In this graph, the trust data are taken from a more recent (2005-07) *World Values Survey*, and income inequality is measured by the gini coefficient, averaged over the period 1992-2007, as reported in the 2009 UN *Human Development Report*.

In my expanded data set, the relevant data on trust are available for 26 countries (Figure 12c). Here too we find a reasonably strong pattern (the adjusted $R^2 = 0.382$, and the significance level is better than 0.001), further reinforcing the conclusion that inequality and trust seem to co-vary. Again, the Scandinavian countries form a distinct cluster, but even if the Scandinavians are taken out, the association still holds, albeit less convincingly.[53]

This is not quite the end of the story, however. Looking at Figure 12c, it is obvious that many of the countries with the lowest levels of trust (Turkey, Chile, Trinidad, Mexico) are among the poorest nations in our sample, and they are often also among the least equal. It will be recalled that Wilkinson and Pickett insist that, beyond a certain point, wealth is irrelevant in influencing the quality of life, but it seems from Figure 12c that, as far as levels of trust are concerned, this may not be the case.

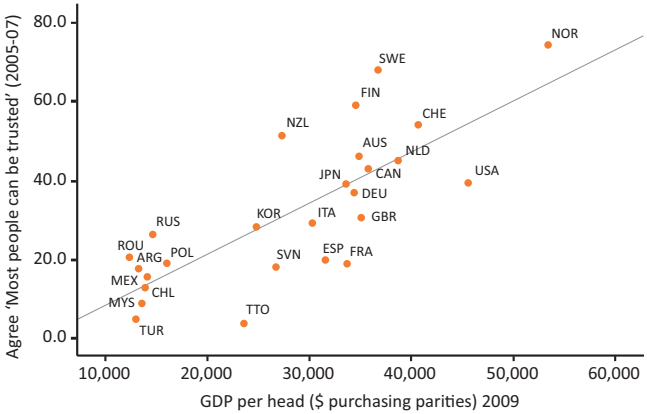
We can test for this by looking at two pieces of information simultaneously – the level of inequality in a country, and its wealth (measured by GDP per head of population). When we do this (in a multiple regression model), we get a good, strong predictive model accounting for almost two-thirds of the variance in countries' trust levels, but we find that most of the explanatory work is being done by GDP, not by income inequality.[54] GDP has more than twice the impact on trust levels than income inequality does. Indeed, inequality ceases to achieve statistical significance once GDP is taken into account. As Figure 12d shows, we can construct a strong model predicting trust from GDP alone – we do not need to add information about the income distribution of these countries.

It seems from all this that inequality may have some association with trust, but prosperity matters more. As countries get wealthier, trust levels increase (which runs counter to Wilkinson and Pickett's argument that GDP ceases to have an impact on wellbeing once living standards reach an adequate level).

53 The adjusted R^2 without the Nordic countries = 0.199, p = 0.019.

54 The adjusted R^2 for this multiple regression model is 0.620, with $p<0.001$. The Beta coefficient for GDP per head is 0.620 (p=0.001), as compared with -0.261 (p=0.106) for income inequality.

Figure 12d: The association between trust and GDP per head in an expanded sample of 26 countries[55]



Conclusion: Trust may have some weak association with inequality, but the effect of GDP is greater.

2.5 Social capital: other measures

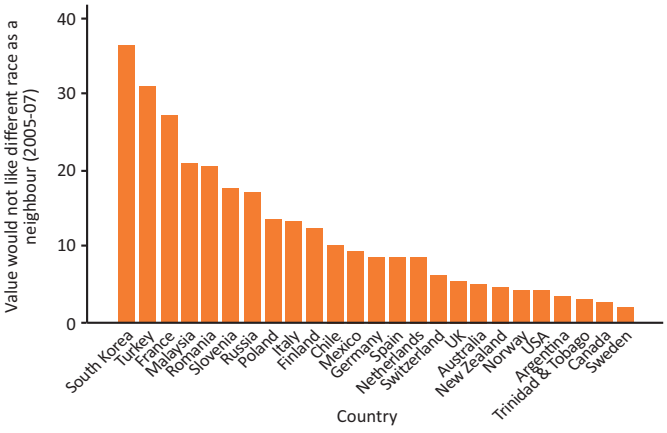
Although Wilkinson and Pickett talk approvingly of Putnam's concept of 'social capital', the only measure of social capital they use in their book relates to the question whether 'most people can be trusted.' But what happens if we measure social capital by other, equally important, indicators?

One aspect of social capital is racial tolerance. Countries where members of different ethnic groups get along with each other without too much friction are clearly more cohesive and more harmonious than those where ethnic barriers restrict cooperation. Yet when we measure ethnic tensions, the more unequal Anglo countries come out better than many of the more equal nations.

55 GDP per head is measured in US $ Purchasing Power Parities and is taken from the UN *Human Development Report* 2009. Adjusted $R^2 = 0.590$, $p < 0.001$

Asked if they would mind having someone of a different race as a neighbour (Figure 13a), for example, the Brits, Americans and Australians appear much more comfortable than their more egalitarian French, Italian, Finnish, German, Spanish, Dutch and Swiss contemporaries. The average French person is five times more likely to object to an ethnic minority neighbour than the average American. Even if we suspect that many respondents in the Anglo countries are too embarrassed to say what they really think when they answer this question, these results are still significant, for they give us an insight into the acceptability or otherwise of expressing racial bigotry in these different countries.

Figure 13a: Opposition to having a neighbour from a different ethnic group[56]



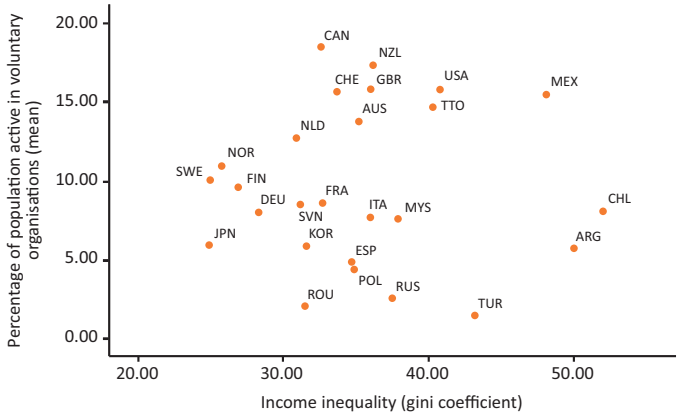
Another important indicator of social capital is voluntary membership of non-governmental organisations (indeed, for Putnam this is a key indicator, for this tells us something about the vitality of civil society). My expanded data set includes information from 26 countries on the proportion of the population actively

56 Source: *World Values Survey* 2005-07

involved in voluntary organisations, and from this, it is possible to construct a simple average of organisational activity in each country.[57] Figure 13b plots each country's score against its level of income inequality.

According to Wilkinson and Pickett's hypothesis in *The Spirit Level*, we should expect voluntary activity in informal organisations to correlate strongly with income equality, for it is a key indicator of social capital. But Figure 13b shows there is no association.[58]

Figure 13b: Voluntary organisational memberships against income inequality (26 countries)



57 The *World Values Survey* provides data on active involvement in religious organisations, sports and recreational organisations, arts and culture organisations, trade unions, professional associations, political parties, environmental groups, and charitable and humanitarian groups. I have created a single indicator of organisational involvement by summing the percentages involved in each type of body and dividing by 8 (the number of organisational categories) to get an average measure of active participation in civil society. Income inequality is measured by the gini coefficient.

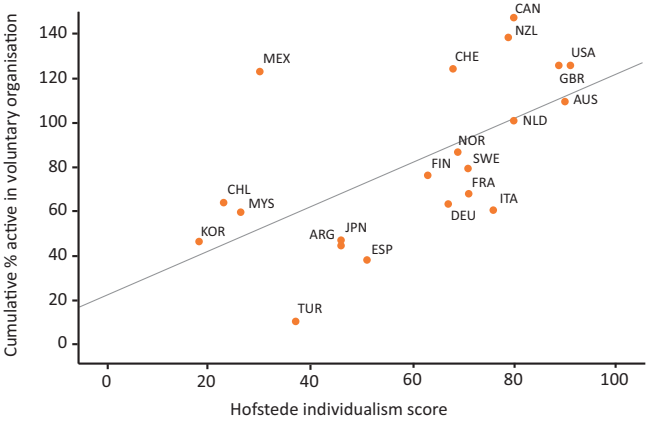
58 Adjusted R^2 = -0.041, p = 0.885

Where we do find an association is not with inequality, but with individualism. In a major study of national cultures carried out 40 years ago, Geert Hofstede analysed the attitudes of more than 100,000 people across 72 different countries, from which he identified a number of distinctive 'cultural dimensions' on which countries vary from one another. One of these dimensions was 'individualism.' In individualistic cultures, he wrote, 'The ties between individuals are loose: everyone is expected to look after

him/herself and his/her immediate family only.'[59] These are precisely the kinds of countries that Wilkinson and Pickett are trying to warn us about, and the Anglo nations come out as the highest scorers on Hofstede's Individualism Index.

Yet when we plot countries' individualism scores against their levels of voluntary organisational activity, we find that far from being fragmented, chaotic places (as the authors of *The Spirit Level* would have us believe), it is the highly individualistic countries which have the most active joiners (Figure 13c). Returning for a moment to our earlier discussion of charitable donations, we find a similar pattern emerging there too – the most generous donors tend to be the most individualistic countries.[60] These findings are not being generated by a third variable, such as GDP, for if we run a multiple regression model predicting organisational activity from both individualism and GDP, it is the former that proves significant while the latter adds nothing to our explanation.[61]

Figure 13c: The more individualistic countries have higher levels of voluntary organisational activity[62]



59 Hofstede, *Culture's Consequences*, p.225

60 The data on charitable giving suggest that generosity might be higher in more individualistic cultures, although on my small sample of just 11 cases, the association falls just short of the 5% threshold for statistical significance ($p = 0.088$ with adjusted $R^2 = 0.21$).

61 Beta (individualism) = 0.648 (p=0.032); Beta (GDP) = -0.046 (p=0.852)

62 Adjusted $R^2 = 0.334$, $p = 0.004$

It seems that egalitarianism is not the only route to social cohesion after all. People come together and cooperate in more individualistic countries too.

Anyone familiar with the classic works of sociology would not be surprised by this result, for it confirms Emile Durkheim's essential insight more than one hundred years ago that societies based on individualistic values can be *more* cohesive than those based on norms emphasising collective conformity.[63] Durkheim understood that social cohesion does not depend on everyone feeling the same as everyone else. All that is necessary for mutual association is that we should believe we can benefit from each other's assistance in achieving our various goals, for this drives us to cooperate through what Edmund Burke called the 'little platoons' of civil society. Provided there are tasks that need doing, or challenges that need meeting, people will come together spontaneously to resolve them. We do not need governments to redistribute our incomes in order to make us want to co-operate.[64]

"Provided there are tasks that need doing, or challenges that need meeting, people will come together spontaneously to resolve them"

Conclusion: Some measures of social capital, including ethnic tolerance and active involvement in voluntary organisations, do not vary with income inequality, but they are higher in more individualistic national cultures.

2.6 Physical health (life expectancy and infant mortality)

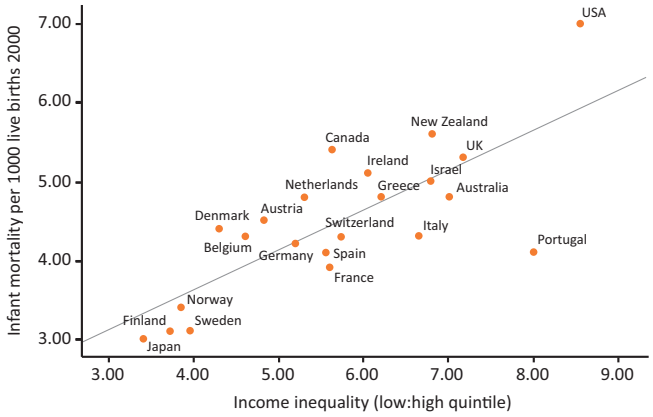
Both authors of *The Spirit Level* are epidemiologists, and Richard Wilkinson has spent many years studying health inequalities. Not surprisingly, the claim that inequality is bad for your health lies at the heart of their book, and it rests on two related indicators: life expectancy and infant mortality rates. The evidence on the latter appears rather more compelling than the evidence on the former.

63 Emile Durkheim, *The Division of Labor in Society* Macmillan, 1933

64 David Willetts has recently made the same point in *The Pinch* where he notes that the individualistic Anglo countries tend to be more vital civil societies: 'The vigour of civil society in England and the USA...depends on some very unusual and shared features of our two countries. England and America share a similar civil society because we share the same (rather unusual) family structure' (Atlantic Books, 2010, p.19)

Figure 14 reproduces their finding on *infant mortality*. It reveals a strong and significant association.[65] A boxplot identifies the USA as an outlier, and we can see from the graph that we also have the familiar Scandinavian cluster at the other end of the distribution. But even if the Nordic quartet is omitted, and the USA is excluded as an outlier that may be distorting the finding, the association still stands up among the remaining cases.[66] We also find a significant association between these two variables in the expanded data set across 39 countries, and this still holds even when controlling for GDP.[67] This therefore looks like a solid finding. The more unequal the distribution of income in a country, the higher its infant mortality rate is likely to be (though this does not necessarily mean that the former is causing the latter).

Figure 14: Wilkinson and Pickett's plot of inequality against infant mortality rate



Much less solid, however, is their result on *life expectancy*. Figure 15a reproduces the key graph. It shows a weak association with income inequality – only 15% of the variance is accounted for, and the relationship is only just statistically significant.[68] It is also clear

65 Based on *The Spirit Level*, fig 6.4, and recreated from the international data set downloaded from The Equality Trust web site. Adjusted R^2=0.565, p<0.001.

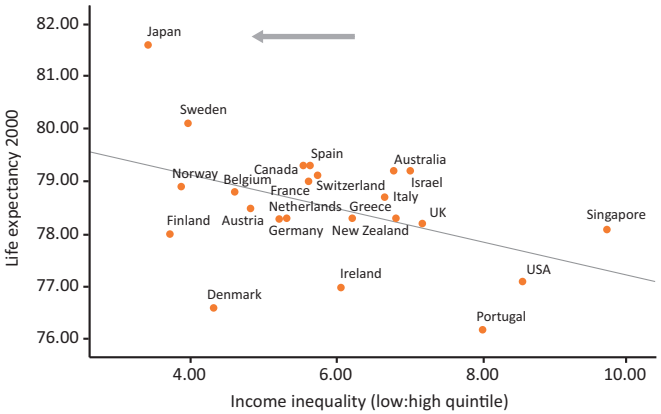
66 Adjusted R^2 =0.237, p=0.027.

67 Botswana, Gabon, Trinidad & Tobago, Mexico, Turkey and Saudi Arabia are all outliers and are excluded from the analysis. For the association between inequality and infant mortality for the remaining countries, the adjusted R^2=0.248, p=0.001. A multiple regression model with GDP per head and income inequality entered simultaneously as independent variables raises the Adjusted R^2 to 0.499 (p<0.001), and both independent variables exert significant effects in this model, although GDP is the stronger of the two: Beta (GDP) = -0.536, p<0.001; Beta (inequality) = 0.344, p=0.008.

68 Based on *The Spirit Level,* fig 6.3, and recreated from the international data set downloaded from The Equality Trust web site. Adjusted R^2 = 0.154, p= 0.036.

from the most casual inspection of this graph that Japan (marked by the arrow) is an outlier which is almost certainly distorting the trend line (a boxplot actually confirms that Denmark, Portugal and Japan are all outliers). This is pertinent, for it has long been recognised that Japan enjoys an unusually high average life expectancy, which has commonly been explained with reference to factors like diet or genes.

Figure 15a: Wilkinson and Pickett's plot of inequality against life expectancy



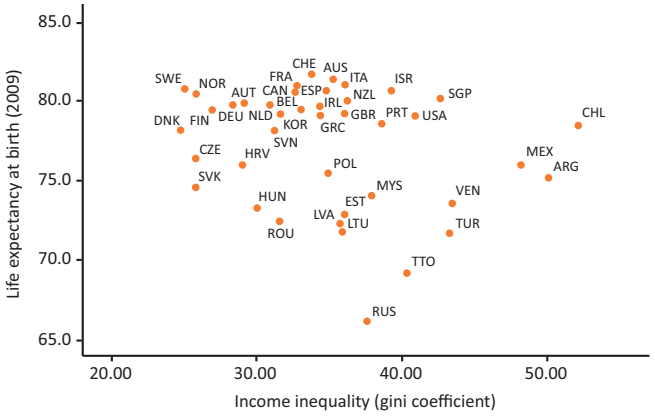
Whatever the explanation for Japanese longevity, it is clearly not the country's compressed income distribution, for there is no association between inequality and life expectancy among any of the other countries in Wilkinson and Pickett's sample. If Japan is removed from their analysis, the apparent association between income inequality and life expectancy collapses and no trend line can be fitted.[69] The highly unequal Singaporeans live longer on average than the highly egalitarian Finns, and the unequal Americans outlive the much more equal Danes. Quite simply, Wilkinson and Pickett have no evidence

69 Adjusted $R^2 = 0.057$, $p = 0.148$.

to link life expectancy to income inequality in these 22 countries, despite their claims to the contrary.

We get the same result on the expanded data set (Figure 15b). Excluding Japan and the two sub-Saharan African countries (whose very low life expectancy makes them extreme outliers at the other end of the distribution) leaves us with 41 countries, but with 97% of the variance left unexplained, there is no association among them between life expectancy and inequality.[70]

Figure 15b: Life expectancy and income inequality in the expanded sample (41 countries)



Conclusion: There does appear to be an association between income inequality and infant mortality, but there is none between income inequality and life expectancy.

2.7 Obesity among adults and children

According to *The Spirit Level*, 'Levels of obesity tend to be lower in countries where income differences are smaller.'[71] But it's not true. What is true is that Americans (and Greeks) tend to be much fatter than people in other countries while the Japanese tend to be much slimmer

70 Life expectancy data are taken from UN *Human Development Report*, 2009; inequality is measured by gini coefficients from the same source. Adjusted R^2=0.031, p= 0.139. If life expectancy is regressed on GDP per head in this expanded sample, we do get a significant finding (adjusted R^2 = 0.511, p< 0.001), but this is only because GDP influences life expectancy among the less prosperous nations (i.e. the relationship is not linear). At per capita GDP above $25,000, there is no association between life expectancy and GDP (Adj R^2 = -0.046, p=0.876

71 *The Spirit Level*, p.91

(a boxplot identifies the USA as an outlier). Obesity rates in other countries fall between these two extremes and exhibit no link to income inequalities. Yet again, therefore, Wilkinson and Pickett's 'finding' has been generated by extreme cases when there is no discernible pattern among all the other countries in between.

Figure 16a reproduces the graph from their book, and Figure 16b shows the same plot when the outlier USA has been excluded.[72] In the first graph, we seem to get a significant association (p=0.007) with a modest strength of association between the variables (Adjusted R^2 = 0.288). But turn to the second graph, and the Adjusted R^2 is halved (to 0.134) while statistical significance disappears (p=0.063). Danes are just as likely to be fat as Kiwis, and Finns are fatter on average than Portugese.

Figure 16a: Wilkinson and Pickett's plot of inequality against obesity

It is much the same story with childhood obesity (this time, there are no Japanese data, so Wilkinson and Pickett rely solely on the heavyweight children of the USA to produce the 'finding' they are after: Figure 17).

72 Based on *The Spirit Level*, fig 7.1, and recreated from the international data set downloaded from The Equality Trust web site.

