

Inequality and Health: Long-Run Evidence from a Panel of Countries

Andrew Leigh

andrew.leigh@anu.edu.au
www.andrewleigh.com
Fellow, Economics Division
Research School of Social Sciences
Australian National University
ACT 0200, Australia
Ph: 011 612 6125 1374

Christopher Jencks

christopher_jencks@harvard.edu
http://ksgfaculty.harvard.edu/Christopher_Jencks
Malcolm Wiener Center for Social Policy
Kennedy School of Government
Harvard University
79 JFK Street, Cambridge MA 02138
Ph: 617 495 0546

This version: May 2005

Abstract

Is mortality higher in countries that are more unequal? To answer this question, we use a new source of data on inequality: tax data on the share of the richest 10 percent of the population. Within countries, changes in top income shares have been shown to proxy changes in other inequality measures, such as the Gini coefficient. Using data on top income shares from Australia, Canada, France, Ireland, the Netherlands, New Zealand, Switzerland, the UK and the US over the period 1905-2002, we investigate the relationship between inequality, life expectancy, and infant mortality. In the absence of country and year fixed effects, we find a positive relationship between inequality and mortality rates. However, in our preferred fixed effects specification, the relationship becomes small and statistically insignificant. Nor do we find support for the hypothesis that changes in the income share of the richest ten percent affect homicide or suicide rates.

Keywords: health, inequality, mortality, top incomes, homicide, suicide

JEL Codes: I12, N30

* Thanks to Philip Clarke, Andrew Clarkwest, David Cutler, Christian Dustmann, Campbell Murray, Joseph Newhouse, S.V. Subramanian and seminar participants at the Australian National University, the University of Adelaide, and the 2005 SOLE/EALE annual meetings for valuable feedback on earlier drafts.

1. Introduction

Do more unequal countries have worse health outcomes? More than 100 articles on this question have been published over the past two decades, but no consensus has emerged (Lynch et al, 2004a). One major problem has been the paucity of reliable historical data on inequality. As a result, studies have examined the relationship between inequality and health at a single point in time. Because economic inequality and mortality are likely to have common determinants, not all of which are measured, the cross-sectional relationship between inequality and health is unlikely to provide an unbiased estimate of the change in mortality when inequality changes.

We investigate this issue using a new source of data on economic inequality: panel data on the share of personal income held by the richest 10 percent of adults in Australia, Canada, France, Ireland, the Netherlands, New Zealand, Switzerland, the UK and the US. In many of these nine countries we have annual data for most of the twentieth century.¹ Our data allow us to control both year and country-specific fixed effects, thereby holding constant both stable unobservable country characteristics and annual changes in mortality that reflect common influences, such as the advent of antibiotics.

The existing literature on inequality and health is surveyed in Deaton (2003) and Lynch et al (2004a). Both reviews conclude that the theoretical stories suggesting a relationship between inequality and health are stronger than the empirical evidence. Five studies that use time series evidence from developed countries to analyze the inequality-health

¹ Our earliest observation is for France in 19056, and our latest observations are New Zealand and the US in 2002.

relationship are especially relevant to our analysis. Wilkinson (1989) and Sen (1999) focus on changes in life expectancy in the UK over the twentieth century and argue that mortality rates fell most rapidly when the income gap between rich and poor narrowed. However, their measures of inequality are relatively inexact, and they do not account for temporal variation in the effect of technological innovation. Focusing on the last thirty years of the twentieth century, Wilkinson (1996) argues that the rise in inequality in the US and UK during the 1980s is the key reason why the rate of decline in infant mortality slowed in the period 1975-85. By contrast, Deaton and Paxson (2001) find no systematic relationship between inequality and health in either the UK or the US from the mid-1970s to the mid-1990s. Similarly, a study by Lynch et al (2004b), which looked at 100-year national trends and 30-year regional trends in the US, found little evidence of a causal relationship between inequality and health.

To preview our findings, our preferred specifications find no statistically significant relationship between income inequality and population health, regardless of whether we measure health by infant mortality or life expectancy, use levels or differences, or control for education and health expenditure. These findings suggest that the relationship between inequality and health is either non-existent or too fragile to show up in a robustly estimated panel specification. The remainder of the paper is structured as follows. Section 2 presents a simple model of the relationship between inequality and health. Section 3 outlines our inequality and health data. Section 4 presents our results, and Section 5 concludes.

2. A simple model of the relationship between inequality and health

There are two basic channels through which inequality might affect an individual's health. These are commonly termed the "absolute income hypothesis" and the "relative income hypothesis." Under the absolute income hypothesis, it might be the case that health depends only on individual income, but that the marginal health gains from an extra dollar of income diminish as income rises. Figure 1 shows a stylized version of such a relationship. A mean-preserving transfer from the richer individual (R) to the poorer individual (P) will raise the health of P by more than it will lower the health of R. Thus more equal societies will have better health holding average income constant. Across countries, the relationship between *average* income and average life expectancy does follow a pattern similar to Figure 1 (Preston 1975; Deaton 2003), but the relationship is almost flat in countries with incomes more than half the US average.

The relative income hypothesis posits that inequality has an impact on health even after holding individual income constant. Several channels have been proposed.

- (a) **Crime:** Inequality has been shown to increase violent crime (Fajnzylber, Lederman & Loayza 2002), which in turn lowers life expectancy.
- (b) **Public spending on healthcare:** Alesina, Baqir and Easterly (1999) show that the average value of public goods to members of a community decreases as heterogeneity increases. Income heterogeneity (inequality) might therefore reduce both public health spending and public provision of individual medical care. Szreter (1988) shows that public sanitation reforms in the UK occurred only when the franchise was extended to the poor.

- (c) **Social capital:** Several studies have found that people in more unequal places tend to be less trusting (Knack and Keefer 1997; Alesina and LaFerrara 2002; Leigh 2003). Trust may, in turn, affect the provision of public health care. Kawachi et al (1997) find a cross-sectional relationship both between inequality and social capital and between social capital and mortality in American states.
- (d) **Relative deprivation:** If individuals measure their well-being by making comparisons with others who are more affluent than themselves, inequality might engender “[l]ow control, insecurity, and loss of self esteem” (Wilkinson 1997). A closely related set of arguments suggests that wider income differentials between rich and poor cause increase in stress.

With only aggregate data on inequality and health, it is extremely difficult to distinguish between the absolute income hypothesis and the relative income hypothesis. A useful way to see this is to combine both hypotheses algebraically. Here, we adapt the model presented in Gravelle, Wildman and Sutton (2002), who begin by hypothesizing that absolute individual income, y , is the only factor affecting an individual’s mortality risk, $m(y)$. The expression $m(y)$ can also be expressed in terms of individual income y and mean income \bar{y} through the following second order approximation:

$$m(y) \approx m(\bar{y}) + m'(\bar{y})(y - \bar{y}) + \frac{1}{2}m''(\bar{y})(y - \bar{y})^2 \quad (1)$$

We now introduce the relative income hypothesis. Suppose that individual mortality also depends on how an individual’s income compares with the mean income in some other

reference population, such as others in the same city, state, nation or workplace. For simplicity, let that reference group be the entire national population, and let the effect of inequality on individual mortality be a linear function of the variance of incomes in the population, $V(y)$.²

$$m(y) \approx \left\{ m(\bar{y}) + m'(\bar{y})(y - \bar{y}) + \frac{1}{2} m''(\bar{y})(y - \bar{y})^2 \right\} + \alpha V(y) \quad (2)$$

Taking expectations of each side:

$$Em(y) \approx m(\bar{y}) + m'(\bar{y})E(y - \bar{y}) + \frac{1}{2} m''(\bar{y})E(y - \bar{y})^2 + \alpha EV(y) \quad (3)$$

Which simplifies to:

$$Em(y) \approx m(\bar{y}) + \frac{1}{2} m''(\bar{y})V(y) + \alpha EV(y) \quad (4)$$

Note from equation (4) that:

- If the second derivative of mortality with respect to income is positive, then there will be a positive relationship between inequality and mortality.
- Using aggregate data, we will be unable to estimate the second derivative of mortality with respect to income, and therefore unable to distinguish between the absolute and

² Alternative assumptions about the form of the relationship between inequality and health are equally plausible but less tractable.

relative income hypotheses.³ However, if we do not find a relationship between inequality and health in aggregate data, it is likely either that both hypotheses are false or that they work in opposite directions, making the net effect close to zero.

