SOCIAL AND POLITICAL DETERMINANTS OF POPULATION HEALTH IN EUROPE

Yannan Hu

ISBN: 978-94-6169-906-0

Social and political determinants of population health in Europe
Doctoral thesis, Erasmus University Rotterdam

Design, lay-out and print: Optima Grafische Communicatie, Rotterdam, The Netherlands

The studies presented in this thesis were supported by the project DEMETRIQ (Developing Methodologies to Reduce Inequalities in the Determinants of Health, grant agreement number 278511), which received funding from the European Union under the FP7 Health programme.
This thesis was printed with financial support of the Department of Public Health, Erasmus Medical Centre and of the Erasmus University Rotterdam.

SOCIAL AND POLITICAL DETERMINANTS OF POPULATION HEALTH IN EUROPE

Sociale en politieke determinanten van de volksgezondheid in Europa

Proefschrift

ter verkrijging van de graad van doctor aan
de Erasmus Universiteit Rotterdam op
gezag van de rector magnificus
Prof.dr. H.A.P. Pols
en volgens besluit van het College voor Promoties.

De openbare verdediging zal plaatsvinden
op woensdag 14 september 2016 om 13:30 uur

door

Yannan Hu

geboren te Hunan, China

Erasmus University Rotterdam

PROMOTIECOMMISSIE

Promotor	Prof.dr. J.P. Mackenbach
Overige leden	Dr. M. Avendaño Pabon
	Prof.dr. E.K.A. van Doorslaer
	Prof.dr. K. Stronks
Copromotor	Dr. F.J. van Lenthe

CONTENTS

Chapter 1

General introduction

1. POPULATION HEALTH IN EUROPE: VARIATIONS, TRENDS, AND DETERMINANTS

1.1 International variations in population health

1.1.1 Patterns and trends

European countries vary substantially in many ways, including their history, climate, political systems, welfare regimes, culture, and health [1]. With regard to the latter, some European countries appear to have the best levels of population health world-wide. For example, in some Mediterranean (Spain, Italy, Cyprus) and continental (Switzerland, Liechtenstein) countries as well as in Iceland and Sweden, life expectancy at birth in 2013 for men was above 80 years [2]. Evidence suggests that this is related to a long and sustained period of improvement in the lives people were able to lead [3]. However, not all countries experienced a sustained health improvement, and countries vary greatly in what they have achieved. For example, in countries in the Baltic region (Latvia and Lithuania) male life expectancy at birth in 2013 was below 70 years; in Central-Eastern European (Bulgaria, Poland, Slovakia) and Balkan countries (Macedonia, Serbia) male life expectancy at birth was below 74 years [2]. While life expectancy was generally higher among women, a similar gap was observed: life expectancy at birth ranged from just below 78 years in Macedonia and Serbia to above 85 years in Spain, France and Italy [2].

Diversity also exists in trends in population health across regions and countries in Europe. Most countries in the North of Europe have had very high levels of life expectancy throughout the 20th century [4]. Trends in life expectancy in Great Britain and Ireland were largely similar to those in the Nordic countries, but levels were consistently lower by a few years [4]. Most continental countries experienced a steep rise in life expectancy, with the Netherlands and Switzerland as two most favourable examples [4]. Overall, Western European countries experienced sustained improvements in life expectancy after World War II [5], with a typical increase of between 6 and 8 years since 1970 [6]. However, the picture for countries from Central and Eastern Europe and the former Soviet Union is very different. Levels of life expectancy in Central and Eastern Europe were similar to those in Western Europe around 1960, but stagnated dramatically and even declined in several countries afterwards [4, 7]. A few years after the collapse of the Berlin wall in 1989, life expectancy in Central and Eastern Europe started to steadily increase again and continued to do so at a rate that is rather similar to the increase in Western European countries [6]. In the countries of the former Soviet Union, life expectancy declined among men after 1960 and stagnated among women, with a short-lived improvement in the late 1980s [4, 8]. Soon afterwards there was a substantial decline, induced by the collapse of the Soviet Union in 1991 [6]. This was particularly dramatic in Russia: between 1990 and 1994 male life expectancy fell by 6 years to a low of 57 years [6, 8].

The between-country variations in political and social developments and population health in Europe offer good opportunities for research on determinants of population health [3, 4].

By linking political and economic factors to population health, we can gain deeper insights into why some countries are more successful than others in improving population health [9].

1.1.2 Potential determinants of European variations in population health

Several factors have been suggested as potential determinants of between-country differences in population health. Economic development, typically measured by gross domestic product (GDP), is perhaps the most straightforward one. Indeed, positive cross-sectional relationships between GDP and life expectancy have been found in European comparative studies [1, 10]. Increases in national income contributed 50% to 75% to the increase in life expectancy between 1960 and 1990, for a large part due to associated declines in mortality from cardiovascular and infectious diseases in this period [10]. Besides economic development, political traditions have been found to be consistently related to population health [9]. Several studies found that more years of social-democratic government are associated with better population health [11, 12], presumably because social-democratic governments are more committed to social policies conducive to health (e.g. preventive health policies related to tobacco and alcohol control [13], and labour market policies aiming at reducing social inequalities [12, 14]). Other national characteristics, such as welfare state characteristics and government effectiveness, have also been found to be potential explanatory factors of national variations in population health [1, 9]. For example, one study found that type of welfare state regime appeared to account for approximately half of the national-level variations of health inequalities in self-perceived health between European countries [15], although existing study results have been inconsistent [16]. Moreover, variations in culture may play a role in explaining differences between European countries in population health. Using three different sets of cultural measures, one study showed that most of cultural scales, especially the Inglehart's "self-expression" scale (i.e. adherence to self-expression instead of survival values) were related to population health outcomes and health behaviours in Europe, and variations in cultural values appeared to account for some of the striking variations in health behaviours between neighbouring countries [17]. For example, while having a similar level of national income as the Netherlands, Belgium appeared to have higher rates of antibiotics consumption. This could be partly explained by the fact that Belgians feel more threatened by unknown situations (i.e. a higher score of "uncertainty avoidance") and can accept more unequal power distribution (i.e. a higher score of "power distance"), since a 'watchful waiting' approach is less acceptable for countries with more 'uncertainty avoidance' and open communications between patients and doctors are less possible in countries with higher 'power distance' [17].

Democracy and income inequality are perhaps two of the most frequently studied but controversial determinants of population health. Population health is one important area where democracy, to the extent that it does promote the public good, can be expected to make a difference [18–20]. With some exceptions [9, 21, 22], many studies have found democratic governments to be associated with higher life expectancy [9, 19, 20, 23, 24] or lower infant mortality [9, 25, 26], even after controlling for some confounding. However, many studies on the relationship between democracy and population health have various shortcomings in their analytic methods, such as failing to examine the robustness against different model specifications and sample changes, or failing to control for country-specific fixed effects or global health trends [19, 22]. Also, studies with less aggregated health measures than life expectancy or infant mortality are lacking, and it is therefore yet unclear which causes of disease explained this. As many of these studies covered a wide range of countries with a similarly wide range of political regimes, including many low-income countries, it is also unclear whether the smaller variation in degrees of democracy observed in Europe today may actually contribute to the observed differences in health [1]. Clearly, this is important from an EU policy perspective.

The question whether income inequality harms population health also has not reached a unanimous answer. The first study published on this issue found a relationship between income inequality and mortality indicators, and interpreted this relation in terms of diminishing health returns with rising individual income [27]. The author suggested that decreasing income inequality by transferring income from the rich to the poor could improve average population health, since the improvement in health of the poor would be larger than the decline in health of the rich. Years later, Wilkinson [28–32] postulated the hypothesis that income inequality was not simply a summary of the balance of income between the rich and poor, but a health risk in its own right [33]. Ever since, numerous studies examined the link between income distribution and population health, but no final agreement has been reached [34–36]. While many cross-national studies found an association between larger income inequalities and poorer population health [27, 28, 37], establishing whether this statistical association reflects a causal effect of income inequality on health has proven to be difficult [38]. Again, methodological improvements such as excluding confounding by other country-level characteristics are necessary. Moreover, limited comparability of the income inequality measures between countries and over time appeared a problem in many studies [39–41], and only few studies investigated disease-specific outcomes [41–43]. Studies specifically assessing the association between income inequality and mortality in a European context, which would be important for policy makers in Europe, are also limited in number.

1.2 International variations in socioeconomic inequalities in health

1.2.1 Patterns and trends

Socioeconomic inequalities in health, measured either in absolute terms, for example as the difference in morbidity and mortality rates between socioeconomic groups (“absolute inequalities”), or in relative terms (for example as the ratio of morbidity and mortality rates among lower as compared to higher socioeconomic groups (“relative inequalities”), are a major challenge to public health in Europe. Socioeconomic inequalities in health have been found in all European countries with available data, and usually amount to between 5 and 10 years difference in life expectancy, and between 10 and 20 years difference in disability-free life expectancy [44]. Historical evidence suggests that socioeconomic inequalities in health are not a recent phenomenon [45]. However, it was only during the nineteenth century when relevant health data became available that socioeconomic inequalities in health were 'discovered' [45]. The recent active interest in this field in Europe can be linked to the publication of the Black Report in England in 1980 [46]. Since then, many studies demonstrated the existence of substantial inequalities in health in different countries [45, 47].

While socioeconomic inequalities in mortality and morbidity are found in all European countries with available data, they vary in size between countries. Socioeconomic inequalities in mortality are generally smaller in Southern European countries and larger in most countries in the Eastern and Baltic regions [47]. For socioeconomic inequalities in cause-specific mortality, three different 'regimes' in Europe were observed: a North-western regime with large inequalities in mortality from cardiovascular disease (men and women) and cancer (men only); a Southern regime with small inequalities in mortality from cardiovascular disease (men and women) and large inequalities in mortality from cancer (men only); and an Eastern regime with huge inequalities in mortality from cardiovascular disease, cancer and injuries (men and women) [45, 47, 48]. Socioeconomic inequalities have also been found for self-reported diseases and disabilities, with higher prevalence rates of less-than-good self-assessed health among lower socioeconomic groups in all countries with available data [47, 49–51]. However, no clear patterns have emerged in the magnitude of socioeconomic inequalities in self-assessed health between European countries [45, 47, 49, 50, 52].

Trends over time in socioeconomic inequalities also vary between countries. During the 1980s and 1990s, inequalities in mortality (especially relative inequalities) widened in many European countries [53, 54]. In more recent periods (i.e. during 1990s and 2000s), relative inequalities in mortality increased in most populations in the North, West and East of Europe, but not in the South of Europe [55]. This increase was mostly due to smaller proportional reductions in mortality among the lower than the higher socioeconomic groups. In the case of Lithuania and Estonia, however, mortality rose among the lower and declined among the higher socioeconomic groups [55]. In absolute terms, reductions in premature mortality were larger among the lower socioeconomic groups in many countries, mainly due to larger absolute reductions in mortality from cardiovascular disease and cancer, and as a result absolute inequalities in mortality often declined, particularly among men [55]. Compared to studies of trends in mortality inequalities, studies of trends in inequalities in self-assessed health are less common. One study covering 10 European countries between the 1980s and 1990s showed a high degree of stability of inequalities in self-assessed health [56]. A comprehensive overview of more recent trends based on a larger set of European countries is still lacking.