Figure 16b: Wilkinson and Pickett's plot of inequality against obesity, excluding the USA

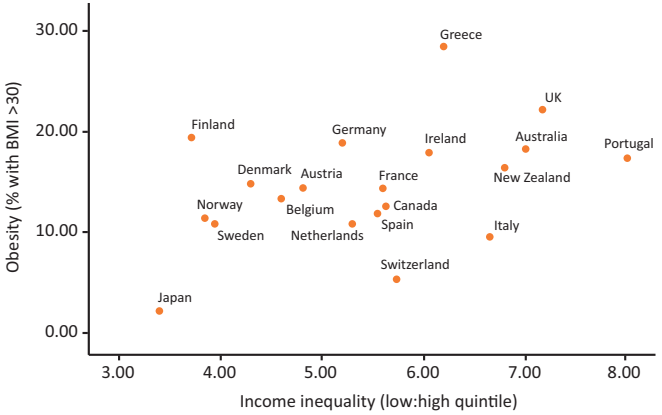
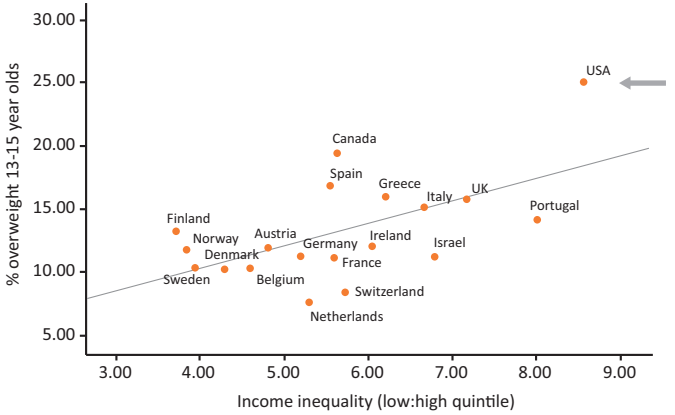


Figure 17: Wilkinson and Pickett's plot of inequality against childhood obesity



With the USA included, they came up with an apparently significant association ($p = 0.008$, Adjusted $R^2 = 0.306$). But with the USA excluded, there is nothing here to report ($p = 0.129$, Adjusted $R^2 = 0.084$).

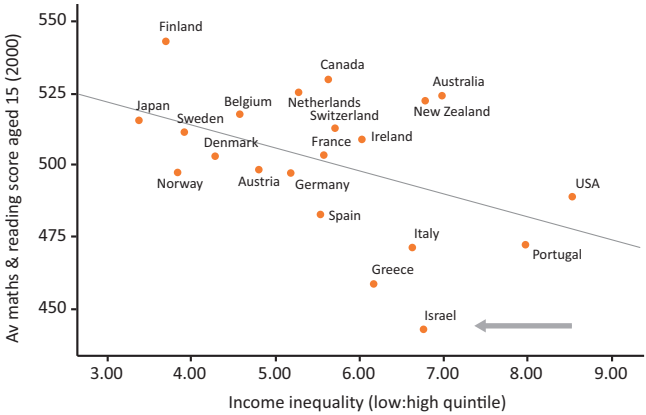
Conclusion: There is no association between obesity and inequality.

2.8 Literacy and numeracy

Combining average maths and literacy scores for 15 year olds, Wilkinson and Pickett claim that education levels are lower in more unequal countries. It is, however, another dubious finding which is reliant on the distorting effect of a single outlying case.

This time, as can be seen in Figure 18a, the outlier is Israel. A boxplot confirms that Israel's poor score makes it a statistical outlier. We might speculate that this has something to do with the quality of education (or the badly disrupted lives) of Palestinian children in that country, but whatever the reason, we should clearly be cautious about accepting a finding that depends heavily on Israel's influence on the trend line.

Figure 18a: Wilkinson and Pickett's plot of inequality against combined maths and literacy scores



This is, however, exactly what Wilkinson and Pickett do. Even with Israel included, the association they report looks flimsy. Inequality explains, at best, only 16% of the variance in combined numeracy and literacy scores, and the relationship between inequality and education scores is teetering on the edge of statistical significance.[73] As soon as we take Israel out, the relationship topples over.[74]

The feeling that the authors should have hedged their 'finding' with some qualifications and caveats is reinforced when we look at the distribution of the cases in Figure 18a. It looks lop-sided. It was noted in Chapter I that regression techniques are quite demanding. They not only require that the slope of the trend line should not be distorted by a few extreme cases, but also that the association between variables be linear (i.e. as the value of x increases, so the value of y should increase or decrease at a fairly steady rate across the whole distribution), and that the variance in the values of the dependent variable, y, form an approximately normal distribution at each value of x. Social statistics never measure up exactly to these pure conditions, but that does not mean we can ignore them, for if they are blatantly breached, we can end up with results which are quite false or misleading.

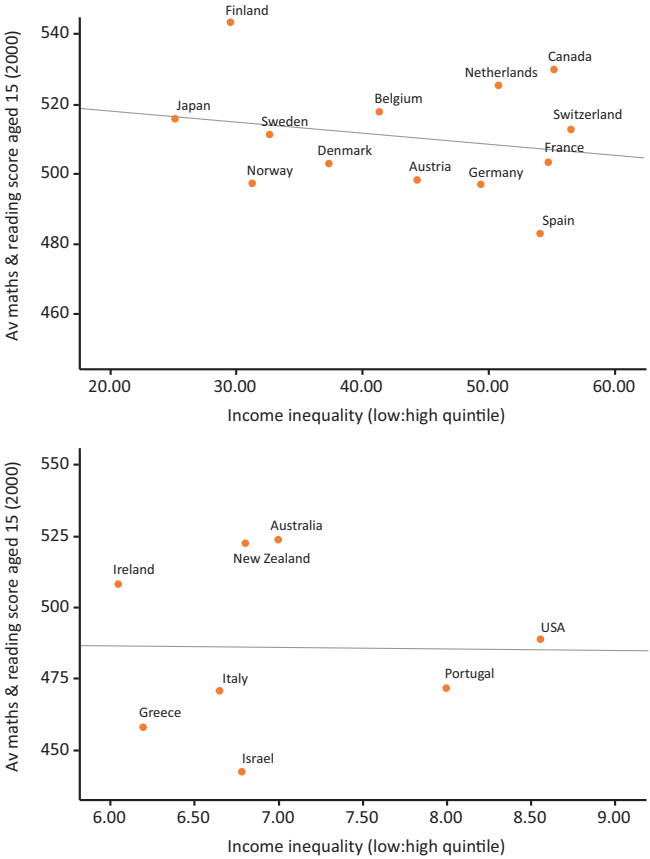
Look again at Figure 18a. At the left-hand segment of the trend line there is a knot of countries below the line which are all very close to it (Japan, Sweden, Denmark, Norway, Austria), but above the line, many of the cases are a long way off it (Finland, Canada, Australia, New Zealand). The line in this section of the graph does not seem to have achieved the 'best fit' which regression analysis seeks – a higher and flatter line would fit these cases far better.

Now look further along the line at the right-hand segment. Here there is only one country (the USA) above the line, but there is a long arc of cases (Germany, Spain, Italy, Greece, Israel) falling away below it. Again, the trend line does not seem to be in the right place.

73 Adjusted R^2=0.159, p = 0.41. The plot is based on *The Spirit Level*, fig 9.2, and has been recreated from the international data set downloaded from The Equality Trust web site.

74 R^2 = 0.128, p = 0.068.

Figure 18b: Wilkinson and Pickett's plot of inequality against combined maths and literacy scores, dividing the sample into more equal and less equal countries

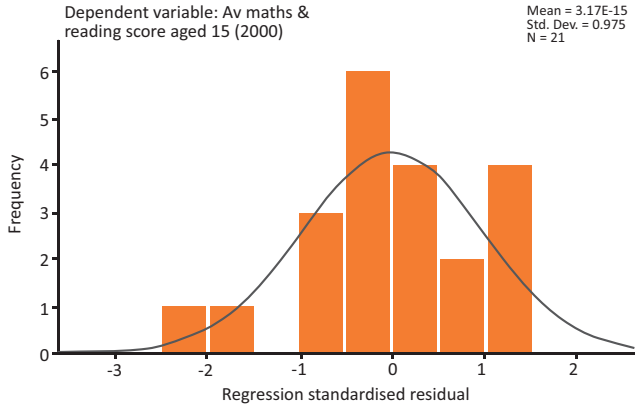


The problem with this graph can be illustrated more clearly if we split it into two. Figure 18b separates the countries with lower inequality (an income ratio below 6, in the top graph) and those with higher inequality (income ratio above 6, in the bottom graph). In both cases, the trend lines are flat. There is no association between these variables in either

part of the graph.[75] So while average scores may be lower in less equal countries,[76] we cannot draw a consistent regression line through them, and it would be misleading to use a trend line to predict a country's education level from its income inequality data.

A histogram of standardised residuals (Figure 18c) confirms that the distribution of scores at different values of the independent variable in Figure 18a is a long way from 'normal.' A key requirement of regression analysis has therefore been violated.

Figure 18c: A histogram of standardised residuals of combined maths and literacy scores, testing for normality at different levels of income inequality



Taking this problem together with the reliance on a single outlier to achieve a statistically significant result, it would have been prudent had the authors concluded that they had no strong evidence in support of their education hypothesis. Instead, they state unequivocally: 'More unequal countries..have worse educational attainment – and these relationships are strong enough for us to be sure that they are not due to chance.'[77]

Conclusion: The reported association between maths and literacy scores and income inequality looks weak and is probably not reliable.

75 For the high equality countries, R^2=-0.167, p=0.983; for the low equality countries, R^2 = -0.038, p=0.470.

76 A Student's t test finds the difference in average scores is on the margins of statistical significance. Mean score for more equal countries (income ratio <0.6) = 511, standard deviation = 16. Mean score for less equal countries (income ratio >0.6) = 486, standard deviation = 30. With unequal variances demonstrated by Levine's test, t = -2.1, which with 9.5 degrees of freedom falls just short of significance (p=0.061).

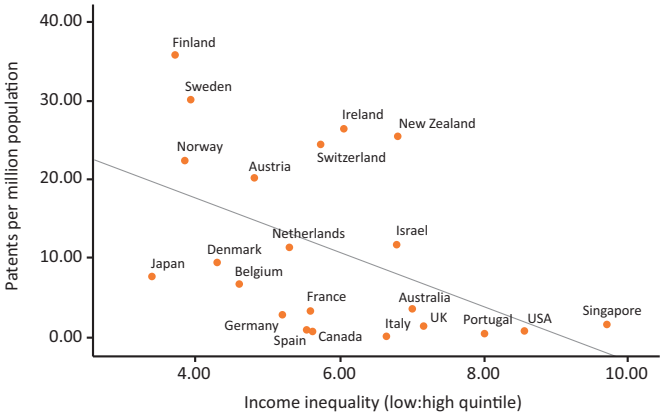
77 *The Spirit Level,* p.105

2.9 Innovation and creativity

We encounter this same problem again later in *The Spirit Level* when the authors look at the distribution of patents across countries.

They want to show, contrary to what is commonly believed, that equality does not deter enterprise and innovation. To do this, they count how many patents are registered each year in the different countries in their sample, and express their results as a proportion of each country's population size. Mapping patents per million population against income distribution, they find that, far from dampening innovation, egalitarianism seems to encourage it (Figure 19a).[78] They conclude that, 'More equal societies tend to be more creative'.[79]

Figure 19a: Wilkinson and Pickett's plot of inequality against patents per million population



But there are three key points to note about this 'finding'.

The first is that *all* the countries that score well on this measure are small: Finland, Sweden, Norway, Austria, Switzerland, Ireland and New Zealand. Perhaps there is something about living in a small

78 Based on *The Spirit Level*, fig 15.3, and recreated from the international data set downloaded from The Equality Trust web site. Adjusted R^2 = 0.203, p = 0.020.

79 *The Spirit Level*, p.225

country that encourages inventiveness. Or perhaps inventors in large countries tend to work collaboratively in companies and universities on a relatively small number of larger projects which result, proportionately, in fewer patents per head of population. A simple patents count may be a flawed measure of a country's inventiveness.

The second point to note about Figure 19a is that the distribution of countries along the trend line looks very lumpy. Most countries, whether equal or unequal, bump along the bottom of the graph at around 10 patents per million or less. A few – the high-scoring, small nations – float well above them in the 20 to 40 per million range, but they are bunched towards the left of the graph. Just looking at this scatter, it seems likely that Wilkinson and Pickett have again violated the regression requirement that variance on the dependent variable (in this case, patents) be roughly normal at different values of the independent variable (income distribution), for there are almost no cases falling above the trend line once the income ratio gets close to 7.

The suspicion that something is wrong can again be checked by inspecting a histogram of standardised residuals, which turns out to be a long way from a normal distribution (Figure 20). But it is also obvious just by looking at Figure 19a that the 'line of best fit' does not 'fit' the more unequal countries, nearly all of which fall on just one side of it. This makes it very unwise to draw a trend line, or to make predictions from it.

The third point to notice about Figure 19a is that the Scandinavians (except Denmark) are once again clustering together at one end of the graph. This again raises the question of whether Wilkinson and Pickett's 'finding' reflects the influence of Nordic cultural distinctiveness rather than income equality. To find out, we can leave the four Nordic countries to one side for a moment and see what happens to the trend line. The answer, predictably, is that it flattens out (inequality explains just 1% of the variance in patents), and the significance of the association collapses (Figure 19b).[80]

80 Adjusted R^2 is 0.013, $p = 0.287$.

Figure 20: A histogram of standardised residuals of patents per million population, testing for normality at different levels of inequality

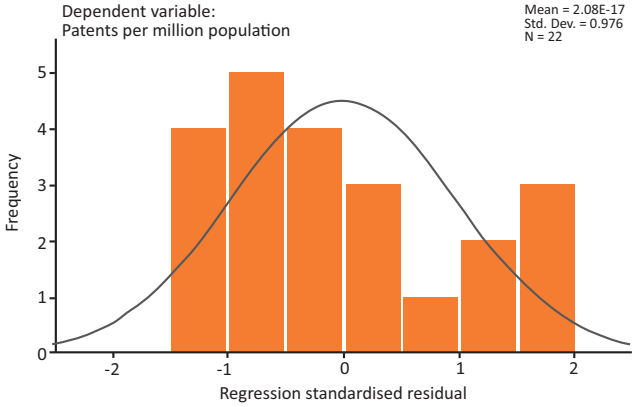
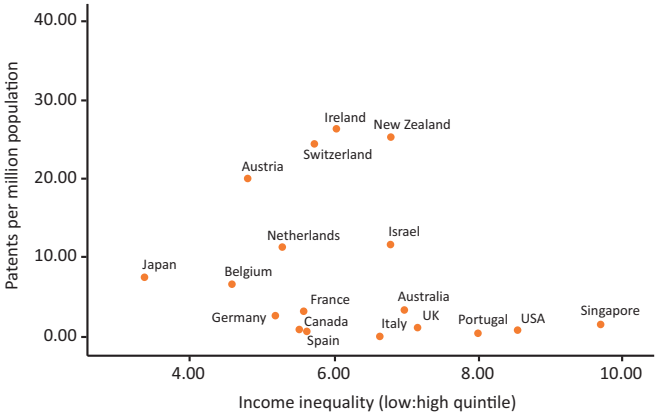


Figure 19b: Wilkinson and Pickett's plot of inequality against patents per million population, excluding the four Scandinavian countries



Basically, therefore, what Wilkinson and Pickett have shown is not (as they claim) that equal countries generate more patents, but that Scandinavia does. Interestingly, using my expanded data set (where I have relevant information on 30 countries), it can also be shown that Scandinavia has won more than its share of Nobel prizes over the last hundred years. Again, this has nothing to do with income distribution (for there is no association between income distribution and Nobel prize success among the other countries included in the analysis). It is simply that Scandinavia punches above its weight on things like patents and Nobel prizes for reasons presumably to do with its historical and cultural uniqueness.[81]

While Wilkinson and Pickett fail to show that equal countries are *more* innovative, it is nevertheless true that egalitarianism does not appear to suppress innovation, for there is no association running in the opposite direction either. Using the expanded data set, I have explored various other indicators of innovation and entrepreneurship including new company start-ups, take-up of e-commerce, investment in R&D and the annual rate of GDP growth, and (allowing for variations in per capita GDP) no association between any of these indicators of economic vitality and the degree of equality or inequality of incomes in a country can be identified.[82]

Although Wilkinson and Pickett are wrong to suggest that equal countries do better on measures like these, it therefore seems that a lot of defenders of free market economics might also be wrong in arguing that radical income redistribution will necessarily choke off the spirit of enterprise and innovation in a country.

Conclusion: Wilkinson and Pickett are wrong to suggest that more equal countries are more innovative. Inequality appears to be neither good nor bad for innovation and creativity.

2.10 Teenage births

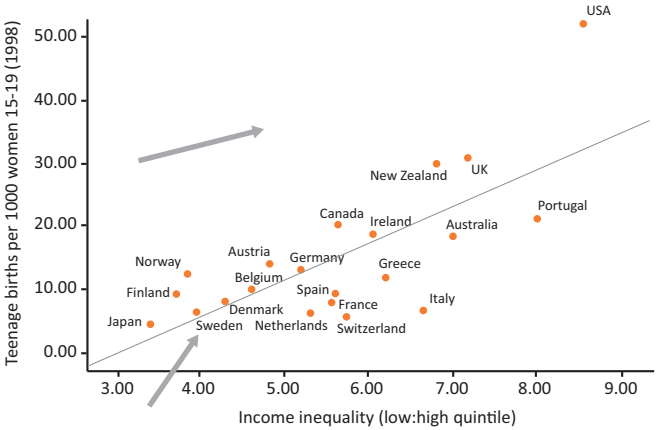
One of the strongest associations reported by Wilkinson and Pickett is that between income inequality and teenage births. 'The teenage

81 The 30 countries on which data are available have won an average of just 0.725 Nobel prizes per million population in the period 1901 to 2002, with a standard deviation of 0.89 (source: http://www.nationmaster.com/index.php). The 4 Scandinavian countries, by contrast, have won an average of 1.95 prizes per million population (and this is dragged down by Finland, with just 0.4). A regression of Nobel prize wins on income inequality, based on a sample of 24 countries but excluding the 4 Scandinavian nations (as well as Switzerland, which is an outlier) produces a non-significant outcome (Adjusted $R^2 = 0.053$, $p = 0.140$).

82 Data in all cases taken from OECD, *Measuring Entrepreneurship* 2009 (http://www.oecd.org/document/31/0,3343,en_2649_33715_41663647_1_1_1_1,00.html.). Data on new company start-ups (companies employing at least 1 other person) are based on an average of manufacturing and services start-ups. Relationship with gini coefficient: Adjusted R^2=0.067, p = 0.155. Use of e-commerce is measured by turnover as % of all commerce in 2008. There appears to be a marginally significant relationship with gini coefficient (Adjusted R^2=0.171, p = 0.050), but controlling for GDP, this disappears (p [GDP]=0.069, p [inequality] =0.165; model fit: Adjusted R^2=0.296, p =0.028). R&D spending shows no association with inequality (Adjusted R^2=0.000, p =0.323). Nor does the annual average rate of GDP growth since 1990 (Adjusted R^2=-0.024, p =0.933).

birth rate,' they say, 'is strongly related to relative deprivation and to inequality,' and they speculate that large numbers of young girls get pregnant in societies where relationships are experienced as fleeting, untrustworthy and unpredictable.[83] Their key graph is recreated in Figure 21a.

Figure 21a: Wilkinson and Pickett's plot of teenage birth rates and income inequality[84]



This is a strong relationship, and unlike the graphs on education and patents, the cases are fairly evenly scattered along the regression line.[85] Having said that, there is a glaringly obvious 'extreme outlier' which shows up clearly on a boxplot. The average teenage birth rate across all these countries is 15 births per thousand women aged 15-19. The average teenage birth rate in the USA, by contrast, is 52 per thousand – more than 3 standard deviations above the sample mean. Clearly, the USA should have been removed from the sample before the regression was computed, for there are unique and special factors contributing to its profile.