One way of attempting to circumvent the aggregation problem is to estimate an equation which includes both \bar{y} and \bar{y}^2 . But if the square of average income does not equal the average squared income (i.e. $\bar{y}^2 \neq \frac{1}{N} \sum y^2$, as will usually be the case), using \bar{y}^2 will not solve the aggregation problem. In what follows, we experiment with specifications that include only \bar{y} and with specifications that include both \bar{y} and \bar{y}^2 .

Gravelle, Wildman and Sutton (2002) point out two further problems with most estimates of the relationship between inequality and health that use aggregate data. First, other country-specific factors may be correlated with both inequality and health. Second, the inequality measure may not capture the full effect of inequality on health. Because we have many observations for each country in our sample, we can control stable country-specific characteristics more effectively than previous studies. But because we only have data on the share of income received by the top income decile, our measure of inequality may not be as good as the measures in the previous literature. There is no consensus on the best measure of inequality for capturing the relationship between inequality and health, but most theories suggest that a measure that covers the full distribution, such as

³ Miller (2001) has shown that this argument only holds if the second-order approximation in equation 1 is exact. But while the second-order approximation is unlikely to be exact, existing data are also unlikely to distinguish the two effects at all precisely.

the Gini coefficient, is likely to do better than the share going to the top 10 percent.⁴ Yet so long as the share of the top 10 percent is positively correlated with the ‘best’ measure of inequality, the top 10 percent’s share should capture some of the relationship between aggregate health and other inequality measures. We return to this issue later.

One final concern is worth noting. Since sicker individuals are less likely to work, countries with lower health standards may have more unequal family incomes. The causal relationship between inequality and health can therefore run either from health to inequality as well as from inequality to health. With a long time series such as ours, we could use Granger causality tests to see whether lagged inequality affected current health or lagged health affected current inequality. But since we find no statistically significant relationship between inequality and health, we do not pursue this approach.

3. Data on inequality and health

Data quality has been a major problem in studies of the relationship between income inequality and health. As Judge, Mulligan and Benzeval note in their review of the literature:

“Many of the studies use multiple sources of income distribution data and/or data from a wide range of years, which makes comparability between countries questionable. Only five of the studies use data based on a measure of equivalent disposable income. In fact, we believe it is the generally poor quality of the

⁴ An exception is Waldmann (1992), who finds a strong positive relationship between the income share of the richest 5 percent of the population and the infant mortality rate. His measures are based on a cross-section of 57 countries, with inequality and infant mortality measured at around 1970.

income data that poses the most serious weakness in most of the studies we have reviewed.” (1998, 569)

Most cross-national studies have used measures of inequality from the Deininger and Squire dataset or the World Income Inequality Database, but Atkinson and Brandolini (2001) have shown that using higher-quality inequality data can substantially alter results. Judge, Mulligan and Benzeval use data from the Luxembourg Income Study (LIS), which uses a consistent measure of income, namely size-equivalized disposable income, for all countries and years. In cross-country inequality regressions. They find no significant relationship between inequality and either life expectancy or infant mortality.

However, using inequality measures from the LIS leads to a dramatic reduction in sample size. Judge, Mulligan and Benzeval have only 16 countries in their sample, and only 10 countries have more than one observation. Even if there were a causal relationship between inequality and health, it might be difficult to discern in samples as small as those available from LIS, particularly if we wish to hold a number of other factors constant.

Here, we measure inequality using the share of pretax income received by the richest 10 percent of the population. These data are drawn from Atkinson and Piketty (2005) and are derived from income reported to the tax authorities. The resulting measure of inequality is particularly sensitive to changes at the top of the distribution. To a large extent, the share of the top 10 percent depends on the share of income going to the top 1 percent. Regressing the top 10 percent share on the top 1 percent share in a specification

that includes country and year fixed effects, the coefficient on the top 1 percent share is 1.39 ($t=32.2$). Leigh (2004) also shows the results of regressing the top 10 percent share on measures of inequality taken from the Luxembourg Income Survey, all of which are based on equivalized after-tax household income. The relationship between the top 10 percent share and measures of inequality based on after-tax household income is positive but weak in a specification that does not include country or year fixed effects. However, once country and year fixed effects are included, the relationship is positive and statistically significant for most inequality measures. One plausible explanation for this pattern is that cross-national differences in both household structure and tax progressivity tend to persist over time; so while pre-tax and post-tax top income shares are only weakly correlated in levels, *changes* in pre-tax and post-tax top income shares are more strongly correlated. We should also note that while the top decile's share of pretax income is related to the Gini coefficient for equivalized post-tax income once we introduce country and year fixed effects ($\beta = 0.746$; $t = 3.5$) and the 90:50 ratio ($\beta = 0.890$; $t=3.2$), it is not related to the 50:10 ratio ($\beta = -0.137$; $t = -0.6$).

We use two measures of population health. The first is life expectancy at birth, which is effectively a composite measure of the probability that individuals of different ages died in the relevant year. Our second health measure is the infant mortality rate, which is the proportion of children born alive who died before their first birthday in the relevant year. We use the log of infant mortality since the rate asymptotes towards zero as population health improves. Partly because infant mortality has a strong impact on life expectancy, these measures are highly negatively correlated with one another. Regressing life

expectancy on infant mortality in a specification with country and year fixed effects, the coefficient on log infant mortality is -7.26, $t=-10.61$. Without country and year fixed effects, R^2 is 0.89.

All specifications control for real per capita GDP. Some specifications also control for the average educational attainment of the adult population, per capita public health spending, and per capita private health spending (in constant US dollars). These three measures are not available until 1960. Appendix Table 1 presents summary statistics. Further details on variable construction may be found in the Data Appendix. Figures 2, 3 and 4 show trends in inequality and health for each country. Figures 5 and 6 show the cross-sectional relationship between inequality and health in 1995, the last year for which we have these statistics for all nine countries. For life expectancy, the gradient is approximately zero and statistically insignificant. For infant mortality, the gradient is positive but still statistically insignificant.

Regressing two series with persistent trends on one another can give rise to the “spurious regression problem.” Before estimating the relationship between inequality and health, therefore, we carry out a standard diagnostic test for a unit root. In the case of the income share of the top decile, this presents a special problem, since the top decile’s share is bounded between 0.1 and 1 and therefore cannot technically have a unit root.

Nonetheless, since the top decile’s share never actually approaches either the upper or lower bound, the series might exhibit nonstationary behavior, and standard OLS regressions could then suffer from the spurious regression problem. We therefore follow

Atkinson and Leigh (2004), and transform our bounded share variable S into an unbounded variable by means of the transformation $\log S/(1-S)$ when conducting unit root tests.