As numerous studies have now described variations in (trends of) socioeconomic inequalities in health, the emphasis of research in this area has gradually shifted from description to explanation. This shift not only stemmed from satisfying scientific curiosities, but also from the need to find entry-points for policies and interventions to reduce health inequalities [45, 57] and to evaluate them. Although some studies attempted to get insights of entry-points for reducing health inequalities by theoretically modifying the population distribution of risk factors [58, 59], variations in the magnitude and trend in socioeconomic inequalities in health in Europe have not been extensively used for research on determinants of health inequalities or the evaluation of strategies aimed at tackling health inequalities.

1.2.2 Potential determinants of European variations in inequalities in health

Ever since the Black report, a social causation mechanism has been seen a main explanation of socioeconomic inequalities in health. According to this mechanism, factors resulting from social stratification and causally related to health contribute to socioeconomic inequalities in health. Existing conceptual models, including the model adopted by the Commission on Social Determinants of Health, mention material, psychological and behavioural factors as well as access to health care as such mediating variables [60]. Apart from country-level characteristics, it is such factors that may differ between countries. Of these, cross-national differences in inequalities in health behaviours and access to health care have probably been investigated most in order to understand cross-national variation in health inequalities. For example, smoking plays an important role in generating health inequalities in Europe [59], as the prevalence of smoking differs strongly between socioeconomic groups in many countries and because smoking is an important contributory factor to major chronic diseases and premature mortality [45, 61, 62]. Similarly, other behavioural factors are also likely to play a role. For example, the distribution of dietary behaviour by socioeconomic status, e.g. consumption of fresh vegetable, meat, fats and oils, which is socially patterned in different European countries could also explain the variations in health inequalities between countries [45].

1.2.3. The reduction of socioeconomic inequalities in Europe

Against the background of an increasing knowledge on the determinants of socioeconomic inequalities in health, an important question is how such inequalities can be tackled. Reducing health inequalities is a major challenge for many European countries, not only from an ethical point of view [60], but also because it offers great potential for population health gains [63]. Several European countries have taken steps to develop strategies aimed at tackling health inequalities.

A few countries, such as the Netherlands and the Nordic countries, had national research programmes as well as high-level advisory committees that issued comprehensive policy advice on how to reduce socioeconomic inequalities in health [45, 64–66]. For example, the national "Program Committee on Socioeconomic Inequalities in Health" in the Netherlands issued a set of 26 specific recommendations in 2001, based on studies covering 12 different intervention fields [67, 68]. However, due to changes in the Dutch government, the recommendations have not been followed [45].

The English government set a unique example of implementing a well-coordinated national program tacking health inequalities, the explicit and sustained commitment of which was regarded as both historically and internationally unique [65, 69]. This contained a number of comprehensive and coordinated policies, which were clearly documented and monitored in a series of reports [70–77]. The high level of government commitment to reducing health inequalities was matched by an equally remarkable commitment to critically review, revise and then re-review its policies [78].

Cross-country comparative studies are of major importance in the evaluation of such strategies, for example as they allow the inclusion of countries without such strategies. This approach can also be used to evaluate more specific policies, which have been implemented without a focus on tackling health inequalities while in fact they may have an equity impact on health, e.g. alcohol policies, education policies and poverty reduction policies. Among them, tobacco control policies are perhaps the most frequently studied. However, it is still uncertain whether tobacco control policies have contributed to a narrowing or widening of socioeconomic inequalities in smoking, especially in European countries during the past two decades. The most consistent evidence from systematic reviews is that higher prices for cigarettes had a disproportionately greater impact on the most disadvantaged smokers and as such contributed to a reduction in inequalities in smoking [79, 80]. The equity impacts of many other tobacco control policies are less consistent; in fact, voluntary, regional and partial smoke free policies might have even increased inequalities in smoking [80]. Again, methodological challenges may have caused the lack of evaluations of the equity impact of such policies.

In more general terms, natural policy experiments (NPE), typically defined as "policies that are not under the control of the researchers, but which are amenable to research using the variation in exposure that they generate to analyze their impact" hold great promise [81–84]. Whereas the evaluation of natural policy experiments for population health becomes increasingly popular, the number of studies evaluating policy effects on health inequalities is still very small. Europe diversity offers a good but currently limited explored setting for this purpose.

2. THIS THESIS

2.1 Research questions

This thesis aims to contribute to the understanding of the potential relationship between social and political factors and population health and health inequalities from an international, European perspective; it also aims to provide potential methods to improve the evaluation of the impacts of social and political factors on health and health inequalities in Europe.

Specifically, we seek to answer the following study questions:

1) Are international variations in population health related to social and political characteristics of European countries, particularly to levels of democracy and income inequality?
2) Are international variations in health inequalities related to policies implemented in European countries, particularly to national policies to tackle socioeconomic inequalities in health and tobacco control policies?

2.2 Methods

As described above, methodological challenges may be at least partly responsible for the scarcity of studies on the social and political determinants of (socioeconomic inequalities in) population health. In fact, there seems to an inverse evidence law: the availability of evidence tends to vary inversely with the potential impact of the intervention [85]. There is a concentration of evidence on the effect of small-scale projects aimed at individual behavioural change, and a dearth of evidence on major policies applied across areas and countries, even though the latter could potentially have a greater population impact.

For the quantitative evaluation of such interventions, two closely related issues need attention: the degree to which groups can be assigned randomly to an intervention or control group, and the degree to which remaining confounding between intervention and control groups can be eliminated. To evaluate the effects of nation-wide policies, for example, countries are needed as control groups; clearly, this cannot be done through a process of randomization. Within countries, policies will be implemented where the need to do so is largest, or where the conditions to be successful are most optimal. As a result, those in the intervention and control groups may differ in many factors, including factors directly related to the assignment of the policy. Analytical methods which deal cleverly with exposure assignment and control for such confounding are needed in order to make reliable causal inferences.

Econometric techniques are increasingly advocated in health research, because of their ability to deal with confounding variables, including potential unobservable confounders. An intriguing question then is whether they can be applied in the evaluation of the impacts of policies on socioeconomic inequalities in health. If so, this would allow extending the evidence of policies aimed at reducing the inequalities substantially.

In this thesis, we use several econometric techniques which have been recognized as potentially useful techniques in public health [86]. By constructing counterfactual outcomes for the exposed units had they not been exposed to the intervention [87], these econometric techniques, such as fixed-effects models and difference-in-differences analyses, try to exclude confounding and improve causal inference from the relationships found. These techniques have not yet been much used in public health, and certainly not for the evaluation of policy effect on health inequalities.

2.3 Structure of this thesis

This thesis is divided in eight chapters. Chapter 1 provides a general introduction into the topic of this thesis. It describes the aims and specific research questions addressed in this thesis and it introduces the data and methods used in this thesis. Further to the general introduction, this thesis is divided into two parts.

The first part includes two chapters represented by chapter 2 and 3, which focus on different aspects of social and political characteristics of countries and their associations with the international variations in population health. In these two chapters, we link changes in democracy and income inequality respectively to improvements in population health as indicated by life expectancy and cause-specific mortality rates in multiple European countries over decades.

In the second part consisting of four chapters, we focus on social and political determinants of socioeconomic inequalities in Europe. In particular, in chapter 4 we estimate the overall trends in socioeconomic inequalities in self-assessed health in seventeen European countries between 1990 and 2010, with special emphases on the comparison between trends in inequalities in self-assessed health and those in mortality, and the correspondence between trends in inequalities in self-assessed health and national policies to tackle inequalities within countries. In chapter 5, we review the analytical methods for the evaluation of natural policy experiments and explore whether and how these methods can be used to evaluate policy effect on health inequalities. This chapter serves as an illustration of the method applied in chapters 6 and 7. In chapter 6, we evaluate the effectiveness of an ambitious programme tackling inequalities in health pursed by the English government, by comparing the changes in trends in health inequalities observed in England after the implementation of its programme to those in other countries without such a programme. Chapter 7 investigates the impact of price and non-price related tobacco control efforts on smoking by socioeconomic group in nine European countries between 1990 and 2007.

This thesis ends with a general discussion (chapter 8) of the findings. We conclude this thesis with summary answers, more specifically address possible methodological limitations and results' implications for public health policy.

REFERENCES

1. Mackenbach, J.P. and M. McKee, *Successes and failures of health policy in Europe: four decades of divergent trends and converging challenges*. 2013, Buckingham: Open Univeristy Press.
2. Eurostat. http://ec.europa.eu/eurostat/statistics-explained/images/d/d3/Life_expectancy_at_birth%2C_1980%E2%80%932013_%28years%29_YB15.png (accessed in January 2016).
3. Marmot, M., et al., *WHO European review of social determinants of health and the health divide.* Lancet, 2012. 380(9846): p. 1011–29.
4. Mackenbach, J.P., *Political conditions and life expectancy in Europe, 1900–2008.* Soc Sci Med, 2013. 82: p. 134–46.
5. Mackenbach, J.P., M. Karanikolos, and M. McKee, *The unequal health of Europeans: successes and failures of policies.* Lancet, 2013. 381(9872): p. 1125–34.
6. Leon, D.A., *Trends in European life expectancy: a salutary view.* Int J Epidemiol, 2011. 40(2): p. 271–7.
7. Zatonski, W., *Closing the health gap in the European Union*. 2008, Maria Sklodowska-Curie Memorial Cancer Center and Institute of Oncology http://www.goinginternational.eu/pdfs/fachartikel/Zatonski_Hauptartikel_KK11.pdf (accessed in January 2016): Warsaw.
8. Leon, D.A., et al., *Huge variation in Russian mortality rates 1984–94: artefact, alcohol, or what?* Lancet, 1997. 350(9075): p. 383–8.
9. Muntaner, C., et al., *Politics, welfare regimes, and population health: controversies and evidence.* Sociol Health Illn, 2011. 33(6): p. 946–64.
10. Mackenbach, J.P. and C.W. Looman, *Life expectancy and national income in Europe, 1900–2008: an update of Preston's analysis.* Int J Epidemiol, 2013. 42(4): p. 1100–10.
11. Mackenbach, J.P., *Political determinants of health.* Eur J Public Health, 2014. 24(1): p. 2.
12. Navarro, V., et al., *Politics and health outcomes.* Lancet, 2006. 368(9540): p. 1033–7.
13. Mackenbach, J.P. and M. McKee, *Social-democratic government and health policy in Europe: a quantitative analysis.* Int J Health Serv, 2013. 43(3): p. 389–413.
14. Borrell, C., et al., *Politics and health.* J Epidemiol Community Health, 2007. 61(8): p. 658–9.
15. Eikemo, T.A., et al., *Welfare state regimes and differences in self-perceived health in Europe: a multilevel analysis.* Soc Sci Med, 2008. 66(11): p. 2281–95.
16. Beckfield, J. and N. Krieger, *Epi + demos + cracy: linking political systems and priorities to the magnitude of health inequities--evidence, gaps, and a research agenda.* Epidemicol Rev, 2009. 31: p. 152–77.
17. Mackenbach, J.P., *Cultural values and population health: a quantitative analysis of variations in cultural values, health behaviours and health outcomes among 42 European countries.* Health Place, 2014. 28: p. 116–32.
18. Besley, T. and M. Kudamatsu, *Health and democracy.* American Economic Review, 2006. 96(2): p. 313–318.
19. Klomp, J. and J. de Haan, *Is the political system really related to health?* Social Science & Medicine, 2009. 69(1): p. 36–46.
20. Wigley, S. and A. Akkoyunlu-Wigley, *The impact of regime type on health: does redistribution explain everything?* World Politics, 2011. 63(4): p. 647-+.
21. Houweling, T.A., et al., *Determinants of under-5 mortality among the poor and the rich: a cross-national analysis of 43 developing countries.* Int J Epidemiol, 2005. 34(6): p. 1257–65.
22. Ross, M. *Does democracy reduce infant mortality?* 2006.
23. Safaei, J., *Is democracy good for health?* Int J Health Serv, 2006. 36(4): p. 767–86.
24. Alvarez-Dardet, C. and A. Franco-Giraldo, *Democratisation and health after the fall of the Wall.* J Epidemiol Community Health, 2006. 60(8): p. 669–71.