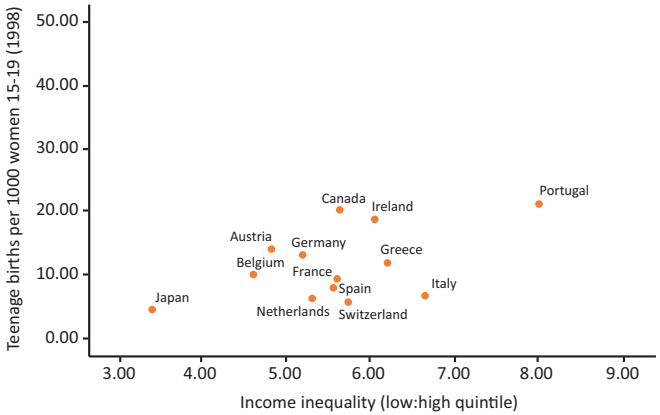
83 *The Spirit Level*, p.121 and following pages.

84 Based on *The Spirit Level*, fig 9.2, and recreated from the international data set downloaded from The Equality Trust web site.

85 Adjusted $R^2 = 0.505$, $p<0.001$.

Even without the USA, however, the association between inequality and teenage births is still significant (although quite a lot weaker).[86] What we really need to know, though, is whether this association holds across all countries, for the familiar cultural clusters can clearly be detected in Figure 21a (indicated by the arrows). At one end of the trend line we find the Scandinavians with low teen births; at the other end, we find the Anglo countries (Australia, New Zealand, the UK, as well as the USA) with high teen births. The key question is, does the association with inequality hold in the eleven other countries?

Figure 21b: Wilkinson and Pickett's plot of teenage birth rates and income inequality with Nordic and Anglo nations omitted



The answer, as we see from Figure 21b, is that it does not. The association is no longer significant.[87] It seems that teenage births are high in the Anglophone countries and low in Scandinavia, but this has little or nothing to do with their levels of income inequality.

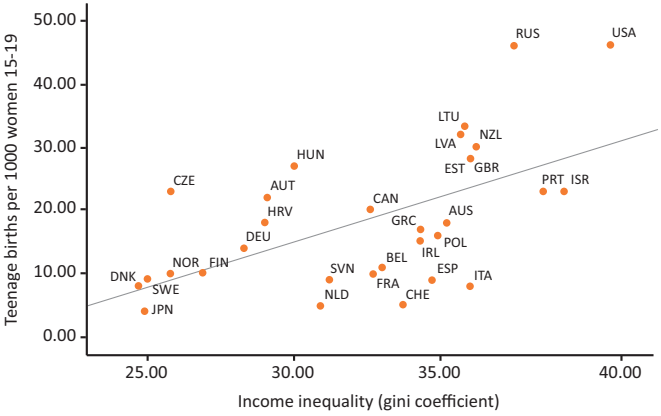
But perhaps we are being unfair on Wilkinson and Pickett, for taking out the Nordic and Anglo nations leaves their analysis with

86 Adjusted R^2= 0.374, p= 0.002.

87 Adjusted R^2 = 0.119, p = 0.160.

only 11 cases. We can rectify this by turning to my expanded sample, where I have 31 countries with data on teen births. Because we are now including countries like Russia, which have teenage birth rates to rival those in the USA, no country appears on a boxplot as an outlier, so we can include them all in our regression model. The result is given in Figure 21c.

Figure 21c: Teenage birth rates and income inequality in the expanded sample (31 countries)

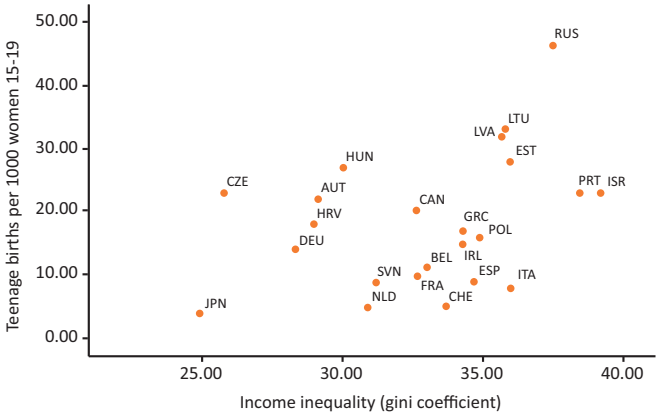


The result is a significant, if modest, association in the direction predicted by Wilkinson and Pickett. As inequality increases, teenage births go up, and income distribution explains almost one-third of the variance.[88] But what happens if we exclude the Scandinavian and Anglo (UK, USA, Australia and New Zealand) countries from this expanded sample (Figure 21d)? We still have 23 countries remaining – more than enough to test whether there is any association. But as with their own, more restricted, sample (Figure 21b), once the Nordic and Anglo blocs are removed, the association between teenage births and inequality ceases to achieve statistical significance.[89]

88 Adjusted $R^2 = 0.315$, $p = 0.001$

89 Adjusted $R^2 = 0.100$, $p = 0.078$

Figure 21d: Teenage birth rates and income inequality in the expanded sample, excluding the Nordic and Anglo countries



Clearly, the English-speaking countries do have a problem with their high teenage birth rates, and the Scandinavian countries seem to perform much better, but these differences probably should not be attributed to differences in their levels of income inequality. They are more likely to reflect specific cultural characteristics.

Conclusion: There is no significant association between inequality and teenage births. The apparent association reported in The Spirit Level is due mainly to the distinctiveness of the Anglo and Nordic countries.

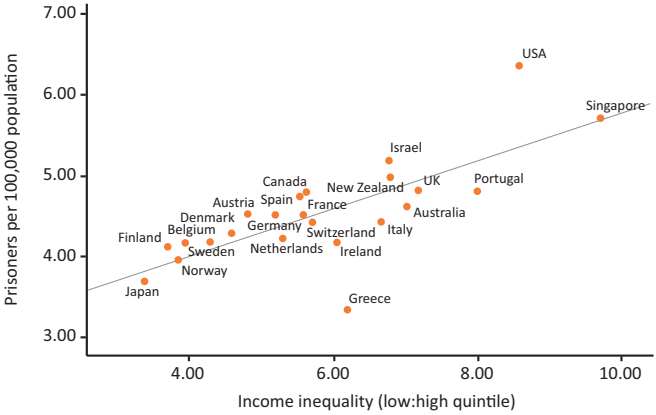
2.11 Imprisonment rates

The association between imprisonment rates and income inequality is one of the strongest analysed in *The Spirit Level* (see Figure 22a). Income distribution accounts for more than half the variance in imprisonment rates.[90] The more unequal a country, the bigger the proportion of its population which is likely to be in prison. Wilkinson and Pickett think this is because egalitarian countries are gentle

90 Adjusted $R^2 = 0.541$, $p < 0.001$. The graph is based on *The Spirit Level*, fig 11.1, and is recreated from the international data set downloaded from The Equality Trust web site.

places, where the emphasis is on 'treating' or 'morally re-educating' offenders, whereas more unequal countries are 'harsher, tougher places' where people feel vindictive towards wrongdoers.[91]

Figure 22a: Wilkinson and Pickett's plot of imprisonment rates and income inequality



So strong is the relationship between inequality and imprisonment rates that for the first and only time in the book, the authors are happy to admit to the existence of outliers when they are discussing this scatterplot. They note: 'Even if the USA and Singapore are excluded as outliers, the relationship is robust among the remaining countries.'[92] Sure enough, a boxplot confirms that these two are indeed outliers, which means they probably should be excluded, but this does not undermine their result. Even with ultra-punitive Singapore and USA excluded, income inequality is still accounting for almost one-third of the variance in imprisonment rates, and the relationship is still clearly significant.[93]

It is notable that we again find the Scandinavian cluster at the bottom end of the distribution, but this finding still stands up even

91 *The Spirit Level*, p.155

92 *The Spirit Level*, p.148-9

93 Adjusted R^2= 0.304, p = 0.006.

if the four Nordic countries are taken out of the analysis as well.[94] Without them, income inequality still explains 20% of the variance in imprisonment rates, so the pattern in Figure 22a cannot be explained away as the result of some peculiar aversion to locking up criminals in Scandinavia. There really does seem to be a link to income inequality.

Two notes of caution should, however, be sounded.

First, this association is not found in my expanded sample of countries. Excluding two 'extreme outliers' (USA and Russia) and two 'outliers' (Latvia and Estonia), we are left with 28 countries. As Figure 22b, reveals, even with the Scandinavians included in this sample, there is no association between income inequality and imprisonment rates among these countries.[95] Looking at this graph, however, it may be that the association is being weakened by some high-imprisonment, former Soviet bloc countries (Lithuania, Romania, Poland, Hungary and Czech Republic). Sure enough, omitting them from the analysis restores the association.[96] The implication of this is that, while inequality may be linked to imprisonment rates in western Europe, in Eastern Europe incarceration rates are very high regardless of income distribution.

The second note of caution is that it does not make a lot of sense analysing imprisonment rates without also looking at crime rates. If they wanted to gauge the 'harshness' of different countries' penal policies, the key question Wilkinson and Pickett should have addressed was not whether more unequal countries lock up a bigger proportion of their *populations*, but whether they lock up a bigger proportion of their *criminals*. When we gather the statistics on this question, we get a very surprising answer.

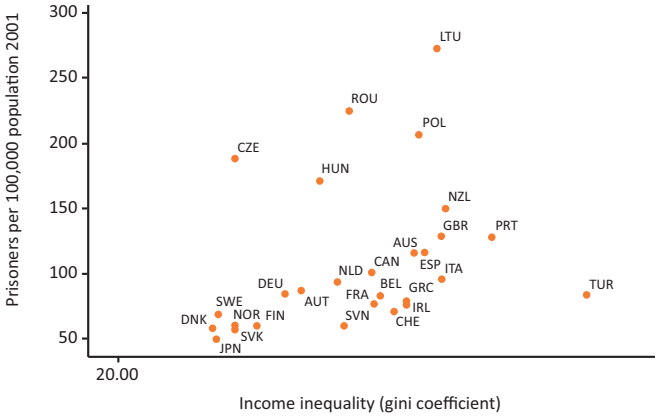
It is notoriously difficult to get reliable crime statistics which can be directly compared across countries, but the International Crime Victim Survey asks people if they have been the victims of various kinds of crimes, and these answers can provide a reasonable guide to people's experiences in different countries.

94 Adjusted R^2= 0.201, p = 0.040.

95 Adjusted R^2= 0.059, p = 0.113.

96 Adjusted R^2= 0.414, p = 0.001

Figure 22b: Imprisonment rates and income inequality in the expanded sample (28 countries)[97]



Figures 23a and 23b plot national imprisonment rates against the number of people in 24 of the countries in my international data set who say they have been victims of burglary and car theft. Ignoring the outliers in both graphs (the USA and Estonia, which both have very high incarceration rates), what is striking about these graphs is that imprisonment rates seem to have no relationship with crime rates. The UK, for example, has the highest rate of reported burglaries of any of these countries, yet its imprisonment rate is not out of line with those of Germany, Spain and Sweden, which have some of the lowest burglary rates. It's much the same pattern with car thefts – Britain and New Zealand again head the league table, yet their imprisonment rates do not reflect this. Regression models confirm that there is no statistically significant association between the number of burglaries or thefts in a country and the size of its prison population.[98]

97 Imprisonment data from Gordon Barclay & Cynthia Tavares, *International comparisons of criminal justice statistics 2001* Home Office 2003 Table B

98 Models exclude USA as extreme outlier. For burglaries, Adjusted R^2=0.069, p=0.120. For car thefts, Adjusted R^2=0.036, p=0.640.

Figure 23: Reported (a) burglary rates and (b) car theft rates compared with imprisonment rates in selected countries[99]

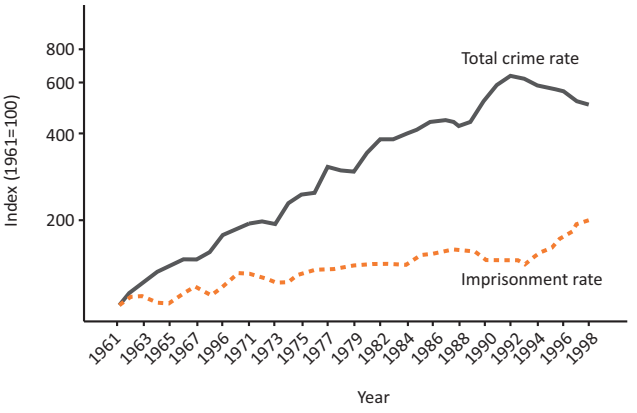
This result runs directly contrary to Wilkinson and Pickett's thesis. They say prison is ineffective in reducing crime ('The consensus among experts worldwide seems to be that it doesn't

99 Burglary and car theft date from: Jan van Dijk, John Van Kesteren, Paul Smit, *Criminal Victimisation in International Perspective* The Hague, Ministry of Justice, 2007

work very well'),[100] and they conclude from this that societies with higher prison populations must be motivated by spite: 'More unequal societies are harsher, tougher places. And as prison is not particularly effective for either deterrence or rehabilitation, then a society must only be willing to maintain a high rate (and high cost) of imprisonment for reasons unrelated to effectiveness.'[101]

But we now see that, far from being 'harsher' and 'tougher' on criminals, as they suggest, the more unequal countries like Britain, New Zealand, Australia and Portugal *are actually locking up a smaller proportion of their criminals than the more egalitarian countries are*. It seems that far from being too hard on crime, the more unequal countries have been too 'soft' on it (which may be one reason why their crime rates are so high).[102]

Figure 24a: UK crime and imprisonment rates, 1961-1999 (1961=100, logarithmic scale)[103]



Take Britain as an example. For many years, the probability that criminals would go to jail declined dramatically in the UK. As Figure 24a shows, the total number of crimes committed in Britain (relative to the size of the population) rose more than six-fold between 1961

100 *The Spirit Level* p.153.

101 *The Spirit Level* p.155.

102 Charles Murray is one exception to the 'expert consensus' to which Wilkinson and Pickett appeal. He argues that, while prison may not be very effective as a deterrent, and it probably does not achieve much by way of reforming criminals, it is effective in incapacitating wrongdoers. Because most serious crime is committed by a relatively small proportion of the population, locking up serial offenders is a very good way of cutting the overall crime rate, for criminals cannot offend while they are in prison. See Charles Murray, *Does Prison Work?* Civitas, 2000. See also Peter Saunders and Nicole Billante, 'Does prison work?' *Policy* vol.18, no.4, 2002-03, pp.3-8

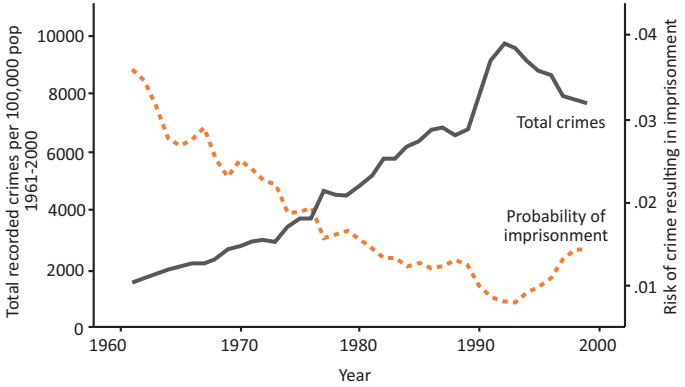
103 Home Office *Recorded Crime Statistics 1898-2001/2* and *Prison Population England & Wales 1999*, Table 1.2a. In 2001-02, the basis for recording crime statistics changed so the series is not comparable after that date. Nevertheless, crime rate has continued to fall since the turn of the century as the prison population has continued to rise. I shall be analysing this in greater detail in a future publication for Civitas.

and 1992, but the prison population (as a percentage of the whole population) did not even double during this period. Only since the mid-1990s has penal policy in the UK been tightened up.

As penal policy got tougher from the 1990s onwards, crime rates at last began to fall. But Wilkinson and Pickett get this causation back-to-front. They accept that in recent years, 'Crime rates in the UK were falling as inexorably as imprisonment rates were rising.'[104] But instead of recognising that it was the increased use of imprisonment that led to this falling rate of crime, they think we have wilfully been locking people up in greater numbers *despite* crime falling (they complain about 'the increased use of custodial sentences for offences that a few years ago would have been punished with a fine or community sentence').[105]

Figure 24b demonstrates that, as the risk of imprisonment fell, the crime rate rose, and it was only when the risk of imprisonment began to rise that the crime rate began to fall back. Wilkinson and Pickett's claim that mean-spiritedness pushed up imprisonment rates at a time when crime was falling confuses outcomes with causes.

Figure 24b: The association between the number of crimes being committed in the UK and the probability of a criminal going to prison, 1961-2000[106]



104 *The Spirit Level* p.147

105 *The Spirit Level* p.147

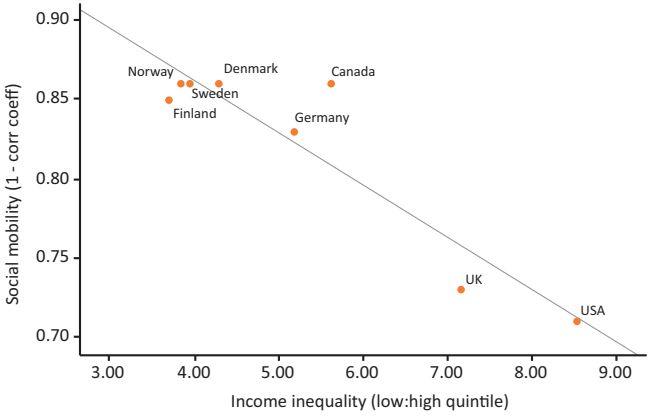
106 *Recorded Crime Statistics 1898-2001/2* and *Prison Population England & Wales 1999*, Table 1.2a.

Conclusion: There does appear to be an association between inequality and imprisonment rates among western countries, but this does not mean the more unequal countries are 'harsher' in their treatment of criminals, which is what Wilkinson & Pickett suggest. On the contrary, criminals in countries like Britain have for many years been much less likely to receive custodial sentences than criminals in more egalitarian countries.

2.12 Social mobility

We have saved the worst until last, for the discussion of social mobility in *The Spirit Level* combines almost every error found in other parts of the book. It includes the distortion of results by outliers, the illegitimate fitting of a trend line to data that won't support it, and the reliance on unique cases to create the illusion of a general pattern where none exists. It then compounds these errors by uncritically accepting data on international mobility rates which should never have been accepted as reliable in the first place.

Figure 25: Wilkinson and Pickett's plot of social mobility rates (measured by movement between income quartiles) and income inequality[107]



107 Based on *The Spirit Level*, fig 12.1, and recreated from the international data set downloaded from The Equality Trust web site.

Wilkinson and Pickett want to claim that social mobility (the movement of people between different social classes or income bands) is greater in more equal countries, and that opportunities in less equal countries are more restricted. Their evidence for this proposition is summarised in Figure 25.

The correlation coefficient they report for this graph is so high (0.93) that this alone might have set off alarm bells in a more cautious and sceptical research team. It translates into an Adjusted R^2 of 0.828 (p=0.001), which suggests that the income distribution of a country determines 83% of the variance in the opportunities available to people to move up or down the income ladder. The truth is very different.[108]

There is a lot of confusion in the academic literature on international social mobility rates, and the OECD warns that such comparisons should be treated with 'a great deal of caution.'[109] Most sociologists think mobility rates between *social classes* are similar across most western countries, and a 2001 review by the UK Government's Performance and Innovation Unit concluded that any differences that do exist are 'modest.'[110] Some economists, however, have recently suggested that mobility between *income quartiles* in Britain and the USA is lower than in Canada and a small number of European countries. They base this conclusion on evidence of a stronger association between parents' and children's earnings here and in the US than in the other nations.[111] It is this research that Wilkinson and Pickett use. But when we inspect these data more closely, the claim is unconvincing.