Panel A of Table 1 presents Augmented Dickey-Fuller tests (Dickey and Fuller, 1981) against the null of a unit root, for the inequality and health variables. For the transformed top income share and life expectancy variables, we cannot reject the null hypothesis of a unit root for any of the countries (even at a 10% level of significance). However, for the infant mortality variable, the Dickey-Fuller test rejects a unit root for all countries. Panel B presents the results of a Johansen (1995) test for cointegration between the top 10% share and each of our two measures of population health. With four exceptions – life expectancy in the Netherlands and Switzerland, and infant mortality and life expectancy in Ireland – the trace statistic is below the 5% threshold at which we would typically judge the series to be cointegrated. Given that there are only 13 observations for the Irish specifications, these results can be safely ignored. In the regressions that follow, we estimate relationships using OLS but show specifications using both levels and differences.

4. Empirical strategy and results

Much of the existing literature relies on comparisons across countries at a single point in time or on changes over time within one or two countries. We therefore begin by estimating an equation similar to those sometimes reported in the literature:

$$m_{jt} = \alpha + \beta(\text{Share10})_{jt} + \gamma Z_{jt} + \varepsilon_{jt} \quad (5)$$

where m is a measure of mortality (life expectancy or infant mortality) for country j in year t , Share10 is the income share of the richest 10 percent of the population, Z is real GDP per capita, and ε is an error term. Standard errors are clustered at the country level.

Columns 1 and 2 of Table 2 present results from this specification, which does not include either country or year fixed effects. For both life expectancy and infant mortality, a rise in inequality is associated with a statistically significant rise in mortality. This accords with Waldmann (1992), who finds a strong positive relationship between the income share of the richest 5 percent of the population and the infant mortality rate in a cross-sectional regression.

However, there are good reasons to think that we need to take account of country-specific and time-specific factors. We therefore move progressively towards estimating an equation of the following form:

$$m_{jt} = \alpha + \beta(\text{Share10})_{jt} + \gamma Z_{jt} + \delta_j + \rho_t + \varepsilon_{jt} \quad (6)$$

in which δ is a country fixed effect, and ρ is a year fixed effect. By including a country-fixed effect, we capture a large set of unobservable country characteristics that might be correlated with both inequality and health. The year fixed effect term is intended to capture nonlinear time trends that are common to all countries, such as wars,

technological innovations that diffuse rapidly, such as measles and polio vaccines, and major epidemics such as influenza and HIV/AIDS.⁵

Columns 3 and 4 of Table 2 add country fixed effects to the regression. This change has minimal impact on the size and significance of the inequality coefficients. However, adding year fixed effects in Columns 5 makes the relationship between inequality and life expectancy both substantively and statistically insignificant. For infant mortality, adding year fixed effect in Column 6 cuts the apparent effect of inequality in half, although it is still significant at the 10% level.

Including both country and year fixed effects causes the relationship between inequality and health to become insignificant for both life expectancy and infant mortality (see columns 7 and 8). In each case, we test the joint significance of the fixed effects, and find that both country and year effects are highly significant. Overall, the results from Table 2 show that the time trend and cross-sectional relationships between inequality and health are not robust to including fixed effects and exploiting only within-country variation in inequality and mortality. This result suggests that the relationship between inequality and mortality may be driven by unobserved factors affecting both inequality and health, rather than being a causal relationship.⁶

⁵ Technological innovations do not, of course, reach all developed countries in exactly the same year. For example, Deaton and Paxson (2001) argue that technological innovations tend to reach the UK about four years after they arise in the US. However, we cannot include a country-by-year fixed effect, since that is the source of the variation in health that we use to identify the effect of inequality. Our results are, however, robust to excluding the years 1914-19 and 1939-45, which are the periods in which year effects vary most across countries.

⁶ For a discussion of the same issue in a different context, see Acemoglu et al (2005).

To further test the relationship between inequality and mortality in a fixed effects specification, Table 3 adds three more time-varying country characteristics to the regression. Columns 1 and 2 of Table 3 include only per capita GDP, making them identical to columns 7 and 8 of Table 2. Columns 3 and 4 add GDP². Columns 5 and 6 restrict the sample countries for which education and health expenditure variables are available, which means eliminating all observations prior to 1960. Columns 7 and 8 add the education and health spending variables in the regression. In all but one case, the relationship between inequality and health is statistically insignificant. The exception is column 3 of Table 3, in which higher top income shares appear to be correlated with *longer* life expectancy. Since this coefficient is only significant at the 10% level, and since one coefficient in eight could easily be significant at this level by chance, we are inclined to treat this result as noise.

Inequality may, of course, take some time to influence mortality. To assess this possibility we estimate the analog of equation (6) in first differences:

$$\Delta m_{jt} = \alpha + \beta(\Delta \text{Share10})_{jt} + \gamma \Delta Z_{jt} + \delta_j + \rho_t + \varepsilon_{jt} \quad (7)$$

We estimate this equation using 1-year, 5-year, and 10-year differences. Most of the theories outlined above imply that changes in inequality should have more impact over a 5-year or 10-year period than over a one year period. If the inequality affects health by affecting public spending on healthcare, by weakening the social fabric, or by creating a sense of relative deprivation, these effects will probably take some years to show up in

mortality statistics. Declines in the absolute income of the poor could, however, have immediate effects on infant mortality, although our measure of inequality would not necessarily capture such declines. If increases in inequality lead to more violent crime, this effect might also be felt within a year or two. Changes in the investment income of the top decile could also have relatively immediate effects on public spending for medical care and public health, creating a positive correlation between inequality and life expectancy.

Tables 4, 5 and 6 present the results of the regressions based on one-year, five-year, and ten-year differences. Columns 1 and 2 of each table show results for the full sample and control only for GDP, while columns 3 and 4 add GDP². Columns 5 and 6 estimate the same difference equations as columns 3 and 4 but only cover the years since 1960, for which we have data on educational attainment and healthcare spending data. Finally, columns 7 and 8 control for average educational attainment, public health spending, and private health spending in the post-1960 sample.

Of the 24 differenced specifications in Tables 4, 5 and 6, the coefficient on the income share of the top decile is statistically significant in only two regressions – column 6 of Table 4 (significant at the 5% level) and column 8 of Table 4 (significant at the 10% level). The two significant coefficients suggest that a rise in inequality is associated with a reduction in infant mortality, but both coefficients are small, and we would expect two coefficients out of twenty-four to be significant at the 10% level by chance alone.

Overall, there is no evidence of a robust relationship between inequality and health in these data.

The coefficients of the controls, in contrast, mostly accord with expectations. When both GDP and GDP² are included in the regression, and the sample covers the full period, GDP is positively associated with better health outcomes, while GDP² has a negative coefficient, suggesting that the protective effect of additional income diminishes as GDP rises. Public and private health spending are typically associated with better health outcomes, although their coefficients are only significant for infant mortality in the 5-year and 10-year differenced specifications. We find little systematic relationship between changes in educational attainment and changes in population health. The only significant effect of education on health is in the 1-year differenced specification, where a rise in education is associated with a *fall* in life expectancy.

Some researchers have suggested that even if there is no overall relationship between inequality and mortality, there may be a relationship between inequality and homicide (Deaton 2003; Lynch et al 2004a) or suicide (Lynch et al 2004b). To test this hypothesis we calculated annual homicide and suicide rates for each country starting in 1950, which is the first year in which such data are available for all countries in our sample. Details on variable construction are provided in the Data Appendix. Table 7 shows the relationship of inequality to the homicide and suicide rates per 100,000 people. As with infant mortality, we use the log of these rates, since they must asymptote towards zero. We tested these relationships using both levels and 1-year, 5-year and 10-year differences.