25. Gerring, J., S.C. Thacker, and R. Alfaro, *Democracy and Human Development*. Journal of Politics, 2012. 74(1): p. 1–17.
26. Navia, P. and T.D. Zweifel, *Democracy, dictatorship, and infant mortality revisited*. Journal of Democracy, 2003. 14(3): p. 90–103.
27. Rodgers, G.B., *Income and Inequality as Determinants of Mortality - International Cross-Section Analysis*. Population Studies-a Journal of Demography, 1979. 33(2): p. 343–351.
28. Wilkinson, R.G., *For Debate - Income-Distribution and Life Expectancy*. British Medical Journal, 1992. 304(6820): p. 165–168.
29. Wilkinson, R.G., *Mistaken Criticisms Ignore Overwhelming Evidence - Commentary*. British Medical Journal, 1995. 311(7015): p. 1285–1287.
30. Wilkinson, R.G., *Unhealthy societies: the affliction of inequality*. 1996, London: Routledge.
31. Wilkinson, R.G., *Commentary: income inequality summarises the health burden of individual relative deprivation*. BMJ, 1997. 314(7096): p. 1727–8.
32. Wilkinson, R.G., *Socioeconomic determinants of health. Health inequalities: relative or absolute material standards?* BMJ, 1997. 314(7080): p. 591–5.
33. Deaton, A., *Commentary: The convoluted story of international studies of inequality and health*. International Journal of Epidemiology, 2002. 31(3): p. 546–549.
34. Lynch, J., et al., *Is income inequality a determinant of population health? Part 1. A systematic review*. Milbank Quarterly, 2004. 82(1): p. 5–99.
35. Wilkinson, R.G. and K.E. Pickett, *Income inequality and population health: A review and explanation of the evidence*. Social Science & Medicine, 2006. 62(7): p. 1768–1784.
36. Mackenbach, J.P., *Income inequality and population health - Evidence favouring a negative correlation between income inequality and life expectancy has disappeared*. British Medical Journal, 2002. 324(7328): p. 1–2.
37. De Vogli, R., et al., *Has the relation between income inequality and life expectancy disappeared? Evidence from Italy and top industrialised countries*. J Epidemiol Community Health, 2005. 59(2): p. 158–62.
38. Avendano, M. and P. Hessel, *The income inequality hypothesis rejected?* Eur J Epidemiol, 2015. 30(8): p. 595–8.
39. Judge, K., J.A. Mulligan, and M. Benzeval, *Income inequality and population health*. Social Science & Medicine, 1998. 46(4–5): p. 567–579.
40. Beckfield, J., *Does income inequality harm health? New cross-national evidence*. Journal of Health and Social Behavior, 2004. 45(3): p. 231–248.
41. Leigh, A. and C. Jencks, *Inequality and mortality: Long-run evidence from a panel of countries*. Journal of Health Economics, 2007. 26(1): p. 1–24.
42. Mellor, J.M. and J. Milyo, *Reexamining the evidence of an ecological association between income inequality and health*. J Health Polit Policy Law, 2001. 26(3): p. 487–522.
43. Babones, S.J., *Income inequality and population health: correlation and causality*. Soc Sci Med, 2008. 66(7): p. 1614–26.
44. Mackenbach, J.P., *Health inequalities: Europe in profile*. 2006, London: An independent, expert report commissioned by the UK. http://ec.europa.eu/health/ph_determinants/socio_economics/documents/ev_060302_rd06_en.pdf (accessed in January 2016).
45. Mackenbach, J.P., *Socialeconomic inequalities in health in high-income countries: the facts and the options*. In: Detels, R. Gulliford, M., Karim, Q.A., Tan, C.C., eds. Oxford Textbook of Global Public Health. 2015, Oxford: Oxford University Press.
46. Department of Health and Social Services, *Inequalities in Health: Report of a Research Working Group (Black Report)*. 1980, Department of Health and Social Services: London.

47. Mackenbach, J.P., et al., *Socioeconomic inequalities in health in 22 European countries.* N Engl J Med, 2008. 358(23): p. 2468–81.
48. Huisman, M., et al., *Educational inequalities in cause-specific mortality in middle-aged and older men and women in eight western European populations.* Lancet, 2005. 365(9458): p. 493–500.
49. Cavelaars, A.E., et al., *Differences in self reported morbidity by educational level: a comparison of 11 western European countries.* J Epidemiol Community Health, 1998. 52(4): p. 219–27.
50. Cavelaars, A.E., et al., *Morbidity differences by occupational class among men in seven European countries: an application of the Erikson-Goldthorpe social class scheme.* Int J Epidemiol, 1998. 27(2): p. 222–30.
51. Dalstra, J.A., et al., *Socioeconomic differences in the prevalence of common chronic diseases: an overview of eight European countries.* Int J Epidemiol, 2005. 34(2): p. 316–26.
52. Eikemo, T.A., et al., *Class-related health inequalities are not larger in the East: a comparison of four European regions using the new European socioeconomic classification.* J Epidemiol Community Health, 2008. 62(12): p. 1072–8.
53. Mackenbach, J.P., et al., *Widening socioeconomic inequalities in mortality in six Western European countries.* Int J Epidemiol, 2003. 32(5): p. 830–7.
54. Shkolnikov, V.M., et al., *Increasing absolute mortality disparities by education in Finland, Norway and Sweden, 1971–2000.* Journal of Epidemiology and Community Health, 2012. 66(4): p. 372–378.
55. Mackenbach, J.P., et al., *Trends in inequalities in premature mortality: a study of 3.2 million deaths in 13 European countries.* J Epidemiol Community Health, 2015. 69(3): p. 207–17; discussion 205–6.
56. Kunst, A.E., et al., *Trends in socioeconomic inequalities in self-assessed health in 10 European countries.* Int J Epidemiol, 2005. 34(2): p. 295–305.
57. Mackenbach, J.P. and M. Bakker, *Reducing inequalities in health: a European perspective*. 2002, London: Routledge.
58. Eikemo, T.A., et al., *How can inequalities in mortality be reduced? A quantitative analysis of 6 risk factors in 21 European populations.* PLoS One, 2014. 9(11): p. e110952.
59. Kulik, M.C., et al., *Smoking and the potential for reduction of inequalities in mortality in Europe.* Eur J Epidemiol, 2013. 28(12): p. 959–71.
60. Commission on Social Determinants of Health, *Closing the gap in a generation. Health equity through the social determinants of health*. 2008: Geneva: World Health Organization.
61. Cavelaars, A.E., et al., *Educational differences in smoking: international comparison*. BMJ, 2000. 320(7242): p. 1102–7.
62. Huisman, M., A.E. Kunst, and J.P. Mackenbach, *Educational inequalities in smoking among men and women aged 16 years and older in 11 European countries.* Tob Control, 2005. 14(2): p. 106–13.
63. Mackenbach, J.P., W.J. Meerding, and K.A. E., *Economic implications of socio-economic inequalities in health in the European Union*. 2007, Luxembourg: services of the European Commission (OIL).
64. Whitehead, M., *Diffusion of ideas on social inequalities in health: A European perspective.* Milbank Quarterly, 1998. 76(3): p. 469-+.
65. Mackenbach, J.P., M.J. Bakker, and E.N.I. Pol, *Tackling socioeconomic inequalities in health: analysis of European experiences.* Lancet, 2003. 362(9393): p. 1409–1414.
66. Judge, K., et al., *Health inequalities: a challenge for Europe*. 2006: An independent, expert report commissioned by, and published under the auspices of, UK Presidency of the EU. http://ec.europa.eu/health/ph_determinants/socio_economics/documents/ev_060302_rd05_en.pdf (accessed 02 Feb 2015).
67. Mackenbach, J.P. and K. Stronks, *A strategy for tackling health inequalities in the Netherlands.* BMJ, 2002. 325(7371): p. 1029–32.

68. Mackenbach, J.P. and K. Stronks, *The development of a strategy for tackling health inequalities in the Netherlands*. Int J Equity Health, 2004. 3(1): p. 11.
69. Mackenbach, J.P., *Socio-economic inequalities in health in western Europe: from description to explanation to intervention. In Siegrist, J. & Marmot, M. (Eds.), Social inequalities in health: New evidence and policy implications*. 2006, New York: Oxford University Press.
70. Department of Health, *Reducing health inequalities: An action report*. 1999: London. http://www.swslim.org.uk/downloads/SL1264.pdf (accessed 02 Feb 2015).
71. Department of Health, *Independent inquiry into inequalities in health (the Acheson report)*. 1998: London. https://www.gov.uk/government/publications/independent-inquiry-into-inequalities-in-health-report (accessed 16 Oct 2014).
72. Department of Health, *Tackling health inequalities: Cross-cutting review*. 2002: London. http://www.dh.gov.uk/prod_consum_dh/groups/dh_digitalassets/@dh/@en/documents/digitalasset/dh_4068003.pdf (accessed 02 Feb 2015).
73. Department of Health, *Tackling health inequalities: A program for action*. 2003: London. http://www.dh.gov.uk/prod_consum_dh/groups/dh_digitalassets/@dh/@en/documents/digitalasset/dh_4019362.pdf (accessed 02 Feb 2015).
74. Department of Health, *Tackling health inequalities: 2007 status report of the program for action*. 2007: London. http://webarchive.nationalarchives.gov.uk/+/www.dh.gov.uk/en/Publicationsandstatistics/Publications/DH_083471 (accessed 02 Feb 2015).
75. Department of Health, *Tackling health inequalities: 10 years on*. 2009: London. http://webarchive.nationalarchives.gov.uk/20130107105354/http:/www.dh.gov.uk/prod_consum_dh/groups/dh_digitalassets/documents/digitalasset/dh_098934.pdf (accessed 02 Feb 2015).
76. Department of Health, *Fair society, healthy lives (the Marmot Review)*. 2010: London. http://www.instituteofhealthequity.org/projects/fair-society-healthy-lives-the-marmot-review (accessed 02 Feb 2015).
77. House of Commons, *Health inequalities*. 2009: London: Health Committee of the House of Commons. http://www.publications.parliament.uk/pa/cm200809/cmselect/cmhealth/286/28602.htm (accessed 02 Feb 2015).
78. Mackenbach, J.P., *Has the English strategy to reduce health inequalities failed?* Soc Sci Med, 2010. 71(7): p. 1249–53; discussion 1254–8.
79. Thomas, S., et al., *Population tobacco control interventions and their effects on social inequalities in smoking: systematic review*. Tob Control, 2008. 17(4): p. 230–7.
80. Brown, T., S. Platt, and A. Amos, *Equity impact of population-level interventions and policies to reduce smoking in adults: a systematic review*. Drug Alcohol Depend, 2014. 138: p. 7–16.
81. Meyer, B.D., *Natural and Quasi-Experiments in Economics*. Journal of Business & Economic Statistics, 1995. 13(2): p. 151–161.
82. MRC, *Using natural experiments to evaluate population health interventions: guidance for producers and users of evidence*. 2011: Medical Research Council [Online] Avilable from http://www.mrc.ac.uk/documents/pdf/natural-experiments-guidance/ (accessed 01 August 2014).
83. Craig, P., et al., *Using natural experiments to evaluate population health interventions: new Medical Research Council guidance*. Journal of Epidemiology and Community Health, 2012. 66(12): p. 1182–1186.
84. Petticrew, M., et al., *Natural experiments: an underused tool for public health?* Public Health, 2005. 119(9): p. 751–7.
85. Nutbeam, D., *How does evidence influence public health policy? Tackling health inequalities in England*. Health Promot J Aust, 2003. 14: p. 154–8.