The most prominent of these economists, Jo Blanden, admits, 'There is a lot of uncertainty about the UK.' The problem is that parental incomes change over time, so if you are comparing people's incomes with those of their parents, it matters a great deal at what age parental earnings are estimated. To deal with this difficulty, economists have 'adjusted' their data on parents' incomes to reflect their occupations or education, but this produces wildly

108 The discussion that follows is based on Peter Saunders, *Social Mobility Myths*, Civitas, 2010

109 Anna Cristina d'Addio, 'Intergenerational transmission of disadvantage' *Social, Employment and Migration Working Paper* no.52, Paris, OECD, 2007, p.29

110 Stephen Aldridge, *Social Mobility: A Discussion Paper* Cabinet Office Performance & Innovation Unit, April 2001, para. 38

111 Jo Blanden, Paul Gregg, Stephen Machin, *Intergenerational mobility in Europe and North America* Centre for Economic Performance, London School of Economics, April 2005

differing estimates, and it is not clear which we should accept. Blanden, for example, reports one correlation between UK sons' and parents' incomes of 0.44, but notes that this seems 'extremely high' (even though it has itself been adjusted downwards from an initial 0.58) when compared with another of just 0.29. She averages the two to get her own figure, but this is completely arbitrary, and we have no way of knowing whether or not it brings her close to a 'true' estimate.[112]

"When the correlation between parents and children's incomes in the UK is compared with those from other countries, it comes out higher than the other European countries examined, and is only exceeded by the USA. However, the standard errors on these estimates are huge"

When the (adjusted, averaged, estimated) correlation between parents' and children's incomes in the UK is compared with those from other countries, it comes out higher than the other European countries examined, and is only exceeded by the USA. However, the standard errors on these estimates are huge (the standard error gives us the likely range within which the real figure for each country lies). These ranges are so great, they nearly all overlap with each other, which means the differences between most countries' estimates are not statistically significant. For example, the USA ranks very poorly while Sweden ranks quite well, but Blanden admits: 'It is impossible to statistically distinguish the estimates for Sweden and the US.' She goes on: 'The appropriate ranking at the top end is difficult with large standard errors on the Australian, French, British and US estimates *making it unclear how these countries should be ranked*.'[113] Blanden also acknowledges that another study has estimated income mobility for 15 countries (not including Britain) and come up with a very different set of results in which the USA is around the middle of the rankings, beating Australia and coming close to Norway.

Clearly, these international comparative statistics are highly suspect. Perhaps Wilkinson and Pickett can be forgiven for not having realised this. They accepted at face value data which looked

112 Jo Blanden, 'How much can we learn from international comparisons of intergenerational mobility?' London School of Economics Centre for Economics of Education *Departmental Paper* November 2009, p.15

113 'How much can we learn from international comparisons of intergenerational mobility?' p.15, emphasis added

useful for their hypothesis, without delving very much into their reliability. Less forgivable is what they did next.

Look at Figure 25. It contains just eight cases. Six of them are clustered together in the top left quadrant of the graph, with the remaining two in the bottom right. Even if we assumed that the statistics on which this scatterplot is based were valid and reliable, it is obvious that a distribution like this cannot support the sort of analysis they go on to perform on it.

What we have here is a dichotomous distribution – two distinct groups. It is tempting to refer to the USA and UK as 'outliers', but with a sample of this size, we might just as well refer to the bunch of six countries at the other end of the regression line as outliers as well! There is no justification for mapping either of these groups of countries onto a graph and drawing a trend line across the acres of empty space that separates them, because there is no possible trend here to detect. They are simply two different categories of countries, one with (apparently) high mobility, the other low.

And now look at the composition of the groups. At one end, the two major Anglophone nations. At the other, the four Scandinavian countries. And making up the numbers, Canada and Germany. Given the cultural distinctiveness of the Scandinavians on the one hand, and the Anglo bloc[114] on the other, a graph like this cannot possibly tell us anything other than that social mobility rates in Scandinavia are higher (according to these highly unreliable figures) than those in the UK and the USA. Beyond this, there is nothing more to say.

Yet Wilkinson and Pickett have no hesitation in concluding that, 'Income inequality causes lower social mobility.'[115] Their data cannot support such a claim.

Conclusion: The data on social mobility used in The Spirit Level are seriously flawed, and the analysis, based on just eight cases, is inadmissible.

114 As elsewhere in this report, I do not include Canada as an 'Anglo' nation because of its significant French-speaking population.

115 *The Spirit Level,* p.168

3. Income inequality and social pathology in the USA

International comparisons do not support the central claim in *The Spirit Level* that greater income inequality produces worse social problems. It is true that the more equal Scandinavian countries often perform better than the less equal Anglo countries on many of the indicators analysed by Wilkinson and Pickett, but this rarely has anything to do with their income distribution. We know this because there is generally no association between income inequality and the phenomena in question when we look at other countries.

But Wilkinson and Pickett have a second string to their bow. As well as presenting evidence comparing different countries, they also produce evidence on how the 50 US states perform on these same measures. Here too, they claim their basic hypothesis is supported, for the more unequal states tend to exhibit higher rates of social pathology.

This is an impressive finding, for not only does it mean the authors have replicated one set of results across two different data sets, but the use of the US states gives them a bigger sample, with 50 cases, which seems to rule out the possibility that distinct cultural differences might be creating the variations they are finding. The international comparisons may have been compromised by the influence exerted by the Scandinavian bloc, at one end of each distribution, and the Anglo bloc, at the other, but with the US states they are making comparisons across a single country.

The USA is, however, a huge and diverse country, and two aspects of its diversity are potentially crucial in their analysis:

- One is the division between the south and the rest of the country. This is important, not simply in economic terms (the north is historically richer, and more industrial), but in historical-cultural terms too. It is only a century and a half since these two waged bitter civil war over the issue of slavery, and it is less than fifty years since the south was desegregated. In the former Confederate states of the 'Deep South', the legacy of this history for social cohesion and pathology should not be discounted.[116]
- The other division links to this – it is the ethnic division between African-Americans and the white population. On most indicators of advantage and disadvantage, black America still lags behind the whites. Moreover, race and ethnicity are important components of social identity for many Americans, and this is likely to affect measures of fragmentation and cohesion between them.

The authors of *The Spirit Level* recognise that race may confound their analysis of inequality across the fifty states, for the states with the highest black populations tend to have greater social problems *and* more unequal income distributions. As they put it, 'In the USA, state income inequality is closely related to the proportion of African-Americans in the state's population. The states with wider income differences tend to be those with larger African-American populations. The same states also have worse outcomes.'[117]

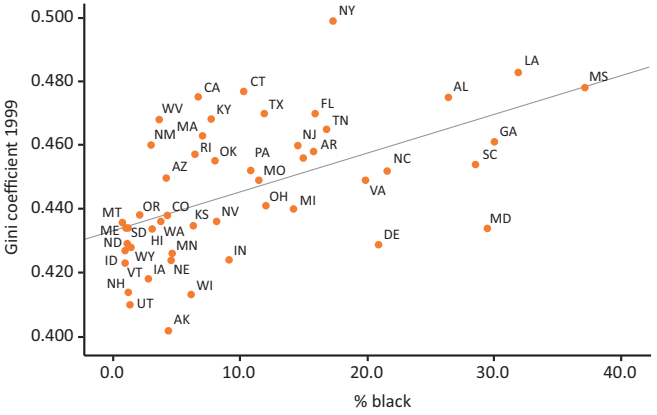
Wilkinson and Pickett do not actually tell us how closely ethnic composition and income inequality co-vary, but on my US states data set, there is a significant and moderately strong correlation ($R=0.54$, $p<0.001$) between the gini coefficient (measuring inequality in each state) and the proportion of African-Americans in the population (see Figure 26).[118] The question we have to ask whenever we encounter states with high levels of social pathology and fragmentation is, therefore, whether it is income inequality or racial composition that is causing the problems.

116 The slave states in 1861 were: Alabama, Arkansas, Delaware, Florida, Georgia, Kentucky, Louisiana, Maryland, Mississippi, Missouri, North Carolina, South Carolina, Tennessee, Texas and Virginia. Delaware, Kentucky, Maryland and Missouri did not join the Confederacy, and West Virginia broke from Virginia to join the Union and abolish slavery.

117 *The Spirit Level* p.185

118 The correlation between the gini coefficient and the proportion of whites is -0.384 ($p=0.006$), and between the gini coefficient and proportion of Hispanics = 0.281 ($p=0.048$).

Figure 26: Income inequality and size of black population in the 50 US states[119]



Wilkinson and Pickett seek to sidestep this question by arguing that racial heterogeneity and income inequality amount to the same thing, so it doesn't matter which is causing the bad outcomes. They suggest that the 'social exclusion' of African-Americans means that 'ethnic divisions may provide almost as good an indicator of the scale of social status differentiation as income inequality.'[120] So throughout their book, they analyse variations between US states, assuming they have been caused by income inequality, even if they have actually been caused by racial differences. In their view, the distinction is unimportant.

This sloppy logic cannot go unchallenged. The hypothesis that ethnic diversity generates bad outcomes is very different in at least two crucial ways from the hypothesis that income inequality generates bad outcomes.

The first is that racial disadvantage cannot simply be reduced to economic disadvantage. Different ethnic groups have different cultures and lifestyles which may generate very different life

119 Gini coefficient is calculated on household income in 1999 and is from US Census Bureau web site (http://www.census.gov/), Table S4: 'Gini ratios by state'. Percent black population in each state from US Census Bureau, *State and county quick facts* 2010

120 *The Spirit Level* p.186

outcomes, irrespective of income differences. For example, higher rates of single parenthood in the African-American community (which have sometimes been linked to the legacy of slavery)[121] can be expected to produce, on average, poorer education outcomes and higher rates of crime in black neighbourhoods, for there is a correlation (in white as well as black communities) between children raised without fathers and poor social outcomes.

Wilkinson and Pickett try to close down lines of investigation like this by suggesting they are 'simply an expression of racial prejudice.'[122] But this is nonsense – it is not racist to recognise that different lifestyles are associated with different ethnic groups. Indeed, Wilkinson and Pickett themselves recognise elsewhere in their book that 'different ethnic groups can have different cultures and values' which influence outcomes such as women's status and marriage rates.[123] Racial and ethnic disadvantage is about more than just incomes.

The second point is that ethnic diversity may itself generate certain social effects. We shall see in Chapter IV how the traditional cultural homogeneity of Scandinavia and Japan might help explain these countries' strong performance on many of the indicators we have analysed. Their strong sense of common identity seems to have its roots in ethnic homogeneity and a history of relative closure to immigration. But what is true of homogeneity in Scandinavia and Japan may be true in reverse of heterogeneity in the more racially-mixed states of the USA – and particularly in the Deep South where old racial divisions run deep. For example, there may be a reluctance to invest in schools and health care in states where the white majority does not want its taxes supporting other ethnic groups from whom they feel socially distant. This will then be reflected in worse education and health outcomes in these states. Although Wilkinson and Pickett are scornful of historical and cultural explanations, it is likely that history and culture have played a crucial role in shaping the outcomes which they analyse.

121 The best-known example of this thesis is the Moynihan Report on the 'Negro Family' in the 1960s. See Lee Rainwater and William Yancey, eds. *The Moynihan Report and the Politics of Controversy*. Cambridge, MIT, 1967

122 *The Spirit Level* p.185

123 *The Spirit Level* p.124

It will be crucial, therefore, whenever we find an apparent association in the states data between income inequality and social outcomes, to test it against the ethnic/racial composition of the states and/or their historical association with slavery and the Confederacy of the Deep South. We can do this by using multiple regression techniques. Rather than simply plotting income inequality against social outcomes and concluding (as Wilkinson and Pickett do) that any association proves causation, we need to compare three hypotheses:

- Income inequality as principal cause (assessed using the gini coefficient);
- Deep South history as principal cause (assessed using a dichotomous 'dummy variable' distinguishing former slave states from former free states);
- Ethnic composition (heterogeneity) as principal cause (assessed by the percentage of the state's population that is African-American).

I shall be using data which I have gathered, because Wilkinson and Pickett's US state data were not available on their Equality Trust web site at the time of writing, and they did not in any case gather data on the ethnic composition of states which we shall need. I shall note any divergence of my findings from theirs, although in most cases my sources are the same as theirs (only a bit more recent). I do not have data on all of their indicators, but we can examine enough of their findings to establish what is going on.[124]

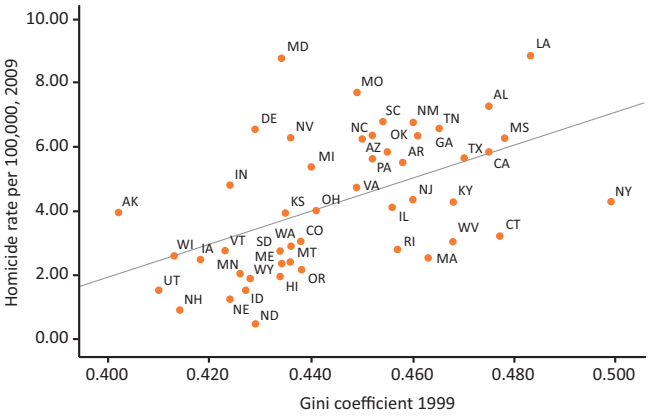
3.1 Homicide

Figure 27a recreates Wilkinson and Pickett's scatterplot of homicide rates against income inequality for the US states. My figures are more up-to-date than theirs, and I actually get a slightly stronger correlation

124 Some of the international indicators analysed in *The Spirit Level* are not used in the analysis of the US states (foreign aid, patents, children's exposure to conflict and social mobility). There are, however, three other indicators (women's status, obesity and trust) which Wilkinson and Pickett do analyse where I do not have appropriate data to replicate their analyses. My inability to analyse trust levels across the US states is unfortunate. Regional data on levels of trust are available on the NORC website, which is where Wilkinson and Pickett say they sourced it, but NORC advise that they will only supply state-level data on payment of US$750, and that this 'can take several months.'

than they do (a Pearson correlation coefficient, r, of 0.52, compared with their result of r=0.42). This translates as an R^2 of 0.27.[125]

Figure 27a: Homicide rates and income inequality (US states



It can be seen from Figure 27a that many of the states with the highest homicide rates are in the south: Louisiana, Maryland, Missouri, Alabama, Tennessee, South Carolina. Closer inspection reveals that the murder rate in the south is almost double that in the rest of the country.[126] However, if we take out the southern states, the association between inequality and homicide still holds among the rest.[127]

The real story on homicide rates is not the distinctiveness of the Deep South, but the importance of race. Compare the tight cluster of cases around the trend line in Figure 27b (a plot of racial composition against homicides) with the much wider scatter in Figure 27a. Clearly the size of the black population is a very strong predictor of a state's homicide rate. Indeed, it explains more than half the variance.

125 R^2=0.269, p<0.001. Florida is missing from my analysis (N=49), while Wyoming is missing from theirs. Their original graph is in *The Spirit Level*, Figure 10.3. My source for homicide rate per 100,000 population: *The Guardian* data blog, 5 October 2009

126 The murder rate in the south is more than 2 standard deviations higher than elsewhere. Mean homicide rate per 100,000 = 6.53 (standard deviation 1.32) in the former slave states and 3.36 (sd 1.66) in the former free states (t=6.38 with 47 degrees of freedom, p<0.001).

127 With the 14 former slave states in the south removed (Florida is already missing), R^2 for the remaining 35=0.217, p=0.005. Among the 14 southern states, however, there is no association: R^2 = 0.004, p=0.835.

Figure 27b: Homicide rates and the size of the black population (US states)

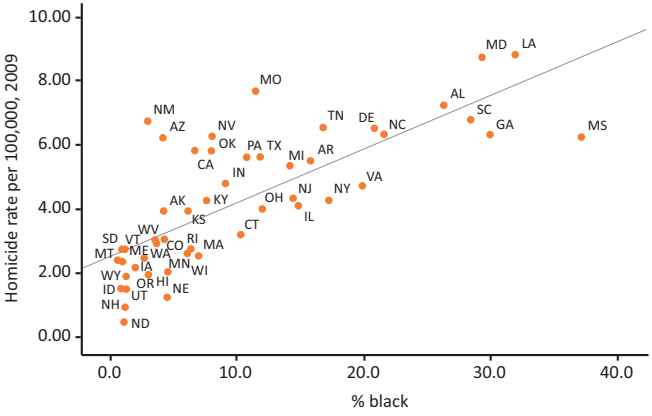
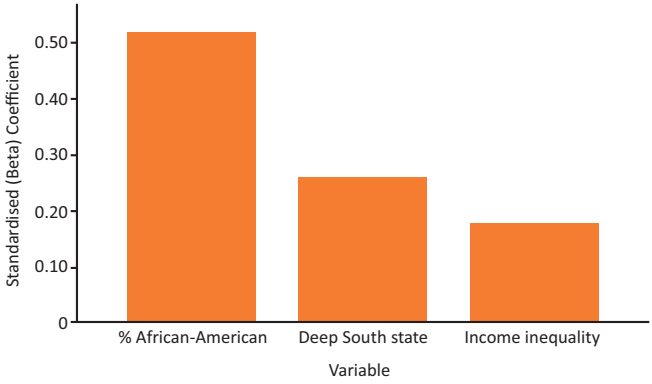


Figure 27c: The relative importance of income inequality, southern location and racial composition in determining a state's homicide rate



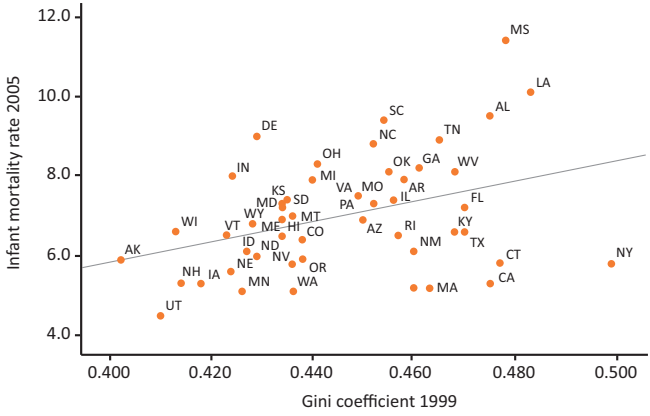
Knowing the racial composition of a state allows us to make a better prediction of its homicide rate than knowing *both* its income distribution and whether or not it is in the Deep South.[128] Indeed, if we construct a model with all three of these measures as independent variables, only racial composition achieves statistical significance.[129] Figure 27c summarises the relative explanatory power of these three variables as demonstrated by this multiple regression model.

Conclusion: Income inequality does not explain a state's homicide rate; the size of its black population is the only predictor we need – and it is a strong one.

3.2 Physical health

As with their international data, Wilkinson and Pickett measure physical health differences across the US states by means of two different indicators: infant mortality and life expectancy. Let us consider each in turn.

Figure 28a: Infant mortality and income inequality (US states)



128 For the association between racial composition and homicide, R^2 =0.572, p<0.001.

129 Model fit R^2 =0.613, p<0.001. Beta (race) =0.479, p=0.005 (p [Deep South] = 0.110; p [inequality] = 0.140).

Figure 28a replicates the plot in *The Spirit Level* charting infant mortality against income inequality. I find a slightly weaker association than the one they report (r=0.38 compared with r=0.43), but the two results are very close and are broadly comparable. The association is statistically significant, but is not strong.[130] A boxplot reveals Mississippi to be an outlier, and if this is removed, the association becomes even weaker and is only marginally significant.[131]

Looking at Figure 28a, it is tempting to suggest that the apparent association between income inequality and infant mortality is actually being driven by the difference between the Deep South and the rest. Even without Mississippi, the worst infant mortality statistics are mainly in the south (Louisiana, Alabama, the Carolinas and Tennessee). The average infant mortality rate in 2005 in the former slave states was 8.2 per thousand births, compared with 6.4 per thousand in the other states of the union, and this is a statistically significant difference.[132]

Tellingly, when we split the sample between the 15 former slave states and the 35 former free states (Figure 28b), we no longer get any association between inequality and infant mortality in either sub-sample.[133] This casts doubt on Wilkinson and Pickett's claim that infant mortality in the USA is a reflection of income distribution, for it disappears when we factor in the north/south split.