The coefficient of the top decile's income share was not statistically significant in any of these specifications.

5. Conclusion

While there is a strong consensus in the literature that the correlation between income and health is positive, there is much less agreement over the relationship between economic inequality and health. This paper has used a new measure of inequality – the income share of the richest 10 percent of the population – to test the relationship between inequality and health over a much longer interval than previous research. By holding constant country and time fixed effects, we have tried to circumvent some of the problems that have plagued past cross-country studies of inequality and health.

Our results showed that higher GDP is associated with better health outcomes, and that this effect declines as GDP rises. Without year fixed effects, we found that more inequality was associated with a decline in health standards. But once we included year effects that were invariant across countries, the relationship between inequality and health became statistically insignificant.

The confidence intervals around our estimates are sufficiently tight to make substantively important *detrimental* effects of inequality on population health quite unlikely. Consider the coefficients from the levels specification, controlling for GDP and GDP², which are shown in Columns 3 and 4 of Table 3. For life expectancy, the 95 percent confidence interval for the effect of a one percentage point increase in the income share of the top 10

percent includes no negative values larger than -0.03 years. This result implies that even a 10-point increase in the income share of the top decile would be unlikely to lower life expectancy by more than 0.3 years. For infant mortality, the upper bound of the 95% confidence interval is +0.6% (an additional 0.2 deaths per 1000, when evaluated at the mean). We also found no significant deleterious relationship between inequality and either homicide or suicide rates.

Our confidence intervals do not allow us to rule out the possibility that inequality *raises* life expectancy by a substantively significant amount, but since our confidence intervals also include zero, and since the literature has focused almost exclusively on the possibility that inequality lowers life expectancy, we do not think our data could justify the conclusion that the true effect is really positive. This conclusion also holds for infant mortality.

One possible explanation for our findings is that we have not measured the type of inequality that affects health. While the top decile's share of total income is highly correlated with the Gini coefficient, it is not correlated with the 50:10 ratio. If inequality at the bottom of the distribution is what matters, we are not measuring the relevant form of inequality. The other possible explanation for our findings is that the underlying relationship between inequality and health is either non-existent or too fragile as to show up in a specification such as this one. This would be consistent with a number of other careful cross-country papers, such as Judge, Mulligan and Benzeval (1998), and Deaton and Paxson (2001).

Data Appendix

Sources of top incomes data

Top incomes data are originally from the common data base provided in Atkinson and Piketty (2005), then converted to calendar year data by Leigh (2004). Note that in Australia and Canada, the tax unit is the individual, while in France, Ireland, the Netherlands, Switzerland and the United States, the tax unit is a married couple or single individuals. New Zealand and the United Kingdom switched from household to individual filing in 1953 and 1990 respectively (our specifications take account of these two shifts, which in any case had only a modest impact on the top 10% share). For Ireland, Brian Nolan notes that the personal income denominator may be overestimated (which would lead the top incomes shares to be underestimated). Although he suggests that one might inflate Irish top income shares by 25%, we opted not to do this.

Sources of life expectancy data

Most life expectancy at birth is taken from the Human Mortality Database (HMD), found at www.mortality.org. There are three exceptions:

- United States data are from the National Vital Statistics Reports, Vol.52, No.14, February 18, 2004, Table 12 (found at www.cdc.gov/nchs/about/major/dvs/mortdata.htm). For 1900-28, the figures are from death-registration states only. From 1929 onwards, they cover the entire US.
- Australian data are from Australian Bureau of Statistics, *Australian Historical Population Statistics*, ABS Catalog Number 3105.0.65.001, Table 48.
- Figures for Ireland are from Central Statistics Office (2004), *Irish Life Tables No. 14, 2001-2003* (available at www.cso.ie).

The following should also be noted:

- In the case of New Zealand, life expectancy from the HMD is available for 1937 onwards for Maori, non-Maori, and the total population, and from 1876 onwards for non-Maori only. We use the ratio of Maori to non-Maori life expectancy in 1937 and 1938 to form a consistent life expectancy series for the entire population from 1876-1936. This method assumes that the ratio of Maori to non-Maori life expectancy was the same in the pre-1937 period as in 1937-38.
- Although our inequality data cover the entire United Kingdom (including Ireland prior to 1921), the HMD only provides mortality figures for England and Wales (omitting Scotland and Northern Ireland). For the period 1999-2002, we update the HMD figures using National Statistics, “Life expectancy at birth by health and local authorities in the United Kingdom 1991-1993 to 2001-2003, including revised results for England and Wales 1991-1993 to 2000-2002” (available at www.statistics.gov.uk). For consistency, we continue to use only figures from England and Wales in 1999-2002.
- Figures for Canada 1997-2002 are from Statistics Canada (www.statcan.ca).
- Life expectancy is linearly interpolated for missing years.

Sources of infant mortality data

The infant mortality rate is measured as probability that a baby born live does not survive until its first birthday. This figure is typically expressed as a rate per 1000 births, and we follow this convention.

Most infant mortality data is taken from the Human Mortality Database (HMD), found at www.mortality.org. We use the tables *Life Tables by Year of Death (1x1)*, and calculate infant mortality as $q(x) * 1000$ for $x=0$, where $q(x)$ is the probability of death between exact ages x and $x+1$. There are three exceptions:

- US infant mortality is from the *Statistical Abstract of the United States*, Table No. HS-13. Live Births, Deaths, Infant Deaths, and Maternal Deaths: 1900 to 2001. Prior to 1960, this excludes Alaska and Hawaii. Beginning 1970, it excludes births to, and deaths of, nonresidents of the United States.
- Australian data are from Australian Bureau of Statistics, *Australian Historical Population Statistics*, ABS Catalog Number 3105.0.65.001, Table 46.
- Irish data is from Vital Statistics, 2001 Annual, p.137

Additionally:

- New Zealand data prior to 1937 are adjusted in the same manner as for life expectancy.
- UK infant mortality data only covers England and Wales.
- Canadian infant mortality for 1997-2002 is updated with figures from the Statistics Canada website.
- Infant mortality for missing years is interpolated log-linearly.

Sources of Homicide and Suicide Data

Homicide and suicide figures are from the World Health Organization Mortality Database (8 December 2004 update), available from www3.who.int/whosis/mort/. This database tabulates deaths by country back to 1950, classified according to the prevailing International Classification of Diseases system (ICD7-ICD10). Homicide and suicide rates are both expressed as rates per 100,000 people. Since the coding changes over time, it is useful to set out here the precise ICD codes that were used here.

Homicide: A149 and B050 from ICD7; A148 and B050 from ICD8; B55 from ICD9; 1102-1103, X85-X99, and Y00-Y09 from ICD10.

Suicide: A148 and B049 from ICD7; A147 and B049 from ICD8; B54 from ICD9; 1101 and X60-84 from ICD10

In missing years, homicide and suicide rates are linearly interpolated for all countries. For Australia and the United States, national homicide and suicide statistics were readily available, and WHO figures were checked against official figures. WHO homicide figures for Switzerland are not credible for the period 1995-2001, so we drop these years from our analysis.