86. Craig, P., et al., *Using natural experiments to evaluate population health interventions: new Medical Research Council guidance.* J Epidemiol Community Health, 2012. 66(12): p. 1182–6.

87. Rubin, D.B., *Estimating causal effects of treatments in randomized and nonrandomized studies.* Journal of Educational Psychology, 1974. 66(5): p. 688–701.

Part I

Social and political determinants of average population health

Chapter 2

Democratization, life expectancy and mortality in Europe, 1960–2008

Johan P. Mackenbach, Yannan Hu, Caspar W.N. Looman

Democratization and life expectancy in Europe, 1960–2008. Social Science & Medicine, 2013. 93: p. 166–175.

ABSTRACT

Over the past five decades, two successive waves of political reform have brought democracy to, first, Spain, Portugal and Greece, and, more recently, Central and Eastern European countries. We assessed whether democratization was associated with improvements in population health, as indicated by life expectancy and cause-specific mortality rates.

Data on life expectancy at birth, age-standardized total and cause-specific mortality rates, levels of democracy and potential time-variant confounding variables were collected from harmonized international databanks. In two pooled cross-sectional time-series analyses with country-fixed effects, life expectancy and cause-specific mortality were regressed on measures of current and cumulative democracy, controlling for confounders. A first analysis covered the 1960–1990 period, a second covered the 1987–2008 period.

In the 1960–1990 period, current democracy was more strongly associated with higher life expectancy than cumulative democracy. The positive effects of current democracy on total mortality were mediated mainly by lower mortality from heart disease, pneumonia, liver cirrhosis, and suicide. In the 1987–2008 period, however, current democracy was associated with lower, and cumulative democracy with higher life expectancy, particularly among men. The positive effects of cumulative democracy on total mortality were mediated mainly by lower mortality from circulatory diseases, cancer of the breast, and external causes. Current democracy was associated with higher mortality from motor vehicle accidents in both periods, and also with higher mortality from cancer and all external causes in the second.

Our results suggest that in Europe during these two periods democratization has had mixed effects. That short-term changes in levels of democracy had positive effects in the first but not in the second period is probably due to the fact that democratization in Central and Eastern Europe was part of a complete system change which caused major societal disruptions.

INTRODUCTION

Over the last 50 years the governing systems of many European countries have undergone profound changes, with a clear shift from authoritarian regimes to liberal democracies [1–3]. In 1960, at the peak of the Cold War, only about half of all European countries, mainly in the North and West, had liberal democracies, as defined by representative government operating through law, by regular, free and fair elections based on universal suffrage, and by respect for individual rights including freedom of expression and association [4]. Many other countries still had authoritarian regimes, in which rulers had limited popular accountability, the media were controlled, and political participation was limited [4]. At that time, several Mediterranean countries were still under right-wing autocratic regimes, and all countries in Central and Eastern Europe were under authoritarian regimes led by communist parties.

After two successive waves of political reform, most European countries now have liberal democracies [2, 3]. In the 1970s, Spain, Portugal and Greece shed off their military dictatorships, and around 1990 the communist regimes in most Central and Eastern European countries were all replaced by more democratic forms of government, ranging from fully liberal democracies to 'illiberal democracies' in which elections do take place but rulers exploit their position to prevent a level playing field, for example by interfering with the rule of law and with the media [4] (table 1).

Table 1 Levels of democracy, as indicated by the revised Polity2 index, in European countries, selected years in the period 1960–2008

	1960	1970	1980	1990	2000	2008
Nordic						
Finland	20	20	20	20	20	20
Sweden	20	20	20	20	20	20
Norway	20	20	20	20	20	20
Iceland	20	20	20	20	20	20
Denmark	20	20	20	20	20	20
Britain & Ireland						
United Kingdom	20	20	20	20	20	20
Ireland	20	20	20	20	20	20
Continental						
Netherlands	20	20	20	20	20	20
Belgium	20	20	20	20	20	18
Luxembourg	20	20	20	20	20	20
Germany (FRG)	20	20	20	20	20	20
Switzerland	20	20	20	20	20	20
Austria	20	20	20	20	20	20

Table 1 Levels of democracy, as indicated by the revised Polity2 index, in European countries, selected years in the period 1960–2008 (continued)

	1960	1970	1980	1990	2000	2008
Mediterranean						
France	15	18	18	19	19	19
Spain	3	3	19	20	20	20
Portugal	1	1	19	20	20	20
Italy	20	20	20	20	20	20
Malta			16	20	20	20
Greece	14	3	18	20	20	20
Cyprus	18	17	20	20	20	20
Western Balkans						
Yugoslavia	3	3	5	5		
Slovenia				5	20	20
Croatia				5	18	19
Bosnia-Hercegovina				5		
Serbia				5	17	18
Montenegro				5	17	18
TFYR Macedonia				5	16	19
Albania	1	1	1	11	15	19
Centre & East						
Germany (GDR)	1	1	1			
Poland	3	3	4	15	19	20
Czechoslovakia	3	3	3	18		
Czech Republic					20	18
Slovakia					19	20
Hungary	3	3	3	20	20	20
Romania	3	3	2	15	18	19
Bulgaria	3	3	3	18	18	19
(f) Soviet Union						
USSR	3	3	3	10		
Estonia				10	19	19
Latvia				10	18	18
Lithuania				10	20	20
Belarus				10	3	3
Ukraine				10	16	17
Republic of Moldova				10	17	18
Russian Federation				10	16	14
Georgia				10	15	16
Armenia				10	15	15
Azerbaijan				10	3	3

Notes: Scale from 0 (fully autocratic) to +20 (fully democratic). The original Polity2 index has been converted into an entirely positive scale. For further explanations, see Data and methods section. Germany (FRG) = Federal Republic of Germany before 1990, united Germany after 1990. Germany (GDR) = German Democratic Republic (part of united Germany in 1990 and later).

Whether democracy is more effective in promoting the public good than other forms of government is open to debate [5]. On the positive side, it has been argued that democratic governments can be expected to make decisions in accordance with voters' interests, and thus to be more actively engaged in promoting the public good than authoritarian governments. This advantage may be strengthened by greater public accountability, greater effectiveness in getting things done that require the active participation of the public, greater inclination towards redistributive policies, and greater ability to recruit competent and honest people [6–10]. On the other hand, citizens in democracies may not always vote in accordance with their own interests, democratically elected politicians may have difficulty looking beyond their election horizons, and democratic governments are vulnerable to manipulation and lobbying by corporate interests that stand in the way of promoting the public good [4, 11, 12].

Population health is one important area where democracy, to the extent that it does promote the public good, can be expected to make a difference. The past half century has seen an enormous growth of effective interventions to improve population health, ranging from tobacco control to road traffic safety, and from antibiotics to coronary artery bypass grafts, many of which have contributed importantly to advances in population health [13–16]. Implementation of these interventions has to a large extent been dependent on public policy, e.g. in the form of national health systems or universal health insurance schemes, and in the form of public health services and environmental protection programs [13, 16]. However, the hypothesis that democratization promotes the implementation of effective health interventions, and thereby reduces mortality from conditions amenable to these interventions, has never been directly tested.

Many studies have found democratic government to be associated with higher life expectancy [6–8, 17–23] or lower infant mortality [10, 19, 24–26], even after controlling for some confounding variables, but not all studies did [19, 27, 28]. Some authoritarian regimes have been very effective in improving population health: in a world-wide comparison Cuba and China stand out as autocratically governed countries with remarkably high life expectancies at birth [29, 30], whereas the European experience shows that fascist and communist countries had very rapidly rising life expectancies in the first decades after World War II [31, 32]. Perhaps more important than these counterexamples is the fact that most studies of the relation between democracy and population health did not apply sufficiently rigorous analytic methods [8]. Also, studies looking at less aggregate health measures than life expectancy or infant mortality are lacking, and so it is as yet unclear what the intervening mechanisms are.

We will therefore exploit the abrupt changes from authoritarian to democratic rule in Europe in the 1970s (Spain, Portugal and Greece) and around 1990 (Central and Eastern Europe) to assess whether democratization is associated with improvements in population health, as indicated by higher life expectancy and lower cause-specific mortality. In a previous descriptive study we have identified this as a potentially fruitful area for studying the impact of political conditions on population health [31]. By focusing on causes of death that have become amenable to intervention we hope to find clues for the mediating role of specific public policies (which cannot be studied directly because of a lack of comparable, quantitative trend data on policy implementation [13]). A simple graphical representation of the ideas underlying our analyses can be found in figure 1.

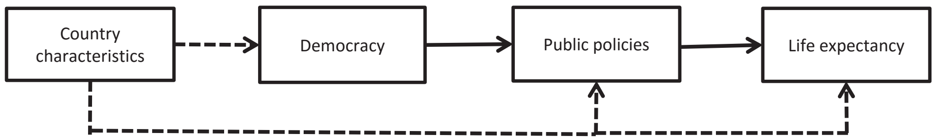


Figure 1 A simple graphical representation of relations between variables included in the analysis
Notes: The leading hypothesis, as set out in the Introduction section, is that democratic government is more effective in promoting the public good than autocratic government, because it is more active and more successful in implementing public policies which benefit the general population, including those that reduce mortality and raise life expectancy. Some of these beneficial effects may be immediate, while others require a slow build-up over years of democratic government. Democracy is associated with other country characteristics that also affect life expectancy, either directly or through public policy. Most of these other characteristics work in tandem with democracy, and without appropriate controls the effect of democracy on population health therefore risks to be overestimated. Some of these other country characteristics are measurable, e.g. national income and level of education, and some are unmeasurable, e.g. historical endowments and cultural values. The strategy for controlling these potential confounders is set out in the Data and methods section (with further details in the appendix).