When we drill down even further into the data, however, we find that it is not really the south/north split that is at the heart of this issue either. As with homicides, the real cause of variations in infant mortality rates appears to be race. This can be seen in Figure 28c which plots racial composition against infant mortality. It reveals a clear trend line with a much steeper slope than in Figure 28a and with states clustering much closer to the line. The proportion of African-Americans in the population of a state explains more than half the variance in infant mortality rates.[134]

130 Infant mortality data from US Census Bureau, *Statistical Abstract of the United States* 2009, Table 111. R^2=0.143, p= 0.007. Their original plot can be found as Figure 6.6 in *The Spirit Level*

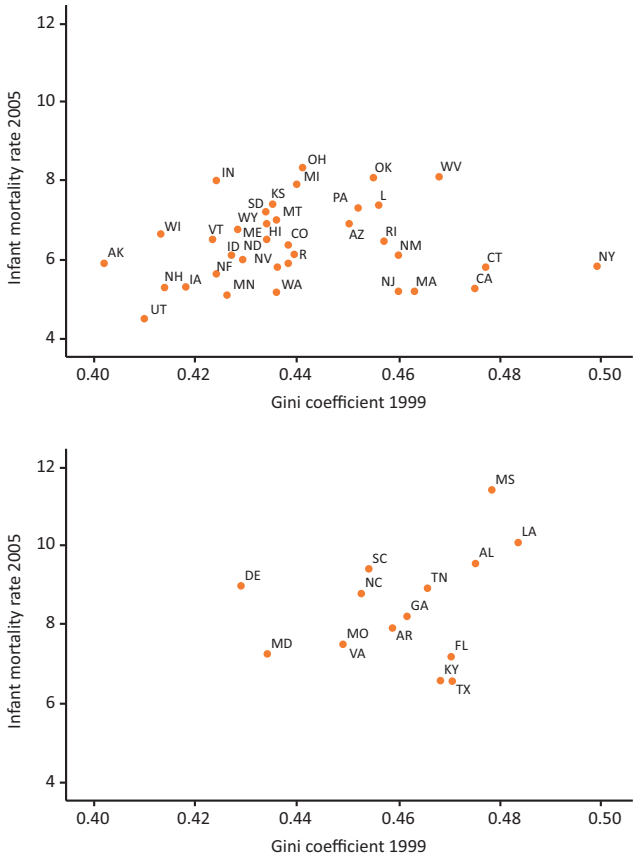
131 R^2= 0.105, sig=0.023.

132 Mean infant mortality rates = 8.179 (std dev 1.12) in the south and 6.394 (std dev 1.01) elsewhere. t= 5.415 with 47 degrees of freedom, p< 0.001.

133 For the 35 former free states, R^2 = 0.091, p= 0.536. For the 15 former slave states (Mississippi re-included) R^2 =0.012, p =0.273.

134 R^2 =0.544, p<0.001. If Mississippi is excluded, R^2 = 0.464, p<0.001.

Figure 28b: Infant mortality and income inequality, Deep South compared with the rest



A multiple regression model confirms that race, not income inequality, drives infant mortality in the US states. Income inequality hardly registers; race is the overwhelmingly important explanatory variable.[135] Figure 28d summarises the relative explanatory power of these three variables as demonstrated by this multiple regression model.[136]

135 With Mississippi excluded, model fit statistics for a model predicting infant mortality from the gini coefficient and racial composition are: $R^2 = 0.465$, $p<0.001$. Beta for gini coefficient = -0.030 (p = 0.812); Beta for proportion of African-Americans in population = 5.566, p< 0.001. Adding the Deep South dummy variable, race still comes through as the sole explanation: $R^2 = 0.487$, Gini coefficient Beta=-0.030 (p= 0.811); Racial composition Beta = 0.517 (p = 0.006); Deep South Beta = 0.233 (p = 0.171).

136 If we add information about the strength of the state economies (measured by GDP per head), we end up with an even stronger model which explains two-thirds of the variance in infant mortality, but race remains the dominant influence in this final model, and income inequality becomes completely insignificant. Adding GDP (but still excluding Mississippi) raises the adjusted R^2 of the model to 0.560 (<0.001). Beta values: race = 0.684 (p<0.001); GDP = -0.360 (p=0.003); Gini = 0.163 (p= 0.217). Whether the explanatory power of race is merely a composition effect (African-Americans have higher infant mortality, therefore states with a high black population have higher rates), or if it tells us something about the impact of racial diversity itself (a high proportion of ethnic minorities in a population results in poorer infant mortality outcomes – perhaps because health care is worse or welfare is less generous), is something we cannot answer from this data set.

Figure 28c: Infant mortality and size of black population (US states)

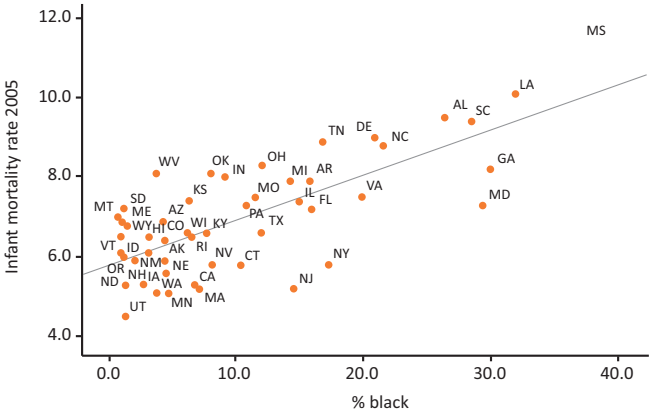
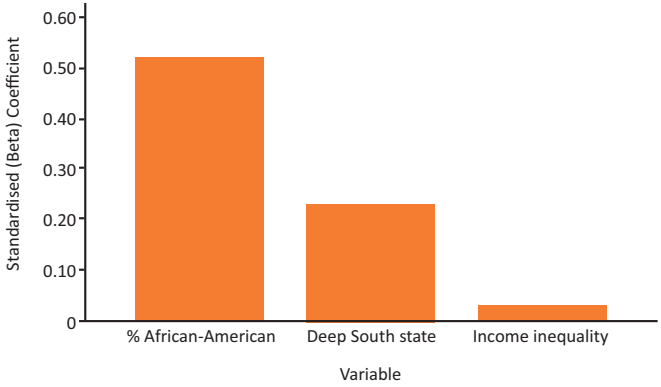


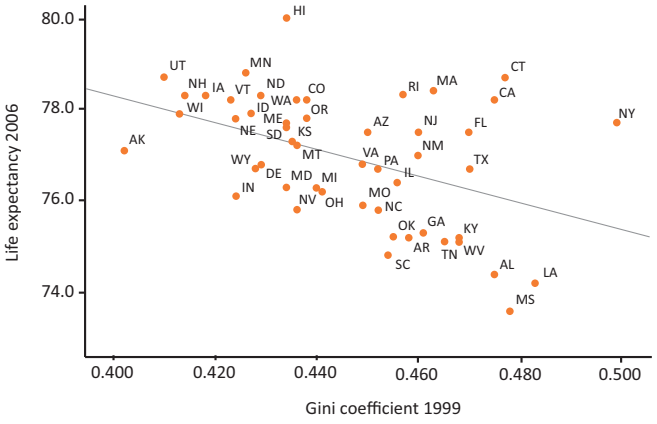
Figure 28d: The relative importance of income inequality, southern location and racial composition in determining a state's infant mortality rate



Conclusion: Infant mortality rates reflect the racial composition of the states (and to a lesser extent their wealth per head), but not their degree of income inequality, which is wholly unimportant.

Wilkinson and Pickett also look at life expectancy in the different states. Figure 29a is a reconstruction of their graph. Although we draw on different data sources, we arrive at exactly the same correlation (r=0.45), which suggests that income inequality explains about 20% of the variance in years of life expectancy across the different states. There are no outliers.

Figure 29a: Life expectancy and income inequality (US states)[137]



We can analyse these life expectancy statistics in almost exactly the same way that we analysed the infant mortality statistics. If we start by comparing life expectancy in the southern states with life expectancy elsewhere, we find it is two years lower – a significant difference.[138] If we then take the Deep South states out of the graph, the association elsewhere collapses (although it is still significant among the 15 southern states).[139] Next, if instead of income inequality, we substitute racial composition

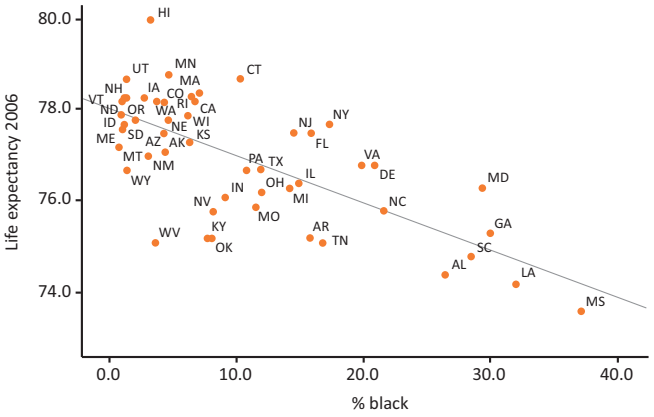
137 Their original graph is in *The Spirit Level* Fig 6.5. R^2 = 0.198, p = 0.001. Life expectancy statistics taken from *Business Week*, September 15 2006, quoting data from Harvard School of Public Health.

138 Mean life expectancy is 75.6 (std dev 1.1) in the former slave states, compared with 77.5 (std dev 1.1) elsewhere (t=-5.84 with 48 degrees of freedom, p<=0.001).

139 Without the 15 former slave states, R^2 = 0.027, p=0.308. Among the 15 southern states, R^2 =0.298, p=0.035

(the percentage of the population made up by African-Americans) as our predictor variable, we again end up with a much stronger model than the one presented in *The Spirit Level* (see Figure 29b).[140] Finally, if we add the wealth of the states (measured by GDP per head) to the mix, we can construct a simple but powerful model in which racial composition is the key explanatory variable, GDP is about half as powerful as race, and income inequality barely achieves statistical significance.[141]

Figure 29b: Life expectancy and size of the black population (US states)



Conclusion: Racial composition (measured by the proportion of African-Americans) is the most powerful predictor of average life expectancy in a state. GDP is also important. Income inequality appears marginal.

2.3 Literacy and numeracy

Wilkinson and Pickett find an association between the average maths/literacy scores achieved by eighth-graders and the extent of income inequality in the state in which they live. Their findings are

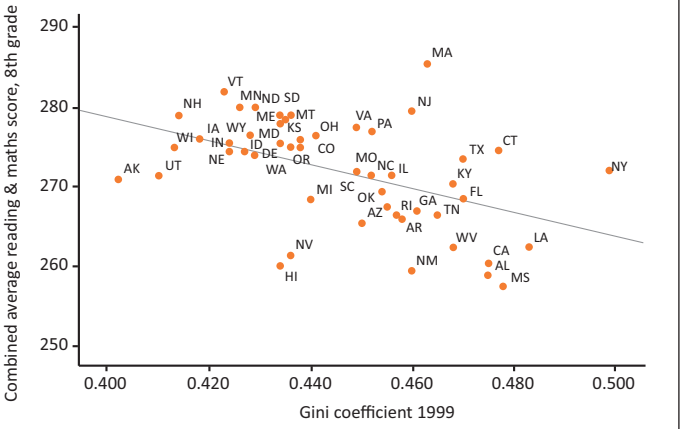
140 R^2 = 0.488, p<0.001

141 R^2 for the model = 0.565, p<0.001. Betas: race = -0.612 (p<0.001); GDP = 0.307 (p=0.009); gini coefficient = -0.264, p=0.050.

reproduced (using more up-to-date data, but with the same result: r=0.47) in Figure 30a.

It is a significant association, but not particularly strong (inequality explains 22% of the variance in education scores). However, in this case, it is hard to find *any* predictor variables that strongly influence the outcome, and inequality appears to do a better job than most.

Figure 30a: Maths and literacy scores and income inequality (US states)[142]



The racial make-up of the states does not appear to be particularly important in shaping students' maths and literacy performance. Like income inequality, it is statistically significant, but its influence is weak.[143] There is a significant difference in education scores between the former free and slave states, but again, this is not particularly striking.[144] And on this occasion, the wealth of the states does not seem to be important either. Indeed, a multiple regression model predicting education scores using all four of these independent variables finds only income inequality has an

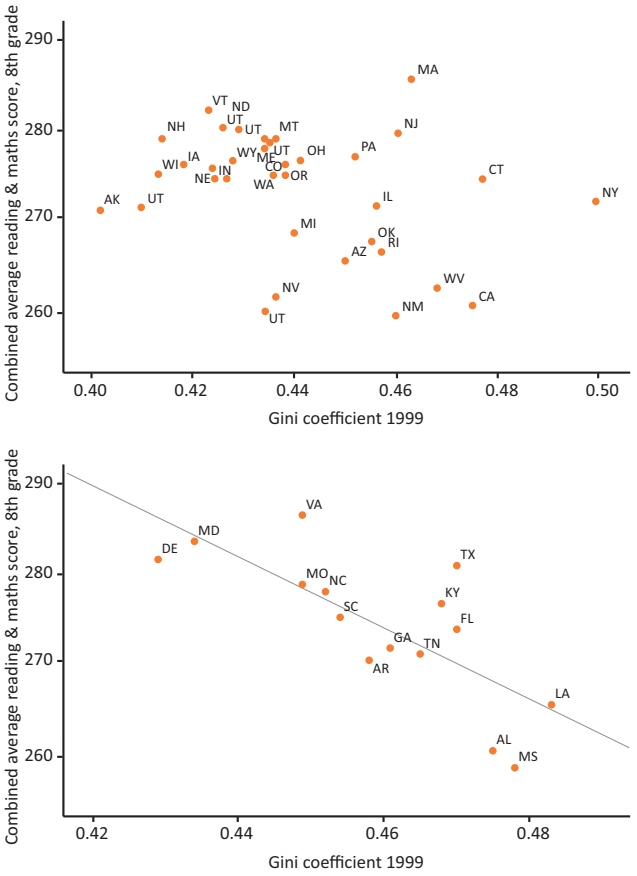
142 Scores are for 2007 and are taken from the National Center for Education Statistics, *Digest of Education Statistics*, 2008. Maths scores (from Table 122) and literacy scores (from Table 135) have been combined and averaged.. R^2=0.221, p=0.001.

143 R^2= 0.150, sig=0.005.

144 Average scores in the Deep South are 269, with a standard deviation of 6, compared with an average score elsewhere of 273, standard deviation 7. t = -2.278 with 48 df, p = 0.027.

effect; GDP, race and the Deep South dummy variable all fail to achieve statistical significance. But even this model only explains one-quarter of the variance.[145]

Figure 30b: Maths and literacy scores and income inequality, comparing southern states with the rest[146]



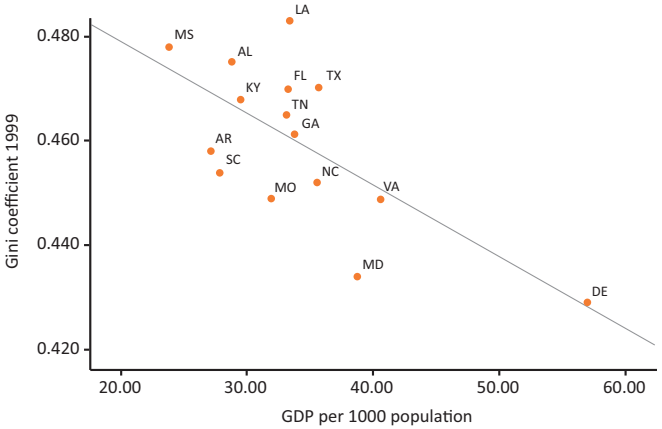
145 Model R^2= 0.250, p=0.008, Betas: gini coefficient = -0.439, p=0.015; GDP=0.129, p=0.390; race = -0.169, p=0. 444; Deep South = -0.029, p=0.977.

146 Among the 15 southern states, R^2=0.558, p=0.001. Among the other 35 states, R^2 = 0.091, p=0.077

A clue to what might be going on is provided if we divide the sample into southern and other states (Figure 30b). The association between income inequality and education scores disappears in the latter, but it strengthens in the former. We found a similar pattern with life expectancy (see note 139).

Why should income inequality affect education outcomes in the Deep South, but not elsewhere in the country? One possibility is that what we are picking up here is the effect of poverty, more than inequality. In the Deep South, the most unequal states are also the poorest (GDP predicts 43% of the variance in the gini coefficient in the 15 southern states – see Figure 31). So the apparent link between inequality and poor education outcomes in the south may be due in part to the lower standard of living characteristic of the more unequal southern states.

Figure 31: Income inequality and GDP per head in the fifteen southern states[147]



To test this, we can construct a multiple regression model predicting education scores in the 15 southern states from both their gini coefficient (income inequality) and their per capita GDP. The result is a very strong model, predicting three-quarters of the variance, in which both

147 Adjusted R^2=0.433, p=0.005

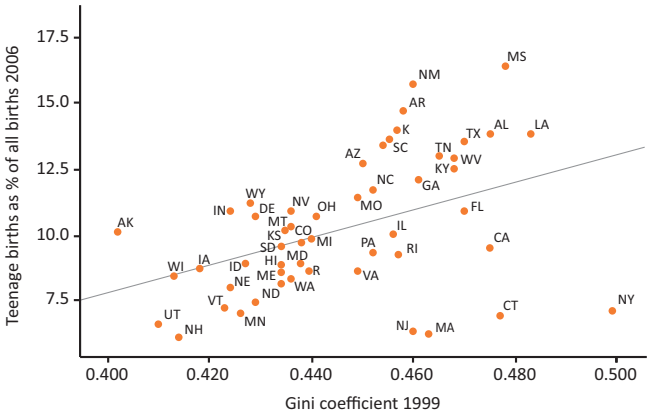
inequality and GDP appear significant.[148] It seems from this that income inequality is linked to educational performance in the Deep South, where it is partially explained by poverty, and may also be explained by variations in the quality of schooling between richer and poorer areas. Elsewhere in the country, however, inequality has little effect.

Conclusion: Education performance is only associated with income inequality in the poorer states of the Deep South.

3.4 Teenage births

In *The Spirit Level*, Wilkinson and Pickett use data on teenage births for their international comparisons, but they switch to include abortions and births to teenagers for their US comparisons. They never explain why they make this switch, and they only mention it in passing. In Figure 32a, I replicate their analysis of the US state data, but for teenage births only (in order to maintain comparability with the international data). Given that their combined birth/abortion figures correlate with income inequality at r=0.46, and my birth-only figures correlate with income inequality at r=0.44, it probably doesn't make much difference which we use.