GDP

GDP is real GDP per capita (measured in 1990 International Geary-Khamis dollars), from Angus Maddison, *The World Economy: Historical Statistics*. Data downloaded from www.eco.rug.nl/~Maddison/

Educational attainment

Educational attainment is the average number of years of schooling for the population aged 15 and over, from: Barro, R.J. and Lee, J.W. 1993. "International Comparisons of Educational Attainment" *Journal of Monetary Economics* 32: 363-394; Barro, R.J. and Lee, J.W. 1996. "International Measures of Schooling Years and Schooling Quality" *American Economic Review* 86: 218-223; and Barro, R.J. and Lee, J.W. 2000. "International Data on Educational Attainment: Updates and Implications" Center for International Development Working Paper 42. Cambridge, MA: CID.

Barro and Lee provide figures every 5 years from 1960-2000, and we linearly interpolate for intervening years (and linearly extrapolate after 2000). Data can be downloaded from www.cid.harvard.edu/ciddata/ciddata.html.

Health expenditure

Health expenditure is from *OECD Health Data 2004* (updated September 24, 2004), downloaded from www.oecd.org/health/healthdata. We use two variables, public health expenditure per capita, and private health expenditure per capita (created as real total health expenditure per capita minus real public health expenditure per capita). Both are supplied by the OECD database in US\$ (converted at purchasing power parity). We then adjust for inflation by converting these amounts into 2003 dollars using the CPI-U-RS.

References

- Acemoglu, D., Johnson, S., Robinson, J. and Yared, P. 2005. "From Education to Democracy?" NBER Working Paper 11204. Cambridge, MA: NBER
- Alesina, A., Baqir, R. and Easterly, W. 1999. "Public Goods and Ethnic Divisions" *Quarterly Journal of Economics* 114: 1243-1284
- Alesina, A and La Ferrara, E. 2002. "Who trusts others?" *Journal of Public Economics*. 85: 207-34
- Atkinson, A and Brandolini, A. 2001. "Promise and Pitfalls in the Use of "Secondary" Data-Sets: Income Inequality in OECD Countries as a Case Study". *Journal of Economic Literature*. 39: 771-99
- Atkinson, A. and Leigh, A. 2004. "Understanding the Distribution of Top Incomes in Anglo-Saxon Countries over the Twentieth Century", *mimeo*
- Atkinson, A. and Piketty, T. (eds) 2005. *Top Incomes over the Twentieth Century: A Contrast Between European and English-Speaking Countries*, Oxford: Oxford University Press
- Deaton, A. 2003. "Health, Inequality, and Economic Development" *Journal of Economic Literature* 41: 113-158
- Deaton, A and Paxson, C. 2001. "Mortality, Income, and Income Inequality Over Time in Britain and the United States" NBER Working Paper 8534. Cambridge, MA: NBER
- Dickey, D and Fuller, W.A. 1981. "Likelihood Ratio Statistics for Autoregressive Time Series with a Unit Root" *Econometrica* 49: 1057-72
- Fajnzylber, P, Lederman, D, and Loayza, N. 2002. "Inequality and Violent Crime". *Journal of Law and Economics*. 45(1): 1-40
- Gravelle, H, Wildman, J. and Sutton, M. 2002. "Income, income inequality and health: what can we learn from aggregate data?" *Social Science and Medicine* 54: 577-589
- Hildebrand, V and Van Kerm, P. 2005. "Income Inequality and Self-Reported Health Status: Evidence from the European Community Household Panel", IRISS Working Paper 2005-01
- Johansen, S. 1995. *Likelihood-Based Inference in Cointegrated Vector Auto-Regressive Models*. Oxford: Oxford University Press
- Judge, K, Mulligan, J and Benzeval, M. 1998. "Income Inequality and Population Health" *Social Science and Medicine* 46(4-5): 567-579

- Kawachi, I., Kennedy, B.P., Lochner K, and Prothrow-Stith, D. 1997. "Social capital, income inequality, and mortality" *American Journal of Public Health* 87(9): 1491-1498
- Knack, S and Keefer, P. 1997. "Does Social Capital Have an Economic Payoff? A Cross-Country Investigation". *Quarterly Journal of Economics*. 112(4): 1251-88
- Leigh, A. 2003. "Does Equality Lead to Fraternity", *mimeo*
- Leigh, A. 2004. "Using Panel Data on Top Income Shares to Analyze the Causes and Effects of Inequality", *mimeo*
- Lynch, J et al. 2004a. "Is Income Inequality a Determinant of Population Health? Part 1. A Systematic Review" *Milbank Quarterly* 82(1): 5-99
- Lynch, J et al. 2004b. "U.S. National and Regional Trends in Income Inequality and Age- and Cause-Specific Mortality" *Milbank Quarterly* 82(2): 355-400
- Miller, Douglas. 2001. "Income Inequality and Mortality in the US: Aggregate Data and Micro Relationships." University of California, Berkeley, *mimeo*.
- Preston, S.H. 1975. "The Changing Relation Between Mortality and Level of Economic Development" *Population Studies* 29: 231-248
- Sen, A. 1999. *Development as Freedom*. New York: Knopf
- Szreter, S. 1988. "The Importance of Social Intervention in Britain's Mortality Decline c.1850-1914: A Reinterpretation of the Role of Public Health" *Social History and Medicine* 1: 1-37
- Waldmann, R. 1992. "Income Distribution and Infant Mortality" *Quarterly Journal of Economics* 107(4): 1283-1302
- Wilkinson, R.G. 1989. "Class Mortality Differentials, Income Distribution and Trends in Poverty 1921-1981" *Journal of Social Policy* 18: 307-335
- Wilkinson, R.G. 1996. *Unhealthy Societies: The Affliction of Inequality*. London: Routledge.
- Wilkinson, R.G. 1997. "Socioeconomic determinants of health: Health inequalities: relative or absolute material standards?" *British Medical Journal* 324(7343): 978 - 978.