DATA AND METHODS

Data

The main independent variable of interest is the Polity2 index as compiled by the worldwide and independent Polity IV project (http://www.systemicpeace.org/polity/polity4.htm). We extracted yearly data on this index from the Quality of Government dataset [33] (http://www.qog.pol.gu.se/data/). The Polity2 index indicates a country's position on a continuous scale from 'strongly democratic' (+10) to 'strongly autocratic' (-10), and is a summary score designed to facilitate time-series analyses. It is based on measurements of the competitiveness of political participation, the regulation of participation, the openness and competitiveness of executive recruitment, and constraints on the chief executive, and has become the standard measure of democracy in the literature [6, 34]. We created an entirely positive scale (from 0 to 20) by adding 10 to each country's original values of the Polity2 index, and labelled this the 'revised Polity2 index'.

As shown in table 1, all countries in the North and West of Europe had fully functional democracies throughout the study-period, but countries in the Mediterranean did not. In addition to Spain, Portugal and Greece, France, Malta and Cyprus also had periods in which the revised Polity2 index was lower than 20. In the Western Balkans and Central & Eastern

Europe, high levels on this index were only reached around the year 2000 and around 1990, respectively, whereas in the former Soviet Union developments towards democracy were highly variable.

In the analysis we controlled for a number of country characteristics that can be expected to be associated with both democracy and life expectancy, and therefore to act as confounders of the relation between democracy and life expectancy (figure 1). We controlled for time-invariant confounders by using country fixed effects models (see Analysis section), and we controlled for time-variant confounders by explicitly including them in our regression equations. The latter were national income, education, independence, armed conflict, and economic freedom. All these variables were obtained from harmonized international data sources. Further details about the rationale, measurement and data sources of the control variables can be found in the appendix.

We collected data on life expectancy at birth, by gender, for the period 1960–2008 from the Human Lifetable Database (www.lifetable.de), supplemented by the World Health Organization Health for All Database (http://data.euro.who.int/hfadb/). Age-standardized mortality data by cause of death were extracted from the International Mortality Data Base of the National Centre for Health Statistics of the Centres for Disease Control (http://www.cdc.gov/nchs/data/dvs/intmort95.pdf), which covers the 1960–1990 period, and from the World Health Organization Health for All Database (http://data.euro.who.int/hfadb/), which covers more recent years. ICD-code numbers are given in appendix table A1. Age-standardization was performed using the direct method and the European standard population.

Some descriptive statistics on dependent and independent variables are presented in appendix table A2.

Methods

We conducted two separate sets of analyses, one for 1960–1990 (capturing primarily the impact of democratization in Spain, Portugal and Greece, in addition to some smaller changes in France and Malta) and one for 1987–2008 (looking at the impact of democratization in Central and Eastern Europe). The first set of analyses was carried out on a dataset covering 29 countries and 31 years (n=899), the second on a dataset covering 43 countries and 22 years (n=946). Splitting up the analysis in this way not only allowed us to make optimal use of the two datasets for cause-specific mortality, but also to deal with changes in the political map of Europe. In the first set of analyses we included the Soviet Union and Yugoslavia, and in the second set we included the newly independent republics emerging from these two countries. Because democracy scores were not available for the newly independent republics before 1990 (Soviet Union) or 1991 (Yugoslavia), we assigned them the score for the Soviet Union and Yugoslavia as a whole for the three or four years preceding independence. Please note that all countries mentioned in table 1 have been included in one or both analyses, even if they did not undergo changes in democracy levels. Countries with stable levels of democracy are relevant for the analysis, not only as a reference point for those where democracy levels changed over time, but also by contributing information on the relation between confounders and life expectancy/mortality.

In the analysis we used both current democracy and cumulative years of democracy. Whereas the first variable captures short-term change, the second captures long-term and cumulative change in levels of democracy. The rationale for the latter is that several of the mechanisms that may underlie the beneficial effects of democracy (e.g. greater public accountability, greater ability to recruit competent and honest people) operate indirectly, by stimulating and facilitating policy changes that need to be built up over time (e.g. comprehensive tobacco control, road traffic safety, health care quality assurance), or require institutional change (e.g. new laws, creation of a modern public health workforce). These policy changes are therefore likely to depend on the gradual accumulation of years of democracy over several decades [7, 17]. On the other hand, some policy changes can be made rapidly (e.g. removing subsidies on unhealthy foods, better enforcement of existing traffic laws, importing modern life-saving drugs), and in case their mortality effects are immediate as well these may roughly coincide with changes in current democracy. For the 1960–1990 period we calculated cumulative years weighted with a country's revised Polity2 index since 1946 (the earliest year for which data are available), for the 1987–2008 period we calculated cumulative years weighted with a country's revised Polity2 index since 1970 (thereby capturing a similar number of years as we did for the first period).

In recognition of the pooled cross-sectional time-series character of our data we used analytic methods appropriate for panel data with clustered errors [35]. A country fixed effects model was used and in addition to the confounders mentioned above, we also controlled for time (in years, measured as a continuous variable). The full model can be written as:

$$healthoutcome_{ij} = \beta_0 + \beta_1 democracy_{ij} + \beta_2 gdp_{ij} + \beta_3 gdp_{ij}^2 + \beta_4 edu_{ij} + \beta_5 inde_{ij} + \beta_6 war_{ij} + \beta_7 free_{ij} + \beta_8 time_i + \beta_9 countrydummies_j + \varepsilon_{ij}$$

where health outcome is measured by life expectancy or age-adjusted mortality rates; democracy is represented by an index of current or cumulative democracy; *gdp*, gdp^2, *edu*, *inde*, *war*, *free* and *time* are control variables representing national income, the square of national income, average years of schooling, transition to independence, armed conflicts, economic freedom and time trend, respectively; country dummies are included to control for unobserved confounding; and subscripts *i* and *j* represent time and country. A model with country dummies is equivalent to a model in which each country's average values (e.g., for democracy) are subtracted from its observations, and we will therefore often refer to the results as indicating that changes in democracy within each country relate to changes in health outcomes. The country fixed effects model is designed to control for time-invariant confounders, such as cultural, social, historical and other conditions which can be considered to have remained relatively constant within the 1960–1990 and 1987–2008 periods.

To facilitate comparisons between the regression coefficients of current and cumulative democracy, we multiplied the former by 20. After this transformation, the coefficients for current democracy indicate the change in life expectancy at birth (in years), associated with a change from full autocracy (revised Polity2 index of 0) to full democracy (revised Polity2 index of 20). The coefficients for cumulative democracy indicate the change in life expectancy at birth (in years), associated with an additional year of full democracy (revised Polity2 index of 20).

Robust standard errors (taking into account correlations between values within each country) were calculated. All regression analyses were performed in Stata 12.0. We tested the robustness of our model by trying several alternative model specifications, and report these sensitivity analyses in appendix table A3.

RESULTS

In the 1960–1990 period, both current democracy and cumulative democracy are positively associated with life expectancy, but only current democracy has a statistically significant effect (table 2a). Among both men and women, a change from full autocracy to full democracy (i.e., 20 points on the revised Polity2 scale) is associated with almost two years of extra life expectancy in our country fixed effects model. Of the control variables, and apart from the country dummies (not shown in the table), only national income, transition to independence and armed conflict have independent associations with life expectancy. The positive association with armed conflict is counterintuitive – the only European country involved in large-scale conflicts in this period was the United Kingdom (Northern Ireland, Falklands war), and our results suggest that its involvement in these conflicts coincided with rapid life expectancy growth.

Figures 2a and 2b illustrate what happened to the gap in life expectancy between countries that democratized in this period, and the full democracies of Western Europe. Until the beginning of the 1970s, both Portugal and Spain had much lower current democracy scores than other Western European countries, and as a result they gradually built up a democratic deficit as shown in the growing gap between their cumulative democracy scores and those of other Western European countries. Despite this growing deficit, however, the gap in life expectancy with the rest of Western Europe continued to narrow. Shortly after the transition to full democracy, both Portugal's and Spain's life expectancy growth accelerated relative to the Western European average, leading to a more rapid narrowing of the gap in the case of Portugal, and the creation of a life expectancy advantage in the case of Spain.

In the 1987–2008 period, the patterns are remarkably different. Now, only cumulative democracy has a positive association with life expectancy, whereas current democracy has a negative association. Results for women, although pointing in the same direction as those

Table 2 Democracy and life expectancy: results of fixed effects regression models
a. 1960–1990.

	Men		Women	
	Current	Cumulative	Current	Cumulative
Current democracy (di)[a]	1.856***		1.836***	
	(0.622)		(0.592)	
Cumulative democracy (dic1)[b]		0.0690		0.0113
		(0.0523)		(0.0487)
National income (gdp)	0.580**	0.457	0.471*	0.481
	(0.243)	(0.312)	(0.239)	(0.281)
National income (gdp2)	-0.00468	-0.00641	-0.00670	-0.00872
	(0.00437)	(0.00643)	(0.00443)	(0.00602)
Education (edu)	0.317	0.361	0.396	0.467
	(0.405)	(0.483)	(0.318)	(0.406)
Independence (inde)[c]	-2.454***		-1.066**	
	(0.422)		(0.397)	
Armed conflict (war)	0.305*	0.286	-0.102	-0.0473
	(0.174)	(0.207)	(0.101)	(0.199)
Time (t)	-0.0142	-0.0174	0.0855	0.0900
	(0.0686)	(0.0716)	(0.0526)	(0.0592)
Constant	60.59***	61.61***	65.99***	66.69***
	(2.916)	(3.753)	(2.165)	(3.003)
Observations	738	711	738	711
R-squared	0.730	0.716	0.870	0.856
Number of countries	26	24	26	24

Notes: Country fixed effects models, estimated using Ordinary Least Squares regression with robust (clustered) standard errors. Current democracy: revised Polity2 score. Cumulative democracy: years since 1946 weighted by revised Polity2 score. Gross Domestic Product: in I$ per capita. Education: average years of education. Transition to independence: dummy variable (0 or 1). Armed conflict: dummy variable (0–3). Time: calendar-year. Robust standard errors in parentheses.

*** p<0.01, ** p<0.05, * p<0.1

a. The coefficients for current democracy have been multiplied by 20 to create comparability with the coefficients for cumulative democracy. They indicate the change in life expectancy at birth (in years), associated with a change from full autocracy (revised Polity2 index of 0) to full democracy (revised Polity2 index of 20).

b. The coefficients for cumulative democracy indicate the change in life expectancy at birth (in years), associated with an additional year of full democracy (revised Polity2 index of 20).

c. This variable has been omitted from the regressions for cumulative democracy because of multicollinearity.

b. 1987–2008.