Figure 32a: Teenage births and income inequality (US states)[149]



148 Adjusted R^2=0.755, p<0.001. Betas: gini coefficient=-0.622, p<0.001; GDP =3.645, p=0.003.

149 R^2=0.190, p=0.002. The corresponding plot in *The Spirit Level* is Figure 9.3

Teenage birth rate is a classic example of a phenomenon where we have to take cultural differences seriously, and Wilkinson and Pickett acknowledge this. They tell us, for example, that in the USA, Hispanic and African-American girls are almost twice as likely to be teenage mothers as white girls. However, they go on to claim that this has little relevance for their findings: 'Because these communities are minority populations, these differences don't actually have much impact on the ranking of countries and states by teenage pregnancy or birth rates, and so don't affect our interpretation of the link with inequality.'[150]

Table 2: States with African-American and Hispanic populations 1 or more standard deviations above the average[151]

% African-American (mean = 10.5, std dev = 9.5)		% Hispanic (mean = 9.9, std dev = 9.8)	
Mississippi	37.2	New Mexico	44.9
Louisiana	32.0	California	36.6
Georgia	30.0	Texas	36.5
Maryland	29.4	Arizona	30.1
South Carolina	28.5	Nevada	25.7
Alabama	26.4	Florida	21.0
North Carolina	21.6	Colorado	20.2
Delaware	20.9		

Considering the ethnic composition of some states, this looks an unsupportable claim for them to have made. In Mississippi, for example, 37% of the population is African-American, and there are eight southern states where African-Americans comprise more than one in five of the population. In New Mexico, California and Texas, more than one-third of the population is Hispanic, and there are another four states where more than one in five are Hispanic. If African-American and Hispanic girls are twice as likely as white girls to become teenage mothers, we

150 *The Spirit Level*, p.124

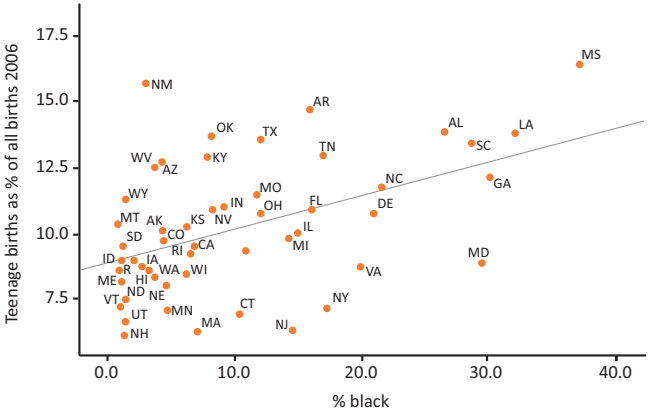
151 Source: US Census Bureau, *State and county quick facts* 2010

should expect this to have a substantial impact on the teenage birth rates of states with high African-American and Hispanic concentrations, and these make up almost one-third of all the states in the Union (see Table 2).

If we plot the teenage birth rates of the 50 states against the proportion of their population that is African-American, we get a graph (Figure 32b) that looks very similar to Wilkinson and Pickett's plot of teenage birth rates against income inequality, only we get a slightly better correlation than they do (r=0.47 against r=0.44).[152] If we then also take account of the size of the Hispanic population (in a multiple regression model), we end up accounting for 27% of the variance in the rate of teenage births in different states – still not strong, but substantially better than Wilkinson and Pickett's 19% of variance explained by their focus on income inequality.[153]

We can improve this model still further if we take account of the affluence or poverty of the states, as measured by their GDP per head. The proportion of variance explained now increases to 38%, indicating that teenage birth rates are influenced by economic as well as ethnic variations.[154]

Figure 32b: Teenage births and the size of the African-American population (US states)



152 R^2 = 0.220, p=0.001.

153 R^2=0.269, p=0.001 (although the % Hispanic variable just ceases to achieve statistical significance: Betas: African-Americans = 0.498 (p<0.001); Hispanic = 0.223 (p=0.081)). A boxplot reveals no outliers on teenage birth rates, so all states are included.

154 Model fit: R^2 = 0.380, p<0.001. All independent variables have a significant impact as shown by the Beta values: African-Americans = 0.602 (p<0.001); Hispanic = 0.453 (p = 0.003); GDP = -0.410 (p=0.006).

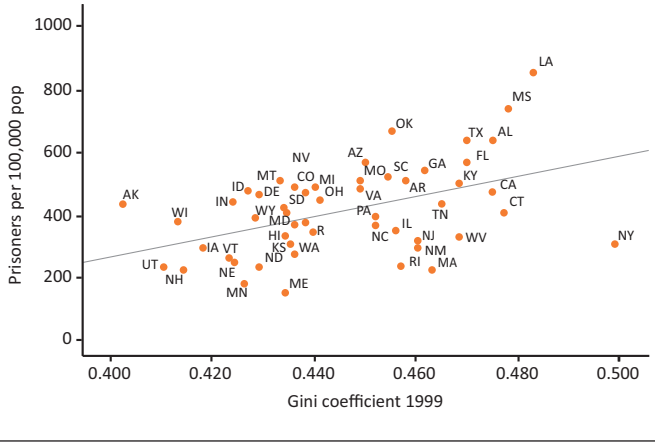
If we further add income inequality to this model, the fit again improves (to 44%), and all the predictor variables play a significant part in influencing the outcome.[155] This is, in fact, the best-fitting model we can achieve, and it suggests that several different factors – ethnicity, poverty and inequality – all contribute independently to a state's teenage birth rate.[156]

Conclusion: Ethnicity (the proportion of a state's population that is African-American or Hispanic) is a better predictor of the teenage birth rate than income inequality, although both factors (as well as the prosperity of a state) appear to have some influence.

3.5 Imprisonment rates

Figure 33a reconstructs Wilkinson and Pickett's plot of income inequality against imprisonment rates using more up-to-date statistics than the ones they use. The result is much the same as theirs (r = 0.47 compared with r = 0.48 in *The Spirit Level*).

Figure 33a: Income inequality and imprisonment rates (US states)[157]



155 Model fit: $R^2 = 0.444$, $p<0.001$. GDP now plays the most significant role in the model, although all variables have fairly similar explanatory power. Beta values: GDP = -0.511 (p=0.001); African-Americans = 0.425 (p=0.004); Hispanic = 0.386 (p=0.008); income inequality = 0.347 (p=0.028).

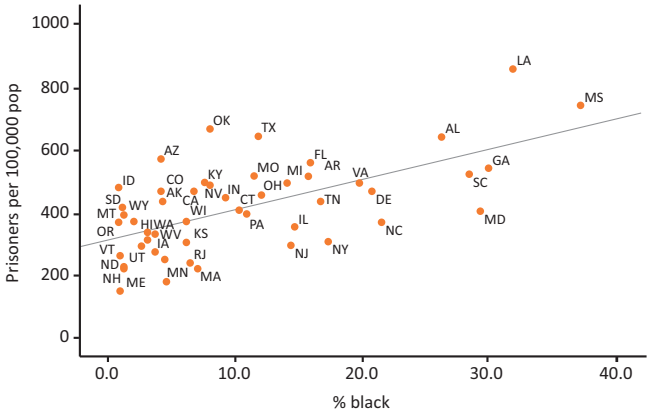
156 As with education outcomes, income inequality seems to be insignificant in influencing the teenage birth rate outside the south, but it retains a powerful effect when analysis is confined to the 15 southern states. In the 35 states outside the old Confederacy, knowing the extent of income inequality offers no help at all in predicting the teenage birth rate. But within the Deep South, this one variable predicts half the variance. Model fit for 35 non-southern states: $R^2 = 0.020$, $p = 0.416$. Model fit for 15 southern states: $R^2 = 0.498$, $p= 0.003$. Adding GDP per head raises the model fit for the southern states to 60% ($R^2 = 0.609$, $p = 0.004$), but GDP fails to achieve statistical significance in this expanded model ($p = 0.090$).

157 $R^2 = 0.220$, $p = 0.001$. Imprisonment rate is based on sentences of more than 12 months and is expressed per 100,000 population. Source: William Sabol, Heather West, Matthew Cooper, 'Prisoners in 2008' Bureau of Justice *Statistics Bulletin* December 2009, Table 10

This does not look a particularly strong association, and it is noticeable that many of the high imprisonment states are in the south (8 of the 10 states with the highest rates are former slave states). In the 15 southern states, the mean imprisonment rate is 543 per 100,000, compared with 355 per 100,000 elsewhere – more than 50% and 1.5 standard deviations higher. This difference is statistically highly significant, and it cannot be ignored.[158]

There is a strong relationship between the size of the African-American population and the imprisonment rate of a state. The correlation coefficient (r = 0.62) is a lot stronger than the one that Wilkinson and Pickett find for imprisonment and income inequality. Race explains 39% of the variance in imprisonment rates across the states (see Figure 32b).

Figure 33b: Imprisonment rates and the size of the African-American population (all US states)[159]



Once we take account of the racial composition of a state, income inequality becomes irrelevant in explaining its imprisonment rate.[160] Nor does the prosperity of a state, measured by its

158 Southern states: mean = 543.20, std dev = 127.3; other states: mean = 354.74, std dev = 113.9. t = 5.175 with 48 degrees of freedom, p < 0.001. If we split the sample between southern and other states, we find no association between inequality and imprisonment across most of the country, but this relationship seems to reappear in the south. For the 35 non-southern states, R^2 = 0.023, p = 0.381. For the 15 southern states, R^2 = 0.569, p = 0.001.

159 R^2 = 0.389, p<0. 001.

160 Fitting a model predicting imprisonment rates from the size of the African-American population and the gini coefficient produces model fit: R^2 = 0.414, p<0.001. Betas: African-Americans = 0.523 (p<0.001); inequality = 0.186 (p=0.167).

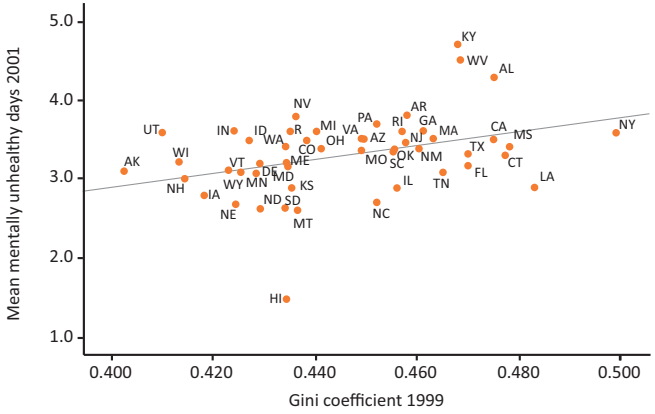
GDP, seem to make much difference either.[161] To explain imprisonment rates in the USA, the best predictor is simply the ethnic composition of the states.

Conclusion: Race explains imprisonment rates across the US states better than income inequality.

3.6 Mental illness – the limiting case

When discussing their US state data, Wilkinson and Pickett do not show their graph for mental illness against income inequality. The reason is that they failed to find a correlation. They say: 'We discovered something rather surprising. Alone among the numerous health and social problems we examine in this book, we found no relationship between adult male mental illness and income inequality among the US states' (they claim there is a weak association for women).[162]

Figure 34: Mental illness and income inequality (US states)



161 Fitting a model predicting imprisonment rates from the size of the African-American population and per capita GDP produces a very similar result: Model fit: $R^2 = 0.415$, $p<0.001$. Betas: African-Americans = 0.518 ($p<0.001$); GDP = 0.211 ($p=0.171$).

162 *The Spirit Level* p.68

They note that mental illness does not vary significantly by race, and they discuss 'the apparent resilience of ethnic minority populations to mental illness.'[163] But they 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.*

For the record, Figure 34 presents my attempt to reconstruct their missing graph. I have based it on a measure of mental health which counts the average number of 'mentally unhealthy days' reported by respondents.[164] This is not a very convincing or robust indicator, but it appears to be the same measure that Wilkinson and Pickett used, and it is about the best available at state level in the USA.

In fact, on my data, there is a weak association with income inequality – an R^2 of 0.132 (p=0.010). However, Hawaii is an extreme low outlier; Kentucky is an extreme high outlier; and West Virginia and Alabama are both high outliers. If these four are removed, no association remains among the remaining 46.[165]

Nor is there any association with ethnicity or the Deep South, nor, indeed, with GDP. This is telling. When there is no association with race, or the Deep South, Wilkinson and Pickett cannot find the results they are looking for on income inequality either.

Conclusion: The reason Wilkinson and Pickett cannot find a relationship between mental illness and income inequality in their American states data is because, unlike the other indicators they examine, there is no relationship between mental illness and race.

163 *The Spirit Level* p.68-9

164 2001 data from H. Zahran et al 'Health related quality of life surveillance' *MMWR Surveillance Summaries,* 2005, 54 (4) 1-35, Table 16

165 $R^2 = 0.070$, p = 0.076

4. Propaganda masquerading as science

In the previous two chapters we have examined in detail many of the key statistical claims on which Wilkinson and Pickett rest their argument that income inequality is responsible for social problems. In the great majority of cases, their claim fails to stand up, usually because it rests on a spurious correlation, and sometimes because the basic rules of regression analysis have been violated.

Table 3 summarises how Wilkinson and Pickett's claims have fared. Those where any element of validity has been found are shown in italics.

166 See note 124.

Table 3: Summary of status of Wilkinson and Pickett's statistical claims

Indicator	Association with income inequality (international data)	Association with income inequality (US State data)
Homicide	No association (spurious correlation depends on specific case of USA)	No association (spurious correlation created by association with race)
Childhood conflict	No association (spurious correlation depends on Scandinavian uniqueness)	Not analysed in *The Spirit Level*
Women's status	No association (spurious correlation depends on Scandinavian uniqueness)	Not tested here
Foreign aid	No association (spurious correlation depends on Scandinavian uniqueness)	Not analysed in *The Spirit Level*
Trust	*Weak association (but GDP is stronger influence)*	Not tested here[166]
Infant mortality	*Strong and significant association*	No association (spurious correlation created by association with race)
Life expectancy	No association (spurious correlation depends on specific case of Japan)	*Mild association – but race and GDP more important*

Indicator	Association with income inequality (international data)	Association with income inequality (US State data)
Obesity (adults)	No association (spurious correlation depends on specific case of USA)	Not tested here
Obesity (children)	No association (spurious correlation depends on specific case of USA)	Not tested here
Literacy & numeracy	Unreliable finding; regression conditions violated	*Association found only in Deep South*
Innovation & creativity	No association (spurious correlation depends on Scandinavian uniqueness)	Not analysed in *The Spirit Level*
Teenage births	No association (spurious correlation depends on Anglo and Scandinavian uniqueness)	*Some association, but race is a stronger predictor*
Imprisonment rates	*Some association, but does not hold in Eastern Europe and does not indicate harsh penal policies*	No association (spurious correlation created by association with race)
Social mobility	No association; data unreliable and regression conditions violated	Not analysed in *The Spirit Level*

Of 20 correlations examined, 14 have been shown to be wholly spurious or invalid. Contrary to Wilkinson and Pickett's claim, income inequality does not explain international homicide rates, childhood conflict, women's status, foreign aid donations, life expectancy, adult obesity, childhood obesity, literacy and numeracy, patents, or social mobility rates. Nor does it explain variations among US states in homicide, infant mortality or imprisonment rates.

In three cases (trust levels internationally, and teenage births and life expectancy in the US states), inequality appears to have some effect, but other variables are more important. In one case (education outcomes in the US), the association with inequality holds, but only in the southern states. In another one (imprisonment rates across countries) the association with inequality breaks down when eastern Europe is included, and the data fail to support Wilkinson

and Pickett's claim that less equal countries pursue harsher penal policies (quite the reverse appears to be the case). This leaves just one case (the association internationally between infant mortality and income inequality) where the evidence unambiguously supports their hypothesis and their claims.

Using a 95% probability threshold as our test of statistical significance, we should expect 1 in 20 correlations to appear significant, simply as a result of chance.

When the association runs the other way

In the course of examining the claims made in *The Spirit Level*, we have encountered a few examples of indicators where less equal nations appear to do better than their more egalitarian neighbours. For example, the Anglo nations score well on 'social capital' when this is measured by activity in voluntary organisations (there is a significant association between the strength of 'individualism' in national cultures and the level of activity in civil society). They also top the international list when it comes to charitable donations, and they appear less bigoted when asked how they feel about having somebody of a different race as a neighbour.

These are not the only examples where more unequal countries appear to perform better than less unequal ones. Wilkinson and Pickett themselves recognise that suicide tends to be 'more common in more equal countries.'[167] This is an interesting admission given that suicide rates were chosen by Emile Durkheim in his classic study of social pathology as *the* key indicator of social cohesion. According to Durkheim, suicide statistics are 'a sign and a result' of a 'collective malady.' A high suicide rate points to a 'state of deep disturbance' in society.[168] More than any other single indicator, he believed, the suicide rate is the canary in society's cage.

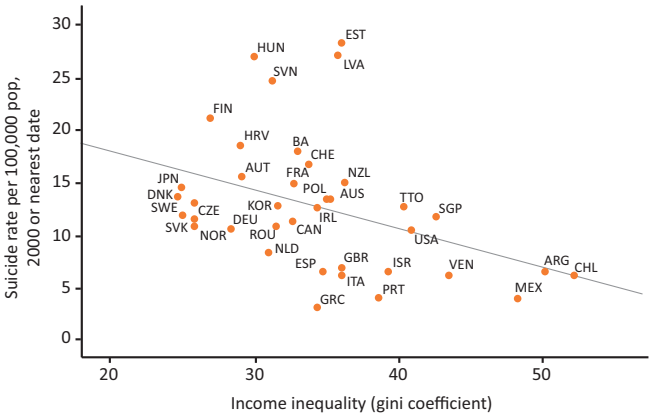
Wilkinson and Pickett do not provide any evidence on suicide rates in their book, nor do any international suicide statistics appear on their web site. In Figure 35, I have therefore created my own

167 *The Spirit Level*, p.175

168 Emile Durkheim, *Suicide* London, Routledge & Kegan Paul, 1952, p.391

scatterplot of income inequality against suicide rates. It is based on my expanded sample, although Lithuania and Russia have been omitted as outliers, and suicide statistics are missing for a number of other countries. We are left with a sample of 23.

Figure 35: Suicide rates and income inequality in 23 countries[169]



The graph confirms that suicide rates tend to get higher as the income distribution gets flatter. It is not a particularly strong association (r = 0.39), but it is statistically significant, and there are no obvious problems with the regression.[170]

Suicide is not the only indicator of social fragmentation which shows a tendency to increase as incomes get more equal. We find the same pattern when we look at divorce rates (Figure 36).

This time we have 36 countries in the sample and no outliers to exclude. Again, we find a modest but significant association with income equality (r = 0.39): as incomes become more equal, divorce rates fall. Given the misery and recriminations that often follow from failed marriages, it is difficult to argue that high

169 Source: E. Krug et al, *World Report on Violence and Health* Geneva, World Health Organisation, 2002, Table A9. Adjusted R^2 (because we are using sample statistics) = 0.130, p = 0.015

170 It may be objected that the official suicide statistics are unreliable. There is an enormous literature in sociology, provoked by Durkheim's work, which suggests that suicide statistics are 'socially constructed' and reflect different norms and classification procedures in different countries. But if we rule out the use of official suicide statistics on these grounds, we should probably rule out all the other official statistics that governments and research agencies collect, including those used by Wilkinson and Pickett. We must, of course, remain alert to possible problems in the statistics we are using, but there is no reason to single out suicide statistics for special concern. I have discussed some of these issues in Alan Buckingham and Peter Saunders, *The Survey Methods* Workbook, Polity Press, 2004, pp.27-35

divorce rates are compatible with a society that is working well. If people cannot rely on their spouses, where does this leave trust in other citizens? And if people cannot even get along with their partners, what does this say about levels of social integration more generally? Figure 36 would appear to represent another awkward challenge to the core claims of *The Spirit Level*.