Figure 1: A Non-Linear Relationship Between Income and Health

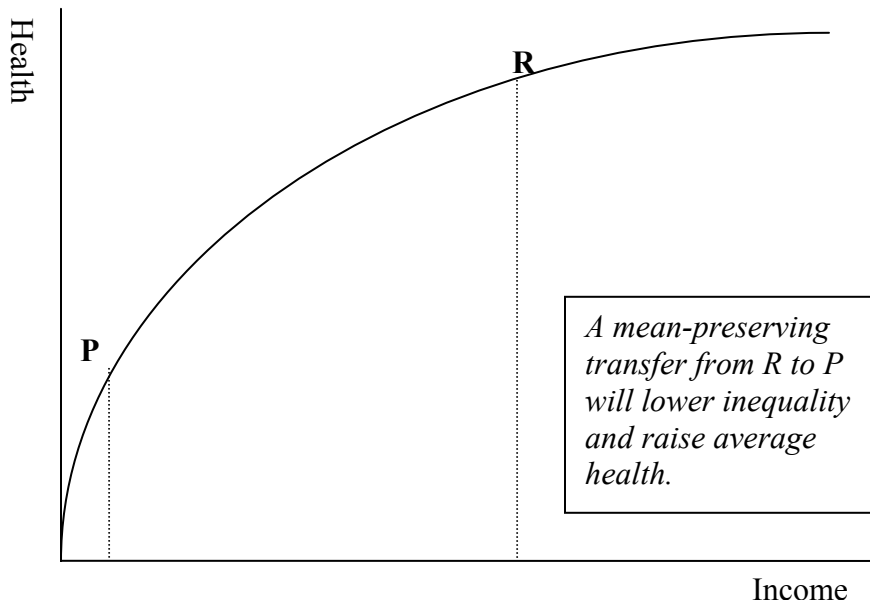
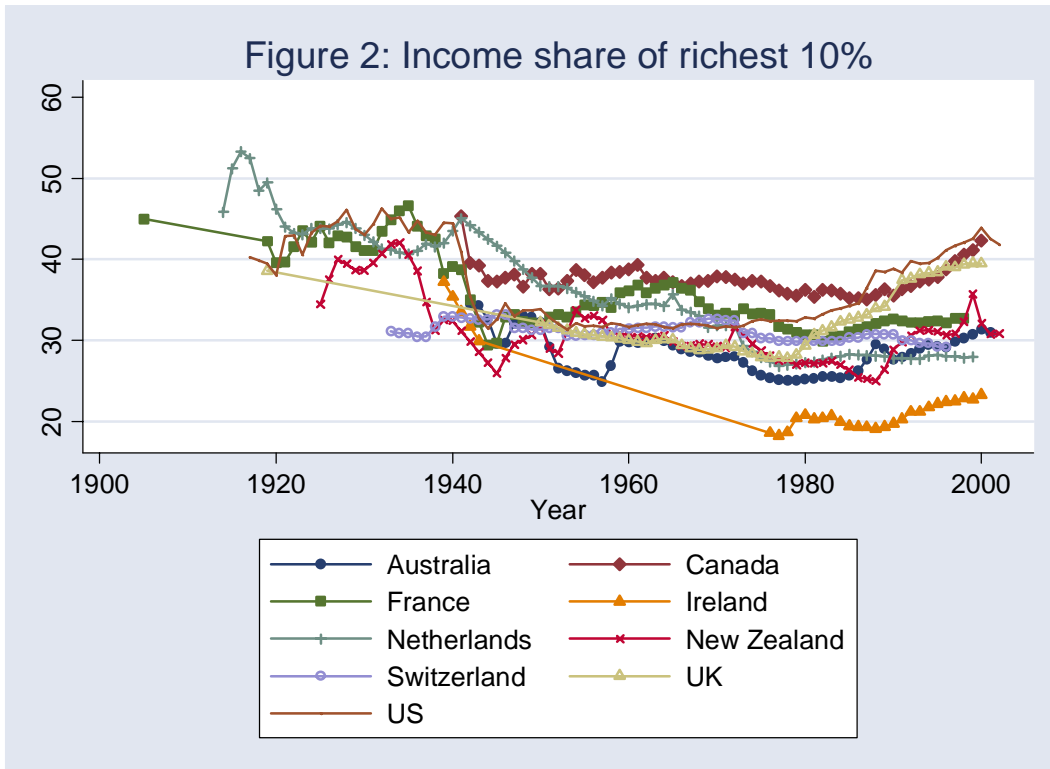


Figure 2: Income share of richest 10%



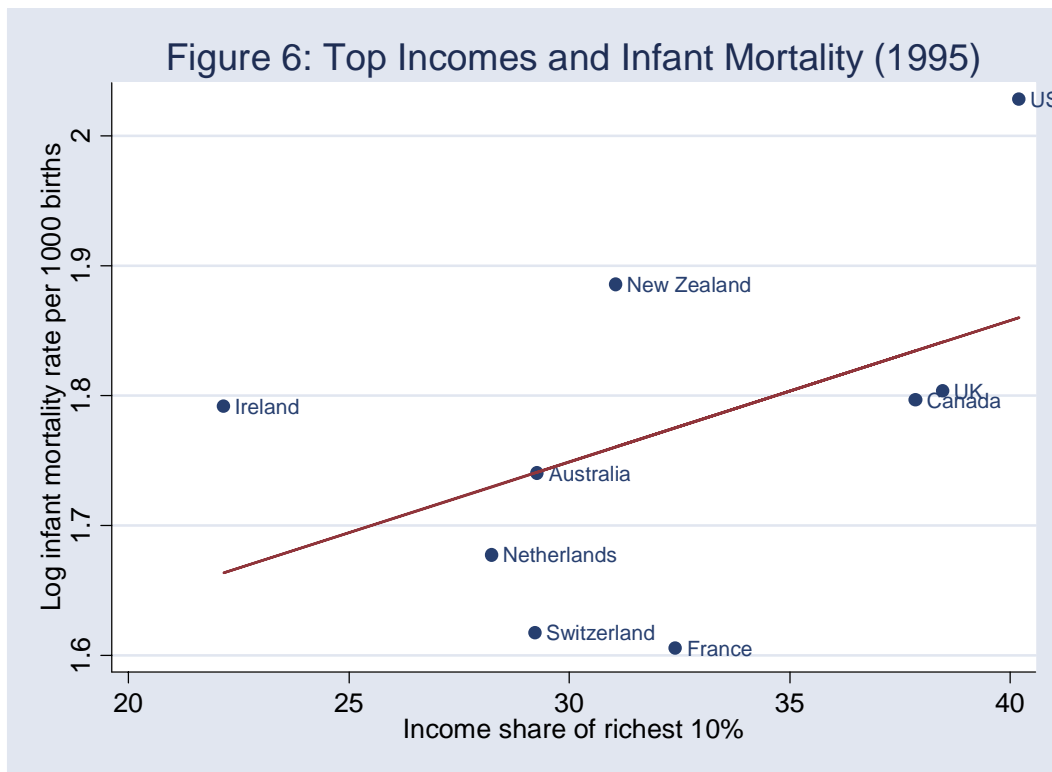
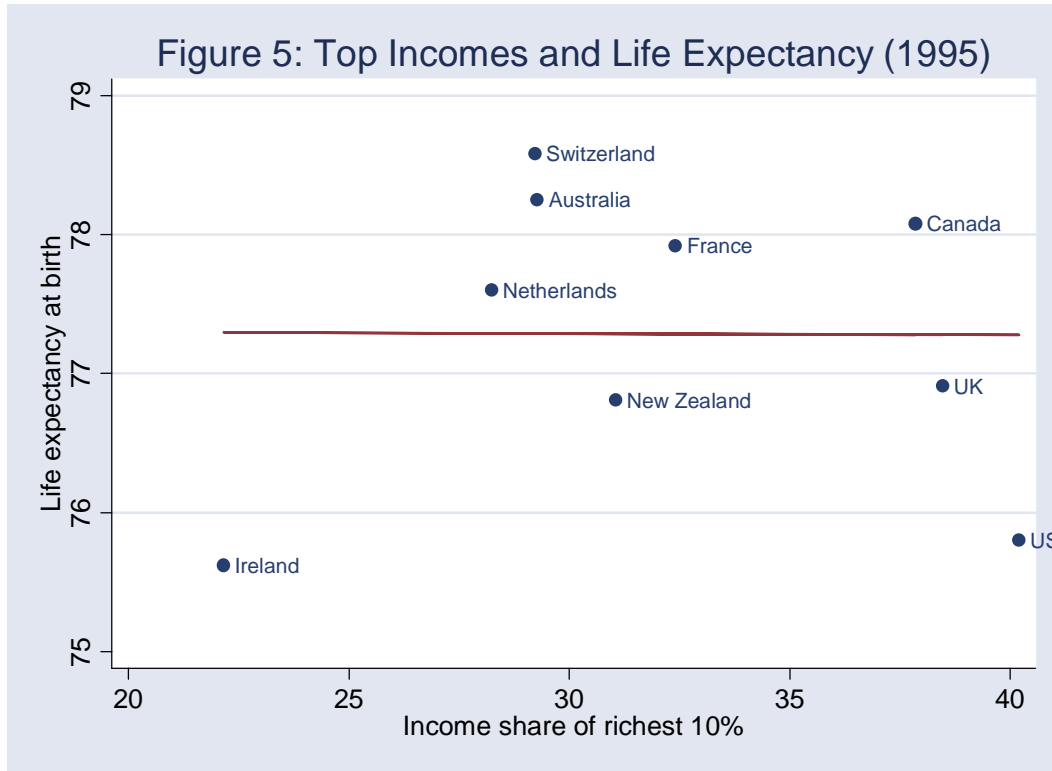