	Men		Women	
	Current	Cumulative	Current	Cumulative
Current democracy (di)[a]	-1.522**		-0.0572	
	(0.674)		(0.412)	
Cumulative democracy (dic1)[b]		0.674**		0.248
		(0.295)		(0.195)
National income (gdp)	0.380**	0.334***	0.278***	0.231***
	(0.146)	(0.107)	(0.0756)	(0.0538)
National income (gdp2)	-0.00362**	-0.00317**	-0.00286***	-0.00234***
	(0.00168)	(0.00127)	(0.000957)	(0.000696)
Education (edu)	-0.223	-0.168	0.00861	0.0300
	(0.231)	(0.203)	(0.142)	(0.127)
Independence (inde)	0.572	0.451	0.173	0.199
	(0.343)	(0.343)	(0.146)	(0.134)
Armed conflict (war)	-0.567*	-0.401	-0.337	-0.282
	(0.323)	(0.271)	(0.200)	(0.172)
Economic liberalization (free)	-0.114	0.0351	0.0583	0.163
	(0.220)	(0.219)	(0.140)	(0.131)
Time (t)	0.182***	-0.481	0.122***	-0.115
	(0.0405)	(0.297)	(0.0253)	(0.193)
Constant	69.87***	60.60***	74.10***	71.08***
	(2.562)	(4.377)	(1.522)	(2.628)
Observations	692	692	692	692
R-squared	0.773	0.795	0.824	0.831
Number of countries	34	34	34	34

Notes: Country fixed effects models, estimated using Ordinary Least Squares regression with robust (clustered) standard errors. Current democracy: revised Polity2 score. Cumulative democracy: years since 1970 weighted by revised Polity2 score. Gross Domestic Product: in I$ per capita. Education: average years of education. Transition to independence: dummy variable (0 or 1). Armed conflict: dummy variable (0–3). Economic liberalization: Economic Freedom of the World index (0–10). Time: calendar-year. Robust standard errors in parentheses.
*** $p<0.01$, ** $p<0.05$, * $p<0.1$
a. The coefficients for current democracy have been multiplied by 20 to create comparability with the coefficients for cumulative democracy. They indicate the change in life expectancy at birth (in years), associated with a change from full autocracy (revised Polity2 index of 0) to full democracy (revised Polity2 index of 20).
b. The coefficients for cumulative democracy indicate the change in life expectancy at birth (in years), associated with an additional year of full democracy (revised Polity2 index of 20).

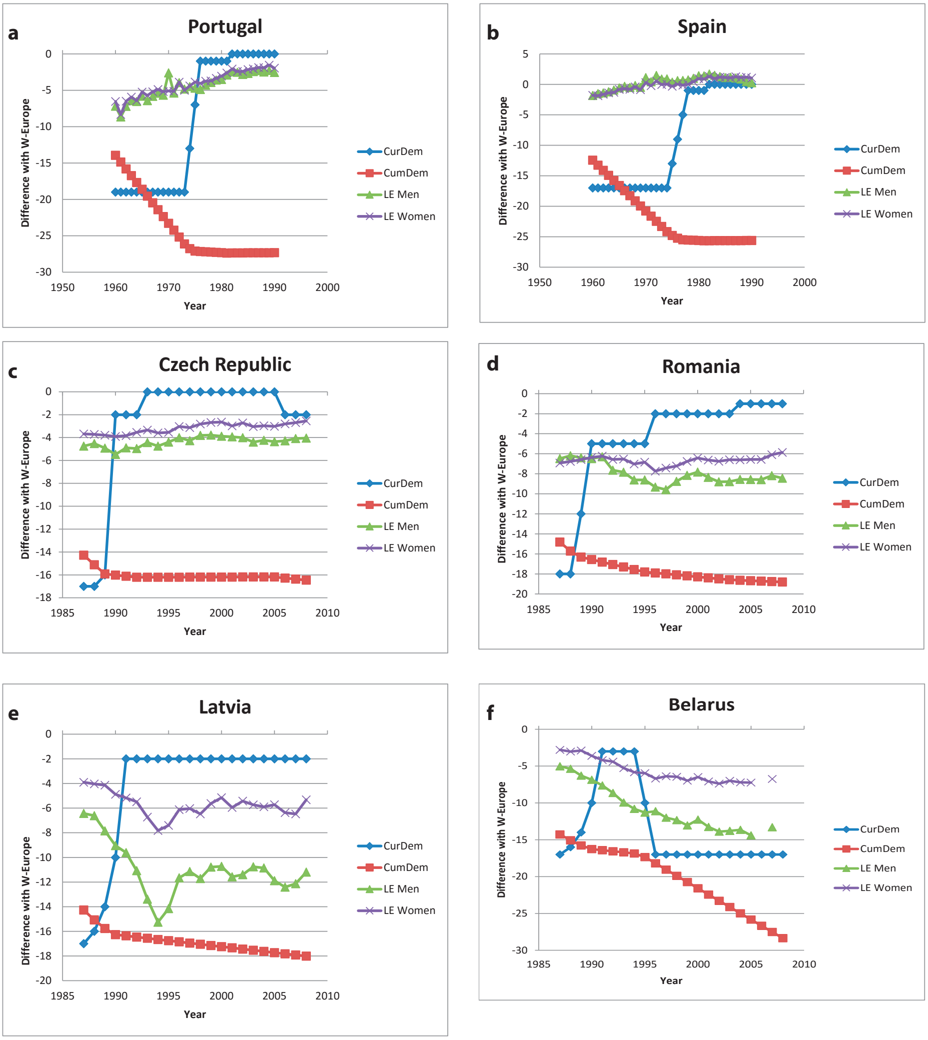


Figure 2 Graphical presentation of some illustrative results
a. Portugal, 1960–1990.
b. Spain, 1960–1990.
c. Czech Republic, 1987–2008.
d. Romania, 1987–2008.
e. Latvia, 1987–2008.
f. Belarus, 1987–2008
Notes: Graphs present differences in male and female life expectancy (LE, in years), current democracy (CurDem = revised Polity2 score, scale 0–20) and cumulative democracy (CumDem = years since 1946 or 1970 weighted by revised Polity2 score) with Western European average.

for men, are not statistically significant. The cumulative democracy results for men suggest that in this period one year of full democracy is associated with around two-thirds of a year of extra life expectancy. National income, armed conflict and time also have independent effects on life expectancy, and in this period the association with armed conflict (occurring mainly in the Balkans and the former Soviet Union) was negative.

Figures 2c to 2f illustrate what happened in this period to the gap in life expectancy between democratizing countries and the full democracies of Western Europe. In Central and Eastern Europe, some counties (exemplified by the Czech Republic) went through a rapid and radical process of democratization, with levels of current democracy rapidly converging with Western Europe, while others (exemplified by Romania) followed a more hesitating course of political reform. In the Czech Republic, the democratic deficit, as indicated by the gap with Western Europe in cumulative democracy, stabilized in the early 1990s, while it continued to grow in Romania. Both the Czech Republic and Romania saw a temporary deterioration of their life expectancy during the period in which current democracy rose rapidly, particularly among men, but recovery started earlier in the Czech Republic, where in the longer run the gap in life expectancy with Western Europe started to narrow, in contrast to Romania where the democratic deficit continued to grow (figures 2c and 2d).

In the former Soviet Union, some countries (exemplified by Latvia) made a rapid transition to full democracy, while others (exemplified by Belarus) fell back into autocratic forms of government. All countries of the former Soviet Union at first saw a severe deterioration of both male and female life expectancy, coinciding with a rapid rise of their current democracy status. As a result, the gap in life expectancy with the Western European average widened, but the health situation soon stabilized in countries in which the gap in cumulative democracy no longer grew rapidly, such as Latvia, while the life expectancy gap continued to widen in countries in which the democratic deficit continued to grow, such as Belarus (figures 2e and 2f).

Results for cause-specific mortality are presented in tables 3a and 3b. In the 1960–1990 period current democracy has a much stronger negative association with all-cause mortality than cumulative democracy. The effect is mainly mediated by lower mortality from all heart diseases (men only), pneumonia (both sexes), liver cirrhosis (both sexes), and suicide (men only). For mortality from motor vehicle accidents, however, we find a positive association, indicating that mortality from this cause rose during periods of democratization. Current democracy is also associated with lower mortality from signs, symptoms and ill-defined conditions, suggesting improvements in cause-of-death classification (table 3a).

In the 1987–2008 period cumulative democracy has a statistically significant, negative association with all-cause mortality among men (table 3b). Cumulative democracy is associated with lower mortality from all causes (men only), all circulatory diseases (both sexes), ischemic heart disease (men only), cerebrovascular disease (men only), cancer of the breast (women),

Table 3 Democracy vs. cause-specific mortality: results of country fixed effects regression models
a. 1960–1990.

	Men		Women	
	Current[a]	Cumulative[b]	Current[a]	Cumulative[b]
All causes	-166.9***	-16.1***	-110.2***	-10.4***
	(38.1)	(4.9)	(34.1)	(2.9)
All heart diseasses	-38.5*	-5.0***	-6.6	-3.3**
	(22.1)	(1.5)	(17.1)	(1.6)
Ischemic heart disease	-17.9	-3.9**	3.2	-2.7*
	(25.5)	(1.4)	(19.0)	(1.4)
Cerebrovascular disease	-8.0	-4.9***	-4.0	-4.1***
	(13.1)	(1.0)	(11.6)	(0.9)
All cancers	9.0	-0.7	-0.6	0.1
	(6.9)	(0.9)	(4.2)	(0.6)
Cancer of lung	0.2	-0.2	-2.2	0.3
	(5.4)	(0.5)	(1.6)	(0.2)
Cancer of breast			-0.4	-0.1
			(0.9)	(0.1)
Pneumonia	-32.5**	1.2	-21.6**	1.1
	(12.9)	(1.0)	(8.1)	(0.7)
Liver cirrhosis	-8.8***	-1.3***	-5.6***	-0.3*
	(2.9)	(0.3)	(1.3)	(0.2)
All external causes	2.9	-1.8*	1.1	-1.0**
	(3.5)	(0.9)	(2.1)	(0.3)
Motor vehicle accidents	12.1***	-0.8*	2.2***	-0.2**
	(2.4)	(0.4)	(0.5)	(0.1)
Suicide	-4.9**	-0.1	-0.8	0.0
	(2.1)	(0.3)	(0.9)	(0.1)
Signs, symptoms and ill-defined	-33.9**	0.0	-33.8***	0.1
	(13.0)	(1.9)	(11.1)	(1.6)

Notes: Country fixed effects models, estimated using Ordinary Least Squares regression with robust (clustered) standard errors, controlling for GDP, GDP squared, education, transition to independence, armed conflict, and time.
*** p<0.01, ** p<0.05, * p<0.1
a. The coefficients for current democracy have been multiplied by 20 to create comparability with the coefficients for cumulative democracy. They indicate the change in age-standardized death rates (in deaths per 100,000 person-years) associated with a change from full autocracy (revised Polity2 index of 0) to full democracy (revised Polity2 index of 20).
b. The coefficients for cumulative democracy indicate the change in age-standardized death rates (in deaths per 100,000 person-years) associated with an additional year of full democracy (revised Polity2 index of 20).

b. 1987–2008.