Figure 36: Divorce rates and income inequality[171]

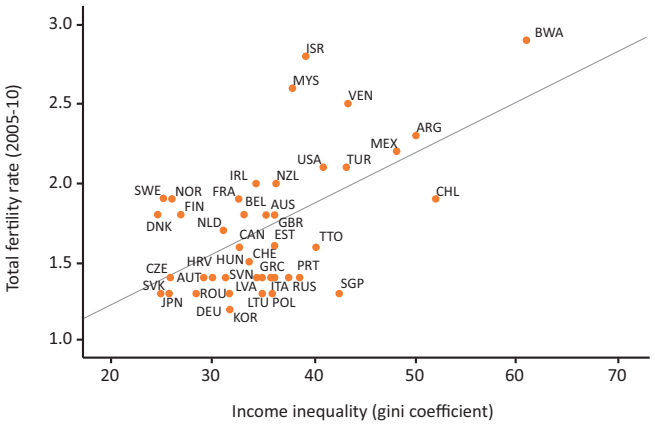
So does Figure 37, which plots fertility rates against income inequality. When people are happy, confident and optimistic, fertility rates are likely to be strong, for having children is an affirmation of faith in the future. When people feel pessimistic and fatalistic, on the other hand, fertility rates are likely to be lower. Given the sacrifices involved in raising children, it is also likely that selfish and egoistic societies will have fewer children than more altruistic ones.

Figure 37 shows that fertility rates are higher where incomes are less equal. Gabon has been excluded as an outlier, leaving 43 countries in the sample. The association between income inequality and

171 Crude divorce rate (divorces per 1000 people) 2002-06, from www.unstats.un.org, Table 25. Adjusted $R^2 = 0.125$, $p = 0.018$.

fertility appears quite strong (r = 0.57) and the statistical significance is robust. Given that many of the higher fertility countries are also among the poorest countries in this sample, it might be suspected that what we are finding is not an association of fertility with income distribution, but an association with GDP. But a multiple regression model predicting fertility rates finds this is not the case. Income inequality predicts fertility, but GDP per head does not.[172] We therefore appear to have a strong and robust finding: people have fewer children when their incomes are more equal.

Figure 37: Fertility rates and income inequality[173]



Alcohol consumption, too, seems to be higher in more equal countries. *The Spirit Level* claims that 'the use of illegal drugs, such as cocaine, marihuana and heroin, is more common in more unequal societies.'[174] This is not one of the claims I investigated earlier, but if we take out the quartet of high drug-use Anglo countries, there is no statistically significant association between drug use and income inequality across the remaining 18 countries in their analysis.[175]

172 Model adjusted R^2 = 0.311, p<0.001. Betas: gini coefficient = 0.630 (p<0.001); GDP = 0.130 (p=0.366).

173 Total fertility rate 2005-10 from UN *Human Development Report* 2009. Adjusted R^2 = 0.314, p<0.001.

174 *The Spirit Level*, p.70. The graph illustrating this claim is Figure 5.3.

175 On Wilkinson and Pickett's original regression, the adjusted R^2 = 0.368, p=0.002. Taking out the USA, UK, Australia and New Zealand leaves 18 countries remaining. Adjusted R^2 falls to 0.117, p=0.091.

More robust (but unexamined by Wilkinson and Pickett) is the association between alcohol consumption and income distribution – only this one comes out the other way around.

Figure 38: Alcohol consumption and income inequality[176]

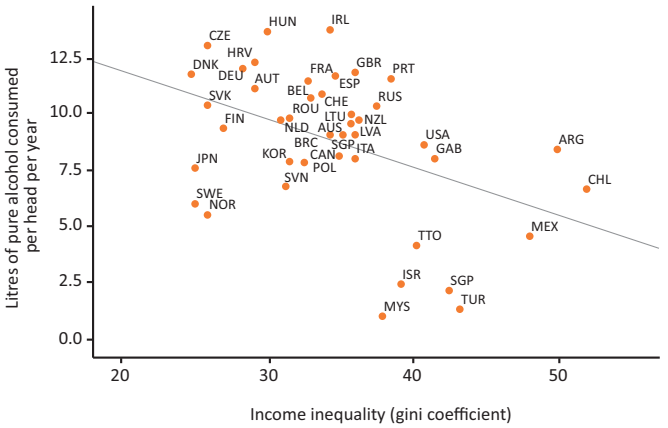


Figure 38 shows that you are more likely to drown your sorrows in alcohol if you live in a more equal country. The association is relatively modest (r = 0.44) but statistically significant. And again, if we control for the possible effect of national wealth, the association still stands up.[177]

Equality and the new 'Social Misery Index'

What are we to make of findings like these? Should we conclude that a more equal income distribution has been demonstrated scientifically to be a 'bad thing' because it is associated with a range of undesirable outcomes, including more suicides, higher divorce rates, a disinclination to have children and a tendency to resort to the bottle?

176 Source: World Health Organisation http://apps.who.int/globalatlas/default.asp. Adjusted R^2 = 0.169, p = 0.004.

177 In a multiple regression with income inequality and GDP per head as independent variables entered simultaneously, adjusted R^2 for the model = 0.152, p = 0.015. Betas: gini coefficient = -0.462 (p = 0.005); GDP = -0.068 (p = 0.665)

Following the logic adopted in *The Spirit Level*, perhaps we should. Wilkinson and Pickett construct what they call a 'Health and Social Problems Index' by combining their various statistics on trust, life expectancy, infant mortality, obesity, homicides and so on into a single measure. They then plot this index against income inequality. Not surprisingly (given that each constituent element correlates with income inequality) they come up with what looks like a strong finding. So pleased are they are with this result, that they reproduce it twice in their book.[178]

Mimicking this procedure, we might construct what we can call a 'Social Misery Index.' We can do it by trawling through the international comparative statistics to find any indicator which varies positively with income inequality (i.e. as inequality increases, the indicator improves). After a quick search on the internet, I came up with the following candidates:

- Racist bigotry (minding if a neighbour is a different race)
- Suicide rate
- Divorce rate
- Fertility rate (reverse coded)
- Alcohol consumption
- HIV infection rate

Only 20 of the 46 countries in my full sample have data on all six of these indicators, but that doesn't matter, for like Wilkinson and Pickett, we can still include countries where data are missing on some of the elements (only three of their 22 countries had full data on all the indicators they included in their index).[179] We do, however, drop four countries (Botswana, Gabon, Saudi Arabia and Venezuela) where data are missing on three or more of the indicators, leaving us with a sample of 42. Like Wilkinson and Pickett, it is then a simple matter to standardise all the scores on each indicator, add them up and average them to give each country a total Social Misery score.

178 It appears as Figure 2.2, and then again as Figure 13.1.

179 For my Social Misery Index, full data are available for Australia, Canada, Chile, France, Germany, Italy, Mexico, Netherlands, NZ, Norway, Poland, Romania, Russia, Slovenia, South Korea, Spain, Sweden, Switzerland, Trinidad & Tobago, UK. Another 16 countries have data missing on race but present on the other 5 (Austria, Belgium, Croatia, Czech, Denmark, Estonia, Greece, Hungary, Ireland, Israel, Japan. Latvia, Lithuania, Portugal, Singapore, Slovakia). Three (Argentina, Finland, USA) are missing divorce data but are present on the other 5, and one (Turkey) is missing suicide data but is present on the other 5. Libya is missing suicide and race statistics but is present on the other 4, and Malaysia is missing suicide and divorce data but is present on the other 4.

Not surprisingly, given what we already know about its component parts, our new Social Misery Index correlates positively and significantly with income inequality ($r = 0.50$, $p = 0.001$).[180] Russia and the Baltic states prevent the association from being even stronger than it is, for they appear exceptionally miserable even though they are only moderately egalitarian (maybe it is something to do with the dark Russian soul or the cold winters). Wilkinson and Pickett leave out a host of countries which should have been included in their sample according to their selection criteria, so let's massage our index a little by dropping these 4 troublesome cases. This gives us a much tidier result, and a stronger finding ($r = 0.64$). It is summarised by the plot in Figure 39.

> "As countries become more equal, life gets more miserable"

Figure 39: Social Misery is worse in more equal societies

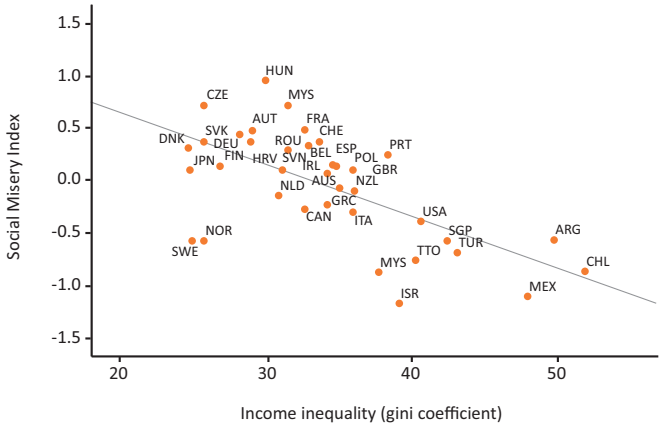


Figure 39 seems to show that, as countries become more equal, life gets more miserable. This conclusion is based on a strong and significant association found across 38 countries, including all

180 Adjusted $R^2 = 0.232$.

those covered by *The Spirit Level*, and many more that should have been included by Wilkinson and Pickett but were somehow overlooked.[181]

I could try backing up these empirical 'findings' with some theoretical speculation, mirroring Wilkinson and Pickett's musings about our genetic inheritance as humans. For example, I might argue:

> *Way back in the mists of evolutionary time, those individuals who were genetically programmed to share their goods died out, because they gave away food to others but did not get any back. This has left human beings with a deeply-embedded instinct to keep hold of their possessions, which is why egalitarian governments bent on income redistribution cause such misery (even among those who do not own very much). Income equalisation causes psychological distress which is then reflected in high scores on the Social Misery Index. A more unequal society provides us with a better-fitting shoe.*[182]

But the truth is that my Social Misery Index actually means very little, just as Wilkinson and Pickett's Health and Social Problems Index means little. My additional story-telling does not make my 'finding' any more compelling or persuasive than Wilkinson and Pickett's does theirs. The likelihood is that the modest differences in income inequality that we see between relatively wealthy countries have little or no impact on their people's wellbeing, either positively or negatively. Fine gradations of inequality do not cause more homicides, obesity or educational backwardness, any more than marginal increases in equality cause more divorces, suicides or HIV infection. As every social science undergraduate is taught when they first encounter statistics, correlation does not demonstrate causation.

Inequality over time

If the central argument of *The Spirit Level* were true, the deleterious effects of income inequality should show up in historical as well as geographical comparisons. In particular, if we see inequality

181 Adjusted $R^2 = 0.389$, $p < 0.001$

182 This theory is not entirely fanciful, although it runs directly counter to that advanced in *The Spirit Level*. It is broadly consistent, for example, with the argument developed by Richard Dawkins in *The Selfish Gene* (Oxford University Press, 1976), although we would need to add that altruism between close kin (especially parents and children) would also have been selected for in human evolution. I have discussed the idea of a 'possessive instinct' in Peter Saunders, *A Nation of Home Owners* (Unwin Hyman, 1990), pp.69-84

rising faster in one period than in another, or rising at one time and falling at another, then we should expect to see some evidence of changes in homicide rates, life expectancy, literacy levels, and so on.

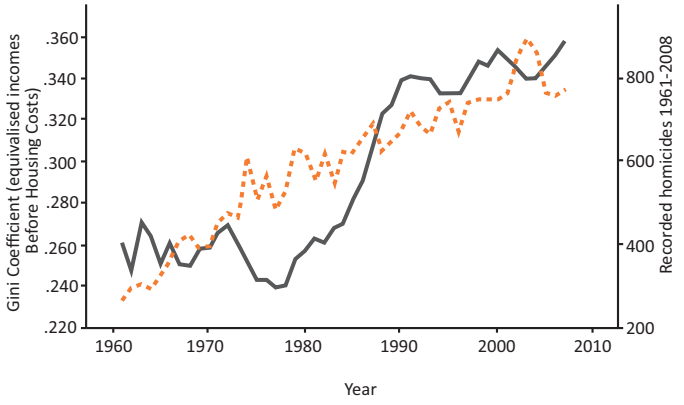
In reality, however, we see no such thing. The UK is an excellent test-bed for this, for as we saw in Figure 2, income inequality remained fairly constant here in the 1960s, fell a bit in the 1970s, rose quite sharply in the 1980s, and then more-or-less stabilised at its new, higher level. This provides us with enough variation over a sufficiently long time period to test Wilkinson and Pickett's belief that social problems intensify when inequality increases.

Many social problems have increased markedly in Britain since 1960, but when we look at the trends, the increase generally predates the increase in income inequality by some 20 years. It also tends to rise at a steady rate throughout the 1970s (when inequality was falling) and the 1980s (when inequality was sharply rising). None of this is consistent with the claim that inequality is the cause.

Take the *homicide rate*. This is a probably the most reliable indicator of crime that we have, for suspicious deaths are almost always reported to the police, and changes in the law and policing practices have little impact on the way murder is defined over time. The historical trend line is, therefore, a pretty accurate record of real changes in the murder rate.

We see from the dotted line in Figure 40 that this rate has been climbing steadily in the UK over the last 50 years. But this seems to have been unrelated to changes in income inequality (the solid line). Murders rose in the 1960s, when inequality was fairly stable; they kept rising in the 1970s, as inequality fell; they continued rising at much the same rate in the 1980s, as inequality rose; and they have continued rising in the last 15 years (although at a slightly slower average annual rate) as inequality has tended to flatten out.

Figure 40: Homicide rate and income inequality in Britain since 1961[183]

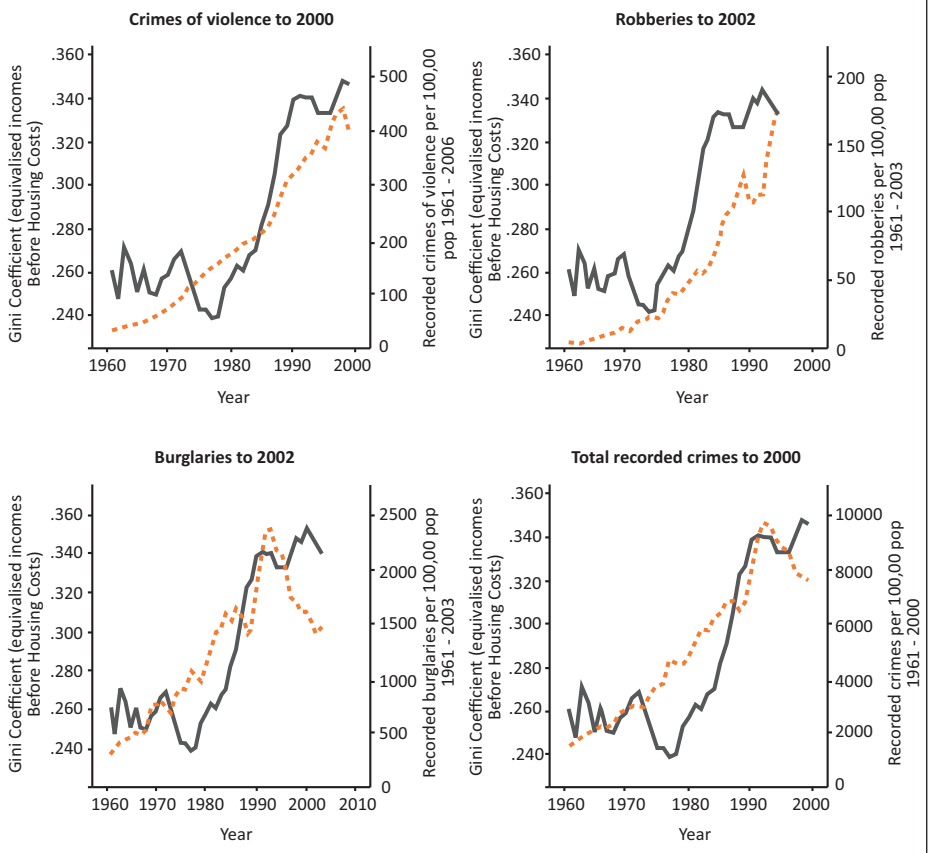


It is much the same story with other serious crime indicators (although the statistics here may be a bit less reliable, and the method for recording and measuring serious crimes changed in Britain in 2002, so it is difficult to construct a continuous series beyond that date). In Figure 41, I show the graphs for the period from 1961 to around 2000 for various categories of crime (depicted in the dotted lines), mapping each one against the trend of income inequality (shown by a solid line in each case).

It is clear in every graph in Figure 41 that the crime rate has been varying independently of changes in the distribution of incomes. Crimes of violence rose throughout the period but then fell in the late 1990s, while robberies rose throughout the period and then escalated in the later years. Burglaries rose until the early 1990s, then began falling significantly, a pattern which is to some extent replicated by the figures for all crimes. Whatever has been driving these trends (and we argued earlier that penal policy may have something to do with it), it is not income inequality.

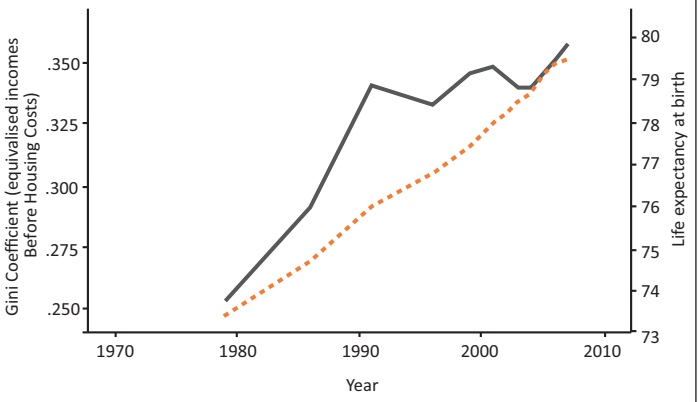
183 Homicides from Home *Office Recorded Crime Statistics* (http://www.homeoffice.gov.uk/rds/index.html)

Figure 41: Various categories of crime and income inequality in Britain since 1961



What is true of social pathology is also true of social wellbeing. Look, for example, at the trend in UK *life expectancy* (Figure 42a). It has been rising at a steady rate since 1979 (dotted line, plotted on the right-hand axis), and changes in the gini coefficient (solid line, left axis) appear to have made absolutely no difference to it.

Figure 42a: Life expectancy and income inequality in Britain since 1979[184]

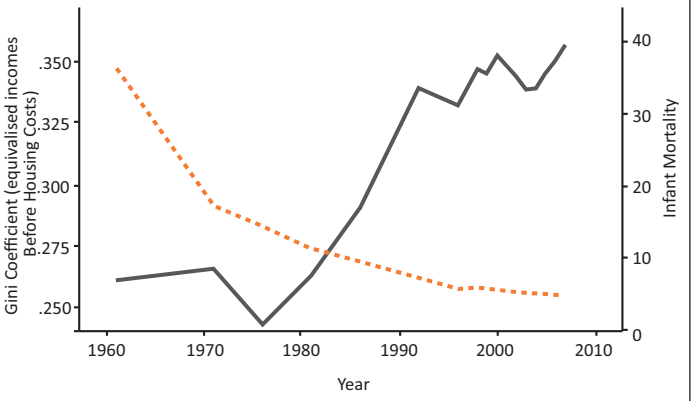


It is much the same story with *infant mortality*. This time we have data going back to 1961. These statistics confirm that infant mortality (the dotted line in Figure 42b) has fallen steadily throughout the last 50 years while income inequality (solid line) rose slightly, fell significantly, rose strongly and then began to level off. There is clearly no connection between the two, for variations in the latter have no resonance in the former.

International comparative research confirms that indicators like these have no connection with income inequality. During the 1980s and 1990s, income inequality rose more in some western countries than in others, so if Wilkinson and Pickett were correct, we should expect to find stronger improvements in life expectancy and infant mortality rates in the countries where inequality grew least. But we don't find this. Indeed, if anything, the evidence comes out the other way around.

184 Life expectancy data from Office of National Statistics *Vital Statistics online* (http://www.statistics.gov.uk/statbase/product.asp?vlnk=539). Male and female expectation of life at birth have been averaged. Income inequality from Institute for Fiscal Studies website, using gini coefficient before housing costs.