Table 1: Stationarity Tests

$$\text{Panel A: Dickey Fuller Tests } \Delta Y_t = \gamma Y_{t-1} + \sum_{i=0}^{10} \beta_i \Delta Y_{t-i+1} + \varepsilon_t$$

Country	Transformed top 10% share	Log infant mortality	Life expectancy	N
Australia	-0.896	-2.711***	2.144	48
Canada	-0.987	-1.975**	1.567	48
France	0.500	-2.847***	1.664	68
Ireland	0.889	-1.851*	1.704	13
Netherlands	1.089	-3.091***	1.948	74
New Zealand	-0.106	-1.691*	1.732	49
Switzerland	0.966	-1.817*	1.147	52
UK	-0.804	-2.365**	2.978	59
US	-0.358	-2.331***	2.725	74

Panel B: Johansen Cointegration Tests on Top 10% Share and Health Measures (5 lags)

	Trace statistic for log infant mortality	Trace statistic for life expectancy	5% critical value	1% critical value
Australia	8.896	9.927	15.41	20.04
Canada	4.758	3.884	15.41	20.04
France	12.964	16.382**	15.41	20.04
Ireland	49.551***	39.878***	15.41	20.04
Netherlands	13.115	17.837**	15.41	20.04
New Zealand	4.275	9.941	15.41	20.04
Switzerland	13.998	24.620***	15.41	20.04
UK	6.567	6.232	15.41	20.04
US	5.711	2.830	15.41	20.04

Notes:

1. In Panel A, Y is the variable to be tested. In the case of the top 10% share, $Y = \log S/(1-S)$, where S is the share of the top 10% group.
2. ***, ** and * denote rejection of the null hypothesis of a unit root at the 1%, 5% and 10% levels respectively (in all cases, the 10% critical value is -1.610).
3. All specifications include 11 lags of the differenced variable, chosen according to the Schwert criterion.
4. Since the New Zealand and UK series both have breaks, the tests are for New Zealand data after 1953, and UK data before 1990.

Table 2: Top 10% Share and Health: Levels Specification With and Without Fixed Effects

	[1]	[2]	[3]	[4]	[5]	[6]	[7]	[8]
Dependent variable:	LE	IM	LE	IM	LE	IM	LE	IM
Income share of richest 10%	-0.3452*** [0.0840]	0.0429*** [0.0085]	-0.3835*** [0.0460]	0.0318*** [0.0063]	-0.0090 [0.0732]	0.0138* [0.0073]	0.0508 [0.0718]	0.0041 [0.0068]
Real GDP per capita (\$1000s)	0.8527*** [0.1026]	-0.1193*** [0.0136]	0.8673*** [0.0782]	-0.1299*** [0.0110]	0.2070 [0.2130]	-0.0221 [0.0225]	0.2363 [0.1634]	-0.0267 [0.0167]
Country FE?	No	No	Yes	Yes	No	No	Yes	Yes
F-test			22***	12***			30,568***	130,000***
Year FE?	No	No	No	No	Yes	Yes	Yes	Yes
F-test					22,679***	94***	963***	149***
Observations	593	584	593	584	593	584	593	584
R-squared	0.75	0.86	0.84	0.93	0.91	0.95	0.95	0.98

Note:

1. Robust standard errors, clustered at the country level, in parentheses. *, ** and *** denote statistical significance at the 10%, 5% and 1% levels respectively.
2. Dependent variables: LE is average life expectancy at birth, IM is the log of the infant mortality rate (per 1000 live births).
3. F-test is a test for the joint significance of the country fixed effects or year fixed effects.
4. All specifications include a dummy variable for the change from joint to individual filing in New Zealand in 1953, and in the UK in 1990 (see Leigh 2004 for details).

Table 3: Top 10% Share and Health: Levels

	[1]	[2]	[3]	[4]	[5]	[6]	[7]	[8]
Dependent variable:	LE	IM	LE	IM	LE	IM	LE	IM
Income share of richest 10%	0.0508	0.0041	0.1915*	-0.0082	0.0566	-0.0065	0.0304	-0.0014
	[0.0718]	[0.0068]	[0.1012]	[0.0067]	[0.0850]	[0.0125]	[0.0675]	[0.0101]
Real GDP per capita (\$1000s)	0.2363	-0.0267	1.3704**	-0.1283**	0.2286	-0.0599**	0.2115	-0.0523*
	[0.1634]	[0.0167]	[0.5470]	[0.0470]	[0.1468]	[0.0245]	[0.1359]	[0.0253]
Real GDP per capita squared (\$1000s)			-0.0373**	0.0034**	-0.0081	0.0015**	-0.0078	0.0015**
			[0.0156]	[0.0014]	[0.0053]	[0.0006]	[0.0043]	[0.0005]
Average years of education							-0.2817	-0.0081
							[0.2432]	[0.0271]
Log real public health spending per capita							0.5781	-0.0729
							[0.3518]	[0.0880]
Log real private health spending per capita							0.6002	-0.1173
							[0.3540]	[0.0879]
Country & year FE?	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Observations	593	584	593	584	329	325	329	325
R-squared	0.95	0.98	0.96	0.98	0.98	0.97	0.98	0.98

Note:

1. Robust standard errors, clustered at the country level, in parentheses. *, ** and *** denote statistical significance at the 10%, 5% and 1% levels respectively.
2. Dependent variables: LE is average life expectancy at birth, IM is the log of the infant mortality rate (per 1000 live births).
3. Columns 5 and 6 are restricted to those country-years for which education and health spending variables are available.
4. All specifications include a dummy variable for the change from joint to individual filing in New Zealand in 1953, and in the UK in 1990 (see Leigh 2004 for details).

Table 4: Top 10% Share and Health: 1-year differences

	[1]	[2]	[3]	[4]	[5]	[6]	[7]	[8]
Dependent variable:	Δ LE	Δ IM	Δ LE	Δ IM	Δ LE	Δ IM	Δ LE	Δ IM
Δ Income share of richest 10%	0.0382 [0.0883]	-0.0033 [0.0046]	0.0585 [0.0762]	-0.0048 [0.0041]	-0.0077 [0.0190]	-0.0157** [0.0063]	-0.0103 [0.0207]	-0.0151* [0.0067]
Δ Real GDP per capita (\$1000s)	0.4414 [0.3263]	-0.0144 [0.0170]	1.5064* [0.7978]	-0.0931** [0.0395]	0.2814 [0.3005]	-0.0038 [0.0369]	0.3079 [0.3054]	-0.0071 [0.0410]
Δ Real GDP per capita squared (\$1000s)			-0.0490* [0.0249]	0.0036** [0.0013]	-0.0088 [0.0082]	0.001 [0.0011]	-0.0095 [0.0082]	0.0011 [0.0012]
Δ Average years of education							-0.1970** [0.0751]	0.0269 [0.0349]
Δ Log real public health spending per capita							0.1332 [0.0924]	-0.0238 [0.0333]
Δ Log real private health spending per capita							0.014 [0.0536]	-0.0192 [0.0184]
Country & year FE?	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Observations	579	571	579	571	319	315	319	315
R-squared	0.53	0.28	0.55	0.32	0.33	0.27	0.34	0.28

Note:

1. Robust standard errors, clustered at the country level, in parentheses. *, ** and *** denote statistical significance at the 10%, 5% and 1% levels respectively.
2. Dependent variables: LE is average life expectancy at birth, IM is the log of the infant mortality rate (per 1000 live births).
3. Columns 5 and 6 are restricted to those country-years for which education and health spending variables are available.
4. All specifications include a dummy variable for the change from joint to individual filing in New Zealand in 1953, and in the UK in 1990 (see Leigh 2004 for details).