	Men		Women	
	Current[a]	Cumulative[b]	Current[a]	Cumulative[b]
All causes	69.8	-60.0***	-22.8	-17.3
	(56.2)	(21.5)	(24.5)	(10.3)
All circulatory diseases	9.3	-43.4***	-26.2	-14.4**
	(28.9)	(11.6)	(18.9)	(6.2)
Ischemic heart disease	-4.0	-24.3**	-13.8	-8.3
	(27.4)	(9.3)	(20.6)	(6.0)
Cerebrovascular disease	8.2	-9.7*	-3.4	-4.3
	(14.6)	(4.9)	(11.2)	(4.0)
All cancers	14.5**	-3.2	23.4**	-2.6
	(5.6)	(2.2)	(10.3)	(4.2)
Cancer of lung	5.1	-0.1	-0.0	1.0**
	(3.2)	(1.7)	(0.9)	(0.5)
Cancer of breast			3.8***	-1.8***
			(1.0)	(0.3)
All infectious diseases	-3.7	-1.1	-2.3**	0.2
	(2.2)	(1.2)	(0.9)	(0.5)
Chronic liver disease and cirrhosis	11.9	-2.3	3.2	-0.5
	(7.4)	(2.7)	(2.8)	(1.1)
All external causes	39.0**	-12.6***	5.7*	-3.5**
	(16.6)	(3.7)	(3.3)	(1.3)
Motor vehicle accidents	7.9***	-1.1	1.7**	-0.7
	(2.1)	(0.7)	(0.7)	(0.2)
Suicide	3.8	-0.3	-0.2	-0.3
	(4.4)	(1.1)	(1.0)	(0.4)
Signs, symptoms and ill-defined	-6.2	5.4	-1.6	4.3
	(16.0)	(7.2)	(13.1)	(5.4)

Notes: Country fixed effects models, estimated using Ordinary Least Squares regression with robust (clustered) standard errors, controlling for GDP, GDP squared, education, transition to independence, armed conflict, economic freedom, and time.

*** p<0.01, ** p<0.05, * p<0.1

a. The coefficients for current democracy have been multiplied by 20 to create comparability with the coefficients for cumulative democracy. They indicate the change in age-standardized death rates (in deaths per 100,000 person-years) associated with a change from full autocracy (revised Polity2 index of 0) to full democracy (revised Polity2 index of 20).

b. The coefficients for cumulative democracy indicate the change in age-standardized death rates (in deaths per 100,000 person-years) associated with an additional year of full democracy (revised Polity2 index of 20).

and all external causes (both sexes). As in the 1960–1990 period, current democracy is associated with higher mortality from motor vehicle accidents, but now also with higher mortality from all cancers, cancer of the breast (women), and all external causes.

In appendix table A3 we present the results obtained with alternative model specifications, including models with more and less stringent controls for confounding and for autocorrelation or spatial correlation, models based on dichotomized measures of democracy, and models with lagged effects of democracy. These checks indicate that our main results are robust against alternative model specifications. In the 1960–1990 period, positive associations between current democracy and life expectancy are found in nearly all models, with effects diminishing with longer lag-times. In the 1987–2008 period, positive associations between cumulative democracy and life expectancy are found in nearly all models, and these associations do not derive from lagged effects of current democracy (see appendix A3).

DISCUSSION

Summary of main findings

In the 1960–1990 period, current democracy was more strongly associated with higher life expectancy than cumulative democracy. The positive effects of current democracy on total mortality were mediated mainly by lower mortality from heart disease, pneumonia, liver cirrhosis, and suicide. In the 1987–2008 period, however, current democracy was associated with lower, and cumulative democracy with higher life expectancy, particularly among men. The positive effects of cumulative democracy on total mortality were mediated mainly by lower mortality from circulatory diseases, cancer of the breast, and external causes. Current democracy was associated with higher mortality from motor vehicle accidents in both periods, and also with higher mortality from cancer and all external causes in the second.

Limitations

A major strength of our study is that we have brought together a very extensive dataset, with a virtually complete coverage of our central variables democracy and life expectancy. Our measure of democracy, the Polity2 index, has become the standard in the international literature [6], and has a number of important strengths as compared to the main alternatives, such as the well-known Freedom House index: it has a broader empirical scope, particularly in historical terms, and is less susceptible to measurement error [34]. It also has a number of weaknesses, such as that it does not fully capture the dimension of participation (i.e., inclusiveness of the right to vote) [34], but the latter problem is less relevant in our study-period.

Period life expectancy is a commonly used, reliable and easily interpretable summary measure of mortality conditions pertaining to a particular point in time [36]. A limitation is that tempo effects can distort the measurement of life expectancy in times of rapidly declining or increasing mortality [37, 38], and that at lower rates of mortality larger declines are necessary for one unit increase in life expectancy [36]. Our analysis with age-standardized mortality from all causes produced broadly similar results (table 3), so it is unlikely that problems related to the life expectancy measure have biased our findings.

The cause-specific mortality data are less complete, and, more importantly, of uncertain validity. Over the study-period, the proportion of deaths classified as due to "symptoms and ill-defined conditions" declined substantially, indicating gradual improvement in the accuracy of cause-of-death certification and coding. In country fixed effects models controlling for various confounders, including time, democratization appeared to be associated with declining mortality from this garbage code during the 1960–1990 period (table 3a). This suggests improvements in the accuracy of cause-of-death classification during democratization which may have led to an underestimation of declines in mortality from specific conditions, and therefore to an underestimation of the effect of democratization on mortality trends from these conditions.

The main methodological challenge in studies of the relation between democracy and population health is establishing causality. Randomization to democracy and non-democracy conditions is, of course, impossible, and one therefore has to rely on observational study designs like ours. Any such study may suffer from 'endogeneity problems': democratization is an expression of collective human will, and may therefore be accompanied by many other changes affecting health outcomes (leading to 'omitted variable bias'), and there may even be 'reverse causation', in the sense that better population health promotes democracy instead of the other way around (leading to 'simultaneity bias')[35]. Our country-fixed effects models are designed to remove unobserved confounding by time-invariant confounders like religion, ethnic fractionalization, reliance on natural resources, history of democracy, etc. They do, however, not control for time-variant confounders not captured by our control variables national income, education, transition to independence, armed conflict and economic freedom. An example may be changes in value orientations: democratization in both Southern Europe and Central and Eastern Europe has followed shifts in value orientations [39] which may also have changed other collective and individual behaviours, independent from and parallel to the changes in the political system. We also cannot exclude some degree of reverse causation: for example, setbacks in life expectancy in the former Soviet Union may have contributed to the popular dissatisfaction that provided a fertile ground for autocratic setbacks in countries like Russia and Belarus [40]. This implies that causal inference can only be tentative, and will have to rely on more information than the results of regression models only.

In contrast to previous studies our study focuses on Europe, and has excluded non-European countries. During the study-period, democratization processes have occurred in many Latin American and Asian countries as well [2]. While a restriction to European countries has reduced the potential for confounding, and has allowed us to look in more depth at underlying processes through an analysis of cause-specific mortality, it may also have limited the generalizability of our findings. For example, democratization in Latin America involved countries in a different stage of economic development and without a history of communist economies, and the specific effects on population health may therefore have been different. Although our results for life expectancy are broadly in line with what has been found before in analyses covering a wider range of countries [6–8, 18, 20, 21], we advise caution in extrapolating our results for cause-specific mortality.

Interpretation

There is an increasing body of research studying the impact of political factors, broadly defined, on population health, and democracy is one of the most frequently studied topics within that literature [19, 31, 41]. As mentioned in the introduction, not all previous studies used rigorous analytic methods. For example, more than half of the studies of democracy versus life expectancy or mortality, including the frequently cited paper by Franco [20], are cross-sectional studies [8]. In our dataset, there also is a strong cross-sectional relation between democracy and life expectancy in Europe among both men and women, even after controlling for GDP and education, in both study-periods (results not shown). These results, however, are likely to be confounded by unobserved country characteristics.

Using more adequate analysis techniques we still found a positive association between indices of democracy and life expectancy, like a few previous studies did before us [7, 21–23]. In our country fixed effects models, we essentially relate changes in democracy levels within each country to changes in health outcomes, thereby removing effects of unobserved country characteristics. Unlike previous studies, focusing on Europe allowed us to look at specific causes of death, and these novel results provide important insights into the possible explanation of our findings. In the 1960–1990 period, the positive effects of current democracy on life expectancy as seen in our regression analyses were mediated mainly by lower mortality from heart disease, pneumonia, liver cirrhosis, and suicide, suggesting improvements in health care (indicated by lower pneumonia mortality) in addition to positive psychosocial effects (reflected in lower suicide mortality). In the 1987–2008 period, current democracy had a negative effect on life expectancy, particularly among men, mainly because it was associated with higher mortality from cancer and external causes, perhaps because of short-term disruptions to health care systems and road traffic safety. The positive effects of cumulative democracy on life expectancy in this period were mediated by lower mortality from circulatory diseases, breast cancer and injuries, suggesting gradual improvements in health care (for patients with ischemic heart and cerebrovascular disease and breast cancer) and injury prevention.

How plausible is it that democratization per se has led to these improvements in population health? Or is democratization only a proxy indicator for a wider range of changes occurring in these societies [42]? Despite the fact that the relation between democracy and population health has been studied quite extensively, the explanation of the study findings has generally remained speculative, due to lack of empirical data on mediating factors and mechanisms.

Developing a robust conceptual model, based on theoretical and empirical insights, that could guide further explanatory studies is therefore a clear research priority. We will discuss the explanation of our findings by reviewing in more detail the events and circumstances preceding and following democratization in Europe in the first and second study-periods.

In the first study-period, Spain, Portugal and Greece more or less simultaneously made a transition from military dictatorship to democracy. In the preceding years, all three countries had to some extent been isolated from developments occurring elsewhere in Western Europe, and democratization was followed by intensified interaction with other European countries, including accession to the European Union in the 1980s [3]. All three countries also started a process of modernization of their health care systems, with Portugal establishing a national health system in 1979 [43], Greece in 1983 [44], and Spain in 1986 [45].

In all three countries, the increased interaction with other European countries, and the associated modernization of all societal sectors including health care, are likely to have been the main mechanisms underlying the association between democratization and improvements in population health. That Portugal benefited from democracy despite the chaotic period following the 'Carnation revolution' can probably be explained from the fact that it was lagging behind so much more than the other two countries. This is exemplified by its delayed change away from a salty diet [46] and its remarkable trends in cerebrovascular disease mortality which in contrast to other European countries peaked in the early 1970s and then started a precipitous decline, probably due to dietary changes and better detection and treatment of hypertension [47].