Figure 42b: Infant mortality and income inequality in Britain since 1961[185]



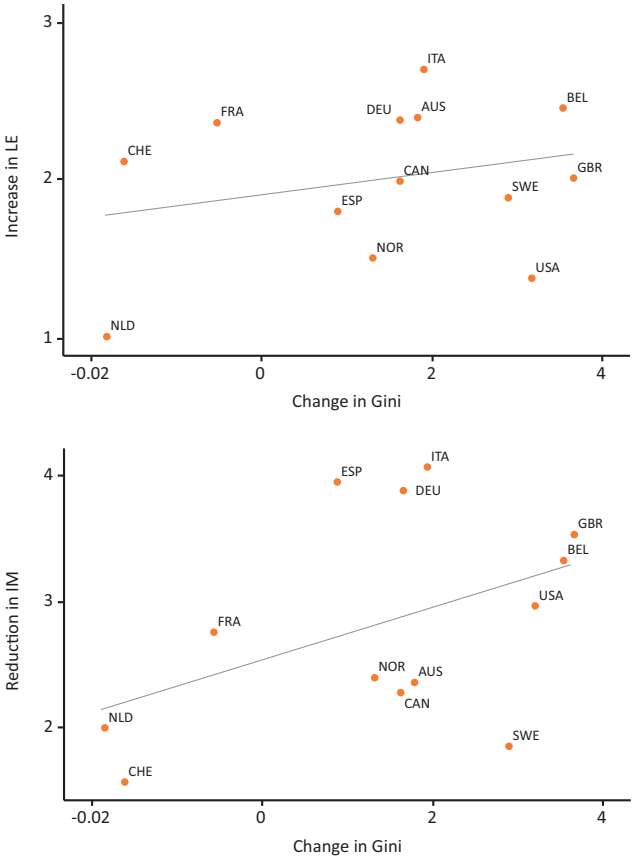
The results are summarised in Figure 43. The top graph shows that, far from reducing, life expectancy actually increased slightly more in the countries where inequality increased most during the 1980s and 1990s. The bottom graph likewise shows that infant mortality rates improved the most in the countries where inequality increased the most.

This evidence has been reviewed by three leading left-leaning economists who conclude: 'Although there are plausible reasons for anticipating a relationship between inequality and health (in either direction), the empirical evidence for such a relationship in rich countries is weak. A few high-quality studies find that inequality is negatively correlated with population health, but 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.'[186]

185 Infant mortality (deaths per 1000 live births) from Office for National Statistics, *Abstract of Annual Statistics* No.145, 2009, Table 5.20

186 Andrew Leigh, et al., 'Health and economic inequality'. Leigh notes on his blog : 'I'm about as anti-inequality an economist as you'll find', and he admits that he began his research 'secretly hoping to find that inequality was bad.' It is to his credit as a serious social scientist that he now openly accepts that the evidence cannot support such a position. http://andrewleigh.com/?p=2400

Figure 43: Increased income inequality did not reduce life expectancy (LE) or increase infant mortality (IM)in 13 European and North American countries in the 1980s and 1990s[187]



Is everybody in the same boat?

One reason *The Spirit Level* has achieved such a big political impact is its claim that: 'Equality is better for everyone.'[188] It is Wilkinson and Pickett's key contention that unequal societies not only suffer from

187 Andrew Leigh, Christopher Jencks and Andrew Smeeding, 'Health and economic inequality' In W. Salverda, B. Nolan, and T. Smeeding, editors, *The Oxford Handbook of Economic Inequality* (2009), p.4 in original paper.

188 This is the sub-title of the book.

lower average life expectancy, higher average infant mortality, more obesity, lower literacy levels, and so on, but that the rich in unequal countries are worse off on these various criteria than less well-off people in more equal nations.

But (with just one exception) the book doesn't actually show this. The exception is its analysis of infant mortality.

As we saw in Table 3, the international infant mortality data provide Wilkinson and Pickett with their strongest and most robust result. It really does seem that infant mortality rates are worse on average in more unequal countries (despite the fact that infant mortality improved most in the eighties and nineties in the countries where inequality grew fastest).

But the authors do not stop there. They also show that in egalitarian Sweden, the survival chances of babies born to lower class parents are actually better than the survival chances of babies born to higher class parents in the UK.[189] This is precisely the sort of evidence they need to back up their claim that 'equality is better for everyone.'

However, *The Spirit Level* presents no equivalent results to this on the other indicators it reviews. All that we are given on these other indicators are statistics for *average* differences between countries and states. We have seen that the analysis of these differences is often faulty, and that spurious 'findings' are presented which do not stand up to scrutiny. But even if these correlations were all valid, they would not actually demonstrate what the authors want to prove.

In an otherwise positive review of *The Spirit Level*, David Runciman pinpoints the problem succinctly, and he claims that the authors persistently 'fudge' it throughout the book: 'Is the basic claim here that in more equal societies almost everyone does better, or is it simply that everyone does better on average? Most of the time, Wilkinson and Pickett want to insist that it's the first... However, most of the data they rely on doesn't exactly say this.'[190]

189 *The Spirit Level*, Fig.13.4

190 David Runciman, 'How messy it all is' *London Review of Books*, vol.31, no.20, 22 October 2009, p.3

Runciman gives the example of imprisonment rates. The book shows that imprisonment rates are much higher in the USA than in less unequal countries (I reproduced their graph earlier as Figure 22a). But why should this mean that relatively affluent Americans would be better off living in a more egalitarian society? Even if 1 in 100 Americans are in jail, 99 out of 100 are not, so even with its very high incarceration rates, it is most unlikely that 'middle America' would benefit if flatter incomes resulted in a smaller prison population.

"It is impossible to judge whether more affluent people in more unequal countries really would be better off if their incomes were redistributed. In most cases, one suspects they would not"

To make its point, the book would need to show that middle class Americans are more likely to find themselves locked up than, say, lower class Swedes. As Runciman says, this may be the case (he notes the 'spectacular' prison terms that are sometimes handed out in the USA to fraudulent bankers and other white collar criminals), but Wilkinson and Pickett give us no evidence, one way or the other. They only discuss differences in *average* imprisonment rates, when what is required is statistics on imprisonment broken down by income groups across the different countries.

As another example, Runciman picks out Wilkinson and Pickett's analysis of international literacy data. They show that, on average, children in Finland perform better on literacy tests than British children do, and that this is true at all points in the social scale. But their data do *not* show that poor Finnish children do better than rich British children. Far from it. The relevant graph indicates that Finnish children from the least advantaged homes have an average literacy score below that achieved by British children from middle-ranking backgrounds.[191] This is a very different picture from that found for infant mortality, and it makes it impossible to argue that middle class British children would gain from a Finnish-style income redistribution.

191 The relevant graph in *The Spirit Level* is Figure 8.4 which plots average literacy scores achieved by children in 4 different countries against the level of education of their parents. At the lowest level of parental education (less than middle school), the average score for Finnish children is just below 280. This is a lot better than the average of just above 240 for British children with equivalent backgrounds. But British children whose parents completed High School also score just under 280 (it is impossible to extract exact figures from this graph), and those whose parents have 'less than college' or 'college and higher' educations score well in excess of this.

Time and again, the book makes the same fudge. Whether it be mental illness rates, homicide rates or obesity rates, the authors appeal to *average* differences between countries to support their argument that *everyone* would be better off if incomes were more equally distributed. Not only is their analysis of these average differences often faulty, but because the authors fail to provide evidence on how the rich and poor fare in these different countries, it is impossible to judge whether more affluent people in more unequal countries really would be better off if their incomes were redistributed. In most cases, one suspects they would not.

As another sympathetic reviewer, John Kay, concludes: 'They do not have the data to support a more general claim that equality benefits the rich as well as the poor... I suspect the claim that equality benefits everyone is just not supportable.'[192]

So what's so special about Sweden and Japan?

The Spirit Level does not show that health and wellbeing vary with the degree of income inequality in a country. Nor does it show that affluent people in unequal countries would benefit if the income distribution were made flatter. But it does show one thing. On a range of social indicators, Japan and the Scandinavian nations tend to come out ahead of the 'Anglosphere' countries such as UK, USA, Australia and New Zealand. The question is: why?

The most likely explanation lies in the history and cultures of these countries, but this is precisely the line of inquiry which Wilkinson and Pickett want to shut down:

- They say cultural factors cannot explain their findings because Japan and Sweden are very different cultures yet score similarly on various sorts of indicators.[193]
- They add that historical factors cannot be significant because countries end up equal or unequal for different historical reasons, but these differences do not change their performance on the various indicators.[194]

192 John Kay, 'The Spirit Level' *Financial Times* 23 March 2009

193 'Japan is, in other respects, as different as it could be from Sweden... Yet despite the differences, both countries do well – as their narrow income differences, but almost nothing else, would lead us to expect.' *The Spirit Level* pp. 183-4

194 'The degree of equality or inequality in every setting has its own particular history,' *The Spirit Level* pp. 188-9

Both these arguments are naive and unconvincing. It is true, of course, that Japan is in many respects a very different society from Sweden and the other Scandinavian countries. But there are common factors (in their histories, and in their cultures) which help explain why they both appear relatively cohesive, and why they both have such compressed income distributions.

Japan and Scandinavia were both 'late developers', agrarian societies which industrialised after Britain and the USA had become the world's leading industrial nations. They remain societies with a strong 'folk' tradition, a resilient sense of collective identity and an emphasis on national belonging and distinctiveness. Historically, they have been relatively closed, ethnically homogenous to a large degree (until recently, in Sweden's case, and still today, in Japan's), with low levels of immigration and very little inter-marriage with 'outsiders'.

In Japan, the 'insider/outsider' (*uchi/soto*) distinction underpins a marked sense of closure and exclusiveness which operates at all scales from the nation downwards. Japanese people 'have a strong sense of *uchi* about themselves as a nation.'[195] Children are taught from an early age to put the needs of the *uchi*, the insider group, above their own, and conformity to the group is reinforced through peer group pressure and the importance of bringing honour rather than shame to one's family and group. These strong collective norms probably help explain the low levels of crime and teenage births that Wilkinson and Pickett found, as well contributing to the consensual culture of Japanese enterprises and the compressed distribution of incomes between shop floor and boardroom.

Coinciding with this egalitarian ethos, however, is the huge significance attached to status hierarchies and pecking orders throughout Japanese society. Companies are ranked, schools and universities are ranked, ministries within the Civil Service are

195 Joy Hendry, *Understanding Japanese Society* 3rd edition, London, Routledge, 2003, p.137

ranked, communities are ranked, and household dynasties (the *ie*) are ranked. Enormous importance is placed on gaining entry to high-ranking institutions. As one observer notes, 'Japan is a society in which hierarchical ranking permeates personal interactions.'[196] It influences where you sit, how you speak, how low you bow, the gifts you exchange.

> "Grafting Sweden's tax and welfare system onto the USA, Australia or the UK would prove extremely difficult, and if attempted would almost certainly result in socially and economically disastrous outcomes"

This hierarchical aspect of Japanese society seems to have been overlooked by Wilkinson and Pickett, yet it poses a major challenge to their theory. They believe that income inequality causes social problems because of the psychological stress that hierarchy creates for individuals. But if this were true, Japan should appear at the opposite end of every one of their graphs, for while its income distribution might be compressed, its status antennae are as finely-tuned as in any society on Earth. If you want to experience the stress that striving for status can generate, have a look at a Japanese crammer school.

Sweden, too, is a remarkably cohesive society with a compressed distribution of incomes. Wilkinson and Pickett are right to point out that Sweden's egalitarianism was achieved through a different route than Japan's (in Sweden, relative income equality was brought about by government tax and welfare policies, rather than by a corporate culture of communalism). As in Japan, however, a relatively flat income distribution should be seen as an expression of common identity, not a cause of it.

The Swedish welfare state was conceived in the inter-war years to express the ideal of the *folksheim*, the 'People's Home,' where everybody could feel a sense of belonging. Housing, education, employment and welfare policies were all designed to underpin and strengthen this uniform national culture. As in Japan, the

196 Ron Dore, *Taking Japan Seriously* London, Athlone Press, 1987, p.86

schools play a crucial rule in transmitting the common culture, and individual deviancy (whether in the form of crime or socially unacceptable behaviour like heavy drinking) is closely monitored. In both countries, rule-breaking tends to be regarded more as a disease to be treated than as an act of individual will to be punished.[197]

Sweden's form of communalism was only possible, however, for as long as the country was ethnically and culturally homogenous, for as in Japan, the Swedish 'People's Home' did not include foreigners and outsiders. As immigration into Sweden has increased, the old bonds of collective sentiment and common identity have begun to fray, and there are clear signs that public support for the high-tax, generous welfare system is eroding as Swedes see funds being diverted to newcomers who are 'not like them' and 'cannot be trusted.'[198]

Now compare these histories and cultures with those of the Anglo countries. Australia, New Zealand and the USA are all settler nations, peopled initially from Britain which itself sat at the heart of a huge global empire. The Anglo tradition has for several centuries emphasised open borders and free trade rather than in-groups and out-groups. And in stark contrast to Japan and Sweden, English culture has for centuries been highly individualistic, viewing the State more as a threat to individual liberties than as the expression of common interests: 'A central and basic feature of English social structure has for long been the stress on the rights and privileges of the individual as against the wider group or the State.'[199] This is reflected to this day in Hofstede's Individualism Index, discussed in Chapter II, where we saw that the USA, Australia and Britain occupy the top three places on the international rankings. Based on their national wealth and location in the world, Hofstede predicted Individualism scores of 92 for Sweden and 66 for Britain. Their actual scores were 71 and 89 respectively.[200]

197 R. Huntford, *The New Totalitarians* London, Allen Lane, 1971. Wilkinson and Pickett are approving of this philosophy of 'treatment' rather than 'punishment', but they might be less enthusiastic if they read some Foucault.

198 See, for example, Maureen Eger ('Even in Sweden: The effect of immigration on support for welfare state spending' *European Sociological Review* vol., 2009, 1-15). She finds 'clear evidence that ethnic heterogeneity negatively affects support for social welfare expenditure – even in Sweden' (p.1). The same phenomenon has been noted in Denmark by Tyler Cohen who points to the distinction between citizenship as a shared culture and citizenship as a set of legal rights and entitlements: 'The welfare state is a means of expressing solidarity with people who are mostly just like you are. Other people with different values cannot be trusted not to abuse the system... In Scandinavia, luckily, the two definitions of "compatriot" largely describe the same group. But immigration is changing this; it drives a wedge between the two definitions, ultimately undercutting support for the public institutions Danes cherish' ('Something rotten in the [welfare] state of Denmark' Economist Free Exchange', www.economist.com/blogs/freeexchange, May 23, 2007)

199 A. Macfarlane, *The origins of English individualism*, Oxford, Basil Blackwell, 1978, p.5

200 *Cultures' Consequences* Table 5.1

Whereas Japan and Sweden retained their peasant roots into the early twentieth century, England had ceased to be a peasant country as early as the fourteenth century. Feudalism disappeared much earlier in England than on the continent, with individuals trading in free markets, buying and selling property (including land), moving around the country, marrying whom they wanted, and freeing themselves from the ties of lords, villages, churches and extended kinship networks. Long before the Renaissance, the Reformation or the Enlightenment, England was a mobile, acquisitive and intensely individualistic country. Alan MacFarlane goes so far as to suggest that: 'The majority of ordinary people in England from at least the thirteenth century were rampant individualists.'[201]

We do not have to make judgements as to the relative merits of Japanese or Swedish society, on the one hand, and the Anglosphere on the other. They each have their strengths and weaknesses. But we do have to recognise the implications of these differences when comparing social outcomes. Wilkinson and Pickett say they are not interested in culture and history. They don't think any of these details matter. But the specific historical and cultural factors we have been outlining go a long way in explaining why these countries appear in their graphs at different ends of their various regression lines.

Like several generations of left-wing Utopians before them, Wilkinson and Pickett think that America or Britain could be made to look just like Sweden, if only the income distribution were changed. As Marx nearly said, change the economic arrangements and the rest will follow. But Sweden and Japan have the income distributions they have because of the kinds of societies they are. They are not cohesive societies because their incomes are equally distributed; their incomes are equally distributed because they are cohesive societies. Grafting Sweden's tax and welfare system onto the USA, Australia or the

201 *The origins of English individualism* p.163. I have discussed this further in *A Nation of Home Owners*, and David Willetts has recently also drawn on Macfarlane's work in his *The Pinch* where he emphasises in particular the significance of limited kinship ties for the vitality of civil society in the Anglo countries.

202 See my 'Australia is not Sweden: National cultures and the welfare state' *Policy* vol. 17, no.3, 2001, 29-32. A generous welfare state superimposed on an individualistic culture which lacks a strong sense of collective responsibility looks like a recipe for extensive fraud and abuse of the system.

UK would prove extremely difficult, and if attempted would almost certainly result in socially and economically disastrous outcomes.[202]

Conclusion

The Spirit Level has little claim to validity. Most of the correlations on which the argument is based do not stand up, and other relevant research does not support it.

Income distribution is a legitimate issue for political debate. But the debate should not be contaminated by wonky statistics and spurious correlations. As I argued in Chapter I, there are strong ethical arguments for and against a more egalitarian income redistribution, and it is appropriate that these arguments should be aired and critically examined. What must not be allowed to happen, however, is for social scientists to pre-empt this debate with spurious claims that the issues can be resolved by the manipulation of a few statistics.

Nearly a hundred years ago, one of sociology's greatest thinkers, Max Weber, delivered a celebrated lecture called *Science as a Vocation*. In it, he expressed his concern that scientists and technical experts were trying to subvert the realm of values, where we all have to wrestle with hard ethical choices, by claiming that 'facts' can resolve our dilemmas for us. *The Spirit Level* is a prime example of what Weber was warning us about (it is telling that Wilkinson and Pickett originally wanted to call their book, *Evidence-based Politics*).[203]

According to Weber, science (even good science) cannot tell us what should be done: '"Scientific" pleading is meaningless in principle because the various value spheres of the world stand in irreconcilable conflict with each other.'[204] He insisted we have to make our own choices between the 'warring gods' of morality, for scientists cannot tell us what to do, or how to 'arrange our lives.' Only 'a prophet or a saviour' can do that, said Weber.[205]

203 *The Spirit Level*, p.ix

204 Max Weber, 'Science as a vocation' In W. Gerth and C. Wright Mills, editors, *From Max Weber* London, Routledge & Kegan Paul, 1948, p.147

205 'Science as a vocation' p.153

The evidence in *The Spirit Level* is weak, the analysis is superficial and the theory is unsupported. The case for radical income redistribution is no more compelling now than it was before this book was published.

Policy Exchange

For several centuries, the left and right have argued about economic inequality. While the left thinks it is wrong for some people to receive high incomes when others get much less, the right thinks it is wrong to flatten out the distribution of resources by taking money away from people who have worked hard or taken risks.

This is essentially an argument about ethical principles. But in 2009, two social scientists published a book which claimed that the debate over inequality could be resolved by looking at empirical evidence. In *The Spirit Level*, Richard Wilkinson and Kate Pickett claimed that egalitarian societies benefit rich and poor alike. Crime rates are lower, infant mortality is reduced, obesity is less prevalent, education standards are higher, average life expectancy is longer, social mobility is more extensive, and so on. They concluded that we would all benefit from a more egalitarian distribution of income.

Not surprisingly, *The Spirit Level* has been enthusiastically welcomed by left-wing social commentators, but its claims have not been subjected to rigorous examination. In this new report, Peter Saunders puts Wilkinson and Pickett's empirical evidence to the test and finds it deeply flawed. Their evidence is weak, their analysis is superficial and most of the correlations in their book do not stand up. Despite the enthusiastic reception this book has received from social commentators, its claims are unsupported. The ethical debate over inequality remains unresolved.

£10.00
ISBN: 978-1-906097-78-3

Policy Exchange
Clutha House
10 Storey's Gate
London SW1P 3AY

www.policyexchange.org.uk