Table 5: Top 10% Share and Health: 5-year differences

	[1]	[2]	[3]	[4]	[5]	[6]	[7]	[8]
Dependent variable:	Δ LE	Δ IM	Δ LE	Δ IM	Δ LE	Δ IM	Δ LE	Δ IM
Δ Income share of richest 10%	0.0427 [0.0998]	-0.0002 [0.0078]	0.090 [0.0689]	-0.0035 [0.0058]	0.0547 [0.0456]	-0.0123 [0.0096]	0.0496 [0.0492]	-0.0096 [0.0089]
Δ Real GDP per capita (\$1000s)	0.6254* [0.3010]	-0.0269 [0.0198]	1.8552*** [0.5520]	-0.1087*** [0.0279]	0.164 [0.1309]	0.0024 [0.0358]	0.2145 [0.1388]	0.0139 [0.0391]
Δ Real GDP per capita squared (\$1000s)			-0.0552*** [0.0161]	0.0037*** [0.0010]	-0.007 [0.0043]	0.0009 [0.0009]	-0.0087* [0.0042]	0.0006 [0.0010]
Δ Average years of education							-0.214 [0.1451]	0.0042 [0.0261]
Δ Log real public health spending per capita							0.2785 [0.2856]	-0.0826** [0.0293]
Δ Log real private health spending per capita							0.0649 [0.0977]	-0.0429** [0.0158]
Country & year FE?	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Observations	531	527	531	527	279	275	279	275
R-squared	0.58	0.45	0.65	0.54	0.5	0.49	0.53	0.51

Note:

1. Robust standard errors, clustered at the country level, in parentheses. *, ** and *** denote statistical significance at the 10%, 5% and 1% levels respectively.
2. Dependent variables: LE is average life expectancy at birth, IM is the log of the infant mortality rate (per 1000 live births).
3. Columns 5 and 6 are restricted to those country-years for which education and health spending variables are available.
4. All specifications include a dummy variable for the change from joint to individual filing in New Zealand in 1953, and in the UK in 1990 (see Leigh 2004 for details).

Table 6: Top 10% Share and Health: 10-year differences

	[1]	[2]	[3]	[4]	[5]	[6]	[7]	[8]
Dependent variable:	Δ LE	Δ IM	Δ LE	Δ IM	Δ LE	Δ IM	Δ LE	Δ IM
Δ Income share of richest 10%	0.0575 [0.0946]	-0.0023 [0.0076]	0.0894 [0.0621]	-0.0047 [0.0052]	0.0569 [0.0532]	-0.0172 [0.0127]	0.0338 [0.0391]	-0.0126 [0.0090]
Δ Real GDP per capita (\$1000s)	0.6224* [0.2802]	-0.0473* [0.0244]	1.5710*** [0.3878]	-0.1143*** [0.0265]	0.5528** [0.1915]	-0.0686 [0.0523]	0.4865* [0.2233]	-0.0436 [0.0584]
Δ Real GDP per capita squared (\$1000s)			-0.0454*** [0.0103]	0.0033*** [0.0009]	-0.0198** [0.0074]	0.0027** [0.0011]	-0.0186** [0.0081]	0.0021 [0.0014]
Δ Average years of education							-0.0735 [0.1384]	-0.0169 [0.0368]
Δ Log real public health spending per capita							0.4864 [0.3301]	-0.1017** [0.0383]
Δ Log real private health spending per capita							0.3117* [0.1614]	-0.0694** [0.0269]
Country & year FE?	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Observations	476	472	476	472	229	225	229	225
R-squared	0.63	0.52	0.69	0.59	0.65	0.56	0.68	0.60

Note:

1. Robust standard errors, clustered at the country level, in parentheses. *, ** and *** denote statistical significance at the 10%, 5% and 1% levels respectively.
2. Dependent variables: LE is average life expectancy at birth, IM is the log of the infant mortality rate (per 1000 live births).
3. Columns 5 and 6 are restricted to those country-years for which education and health spending variables are available.
4. All specifications include a dummy variable for the change from joint to individual filing in New Zealand in 1953, and in the UK in 1990 (see Leigh 2004 for details).

Table 7: Top 10% Share and Homicide/Suicide

Dependent variable:	[1] HOM	[2] SUI	[3] ΔHOM (1 yr diff)	[4] ΔSUI (1 yr diff)	[5] ΔHOM (5 yr diff)	[6] ΔSUI (5 yr diff)	[7] ΔHOM (10 yr diff)	[8] ΔSUI (10 yr diff)
Income share of richest 10%	-0.0385 [0.0271]	-0.0142 [0.0127]	-0.0291* [0.0132]	-0.0021 [0.0056]	-0.0231 [0.0211]	0.0035 [0.0087]	-0.0117 [0.0281]	0.0123 [0.0135]
Real GDP per capita (\$1000s)	-0.0522 [0.0775]	0.0981 [0.0636]	0.0206 [0.0980]	-0.0597 [0.0340]	-0.0059 [0.1284]	0.0084 [0.0636]	-0.0727 [0.1580]	0.0648 [0.1115]
Real GDP per capita squared (\$1000s)	0.0012 [0.0017]	-0.0022 [0.0017]	-0.0013 [0.0027]	0.0013 [0.0009]	0.0006 [0.0036]	-0.0003 [0.0018]	0.0018 [0.0045]	-0.0016 [0.0028]
Country & year FE?	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Observations	420	415	409	404	367	362	318	313
R-squared	0.92	0.87	0.19	0.16	0.23	0.3	0.3	0.41

Note:

1. Robust standard errors, clustered at the country level, in parentheses. *, ** and *** denote statistical significance at the 10%, 5% and 1% levels respectively.
2. Dependent variables: HOM is the log of the homicide rate per 100,000 people; SUI is the log of the suicide rate per 100,000 people
3. For differenced specifications (columns 3-8), GDP is also differenced over the same interval.
4. All specifications include a dummy variable for the change from joint to individual filing in New Zealand in 1953, and in the UK in 1990 (see Leigh 2004 for details).

Appendix Table 1: Summary Statistics

	Mean	SD	N
Income share of richest 10%	33.423	5.934	593
Average life expectancy at birth (years)	70.334	6.442	593
Log infant mortality rate (per 1000 live births)	2.959	0.819	584
Log homicide rate (per 100,000 people)	0.282	0.796	420
Log suicide rate (per 100,000 people)	2.457	0.350	415
Real GDP per capita (\$1000s)	11.442	5.402	593
Average years of education of adults aged 15+	9.196	1.718	348
Log real public health spending per capita (converted to 2003 \$US at PPP)	6.814	0.553	337
Log real private health spending per capita (converted to 2003 \$US at PPP)	6.092	0.823	340