To the best of our knowledge, no previous studies have been made of the impact of democratic reform in Spain, Portugal and Greece. A descriptive study of trends in life expectancy in Spain, Portugal and Greece concluded that most of the closing of the gap as compared to the Nordic countries occurred under authoritarian regime. "By the early 21st century Greece, Portugal and Spain had arrived at basically the same levels of population health as the Nordic nations, despite many years of authoritarian governments during the second half of the 20th century. [...] This raises serious doubts regarding the hypothesis that the political regime [...] exert[s] major influences on population health" [32]. As noted by the same author, however, the second half of the 1970s was a period of more rapid narrowing of the gap than previous periods, and this may well have been induced or facilitated by political reforms. We recommend in-depth, quantitative or qualitative studies linking changes in the political system in these three countries to changes in public policies, and then further downstream to changes in e.g. excessive alcohol consumption (liver cirrhosis), intake of salty foods (cerebrovascular disease), and use of antibiotics (pneumonia) and hypertensives (cerebrovascular disease), to empirically test our tentative explanations.

In the second study-period, countries in Central and Eastern Europe made a transition to various grades of democracy, but even more so than in the case of Spain, Portugal and

Greece this was not only a political reform. It was a double transition (not only political, also economic, with a change from a communist to a capitalist economy), and in some cases even a triple transition (with an additional dissolution of old states and formation of new states) [48]. Economic restructuring was accompanied by declining GDP and rising unemployment. Road injuries soared as a result of increased access to cars and inadequate road infrastructure [49].

Developments were even more dramatic in the Soviet Union which collapsed in 1991 and broke apart in separate republics with different trajectories of political and economic reforms. Within a few years the Baltic countries were on a stable upward track towards liberal democracy, and towards economic integration with the European Union, while democratization in countries like the Russian Federation, Belarus and Azerbaijan soon reversed with the return of some degree of authoritarian government [3]. In several countries of the former Soviet Union rapid restructuring of the economy, following the "shock therapy" recommended by the World Bank, led to a fall in GDP, rising unemployment and high human costs [40, 50].

Because of these double or even triple transitions, and the economic and governance crises that accompanied them, it is unreasonable to expect an upturn of life expectancy during or immediately after the "system change". A more plausible scenario is that the health situation temporarily worsened during this "system change", and only started to improve when countries succeeded in consolidating their democracy [51]. This corresponds to what we found: current democracy, a measure that captures short-term changes, had a negative association with life expectancy in this period, and only cumulative democracy was associated with higher life expectancy (table 2).

In these countries, the association between democratization and life expectancy is likely to reflect the wider policy changes that accompanied, and were promoted by, democratization: health care reform, economic restructuring, road traffic safety programs, health promotion campaigns, etc. [42]. Countries with more advanced democratic institutions are also likely to suffer less from corruption [52], which if present leads to massive inefficiencies in health care [53] and undermines preventive health policies [49]. Like in the case of Spain, Portugal and Greece in the 1970s and 1980s, changes in health care systems are likely to have been partial mediators [54]. These reforms led to a great improvement in access to new health care technologies, which has enabled progress in medical outcomes after decades of stagnation [55], and which may partly explain the observed declines in circulatory disease and breast cancer mortality (table 3b). Declining mortality from conditions amenable to medical intervention also suggests improvements in health care [56, 57]. Here again, in-depth studies of political, policy, risk factor and health changes in single countries could help in testing these and other hypotheses.

Conclusions

Our results suggest that in Europe during these two periods democratization has had mixed effects. That short-term changes in levels of democracy had positive effects in the first but not in the second period is probably due to the fact that democratization in Central and Eastern Europe was part of a complete system change which caused major societal disruptions. Over-all, however, democratization has probably created favourable conditions for reducing mortality from conditions amenable to health intervention, if not in the short then in the longer term.

REFERENCES

1. Judt, T., *Postwar. A history of Europe since 2005*. 2005, New York: Penguin Group.
2. Huntington, S.P., *The Third Wave: Democratization in the Late 20th Century*. 1991, Norman: University of Oklahoma Press.
3. Black, C.E., et al., *Rebirth. A political history of Europe since World War II*. 2000, Boulder CO: Westview Press.
4. Hague, R. and M. Harrop, *Comparative government and politics. An introduction*. Eighth ed. Comparative Government and Politics. 2010, Basingstoke: Palgrave Macmillan. 442.
5. Mulligan, C.B., R. Gil, and X. Sala-i-Martin, *Do democracies have different public policies than nondemocracies?* NBER Working Paper Series. Vol. Working Paper 10040. 2003, Cambridge MA: National Bureau for Economic Research.
6. Wigley, S. and A. Akkoyunlu-Wigley, *The impact of regime type on health: does redistribution explain everything?* World Politics, 2011. **63**(4): p. 647–677.
7. Besley, T. and M. Kudamatsu, *Health and democracy.* . Political Economy, 2006. **96**(2): p. 313–318.
8. Klomp, J. and J. de Haan, *Is the political system really related to health?* Soc Sci Med, 2009. **69**(1): p. 36–46.
9. Sen, A., *Development as freedom*. 1999, New York: Anchor Books.
10. Shandra, J.M., et al., *Dependency, democracy, and infant mortality: a quantitative, cross-national analysis of less developed countries.* Soc Sci Med, 2004. **59**(2): p. 321–33.
11. Friedman, M., *Capitalism and Freedom* 1962, Chicago: University of Chicago Press.
12. Reilly, R.G. and M. McKee, *'Decipio': examining Virchow in the context of modern 'democracy'.* Public Health, 2012. **126**(4): p. 303–7.
13. Mackenbach, J.P. and M. McKee, eds. *Successes and failures of health policy in Europe: four decades of diverging trends and converging challenges*. 2013, Open University Press: Buckingham.
14. Bunker, J.P., H.S. Frazier, and F. Mosteller, *Improving health: measuring effects of medical care.* Milbank Q, 1994. **72**(2): p. 225–58.
15. Mackenbach, J.P., *The contribution of medical care to mortality decline: McKeown revisited.* J Clin Epidemiol, 1996. **49**(11): p. 1207–13.
16. Ward, J.W. and C. Warren, eds. *Silent victories. The history and practice of public health in twentieth-century America.* 2007, Oxford University Press: Oxford.
17. Alvarez-Dardet, C. and A. Franco-Giraldo, *Democratisation and health after the fall of the Wall.* J Epidemiol Community Health, 2006. **60**(8): p. 669–71.
18. Safaei, J., *Is democracy good for health?* Int J Health Serv, 2006. **36**(4): p. 767–86.
19. Muntaner, C., et al., *Politics, welfare regimes, and population health: controversies and evidence.* Sociol Health Illn, 2011. **33**(6): p. 946–64.
20. Franco, A., C. Alvarez-Dardet, and M.T. Ruiz, *Effect of democracy on health: ecological study.* BMJ, 2004. **329**(7480): p. 1421–3.
21. Baum, M.A. and D.A. Lake, *The Political Economy of Growth: Democracy and Human Capital.* American Journal of Political Science, 2003. **47**(2): p. 333–347.
22. Li, Q. and M. Wen, *The Immediate and Lingering Effects of Armed Conflict on Adult Mortality: A Time-Series Cross-National Analysis.* Journal of Peace Research, 2005. **42**(4): p. 471–492.
23. Iqbal, Z., *Health and Human Security: The Public Health Impact of Violent Conflict.* International Studies Quarterly, 2006. **50**(3): p. 631–649.
24. Navia, P. and T.D. Zweifel, *Democracy, dictatorship and infant mortality revisited.* Jorunal of Democracy, 2003. **14**(3): p. 90–103.

25. Zweifel, T.D. and P. Navia, *Democracy, dictatorship, and infant mortality.* Journal of Democracy 2000. **11**(2): p. 99–114.
26. Gerring, J., S.C. Thacker, and R. Alfaro, *Democracy and Human Development.* The Journal of Politics, 2012. **74**(1): p. 1–17.
27. Ross, M., *Is democracy good for the poor.* American Journal of Political Science, 2006. **50**(4): p. 860–874.
28. Houweling, T.A., C.W.N. Looman, and J.P. Mackenbach, *Determinants of under-5 mortality among the poor and the rich: a cross-national analysis of 43 developing countries.* Int J Epidemiol, 2005. **34**(6): p. 1257–65.
29. Roemer, M.I., *Health development and political policy: the lesson of Cuba.* J Health Polit Policy Law, 1980. **4**(4): p. 570–8.
30. Navarro, V., *Has socialism failed? An analysis of health indicators under socialism.* Int J Health Serv, 1992. **22**(4): p. 583–601.
31. Mackenbach, J.P., *Potlitical conditions and life expectancy in Europe, 1900–2008.* Social Science and Medicine, 2013. **82**: p. 134–146.
32. Tapia Granados, J.A., *Politics and health in eight European countries: a comparative study of mortality decline under social democracies and right-wing governments.* Soc Sci Med, 2010. **71**(5): p. 841–50.
33. Teorell, J., et al., *The Quality of Government Dataset, version 6Apr11.*, U.o.G.T.Q.o.G. Institute, Editor. 2011: Gothenburg.
34. Munck, G. and J. Verkuylen, *Conceptualizing and Measuring Democracy: Evaluating Alternative Indices.* Comparative Political Studies, 2002. **35**(1): p. 5–34.
35. Wooldridge, J.M., *Introductory Econometrics. A Modern Approach.* 2006, Mason (OH): Thomson South-Western. 891.
36. Preston, P. Heuveline, and M. Guillot, *Demography. Measuring and Modeling Population Process.* 2001, Oxford: Blackwell Publishing.
37. Bongaarts, J. and G. Feeney, *How long do we live.* Population and Development Review, 2002. **28**(1): p. 13–29.
38. Luy, M., *Mortality tempo-adjustment: An empirical application.* Demographic Research, 2006. **15**(21): p. 561–590.
39. Inglehart, R. and C. Welzel, *Modernization, cultural change, and democracy. The human development sequence.* 2005, Cambridge: Cambridge University Press.
40. Grigoriev, P., et al., *Mortality in Belarus, Lithuania, and Russia: Divergence in Recent Trends and Possible Explanations.* European Journal of Population, 2010. **26**: p. 245–274.
41. Beckfield, J. and N. Krieger, *Epi + demos + cracy: linking political systems and priorities to the magnitude of health inequities--evidence, gaps, and a research agenda.* Epidemiol Rev, 2009. **31**: p. 152–77.
42. McKee, M. and E. Nolte, *Lessons from health during the transition from communism.* BMJ, 2004. **329**(7480): p. 1428–9.
43. Barros, P.P., S.R. Machado, and A. Simoes Jde, *Portugal. Health system review.* Health Syst Transit, 2011. **13**(4): p. 1–156.
44. Economou, C., *Greece: Health system review.* Health Syst Transit, 2010. **12**(7): p. 1–177, xv-xvi.
45. Garcia-Armesto, S., et al., *Spain: Health system review.* Health Syst Transit, 2010. **12**(4): p. 1–295, xix-xx.
46. Mackenbach, J.P., *Bacalhao under the Ponte 25 de Abril: impressions from Lisbon.* Eur J Public Health, 2009. **19**(1): p. 1.
47. Mirzaei, M., et al., *Cerebrovascular disease in 48 countries: secular trends in mortality 1950–2005.* J Neurol Neurosurg Psychiatry, 2012. **83**(2): p. 138–45.
48. Bozoki, A., *Democratization in Central Europe.* Taiwan Journal of Democracy, 2008. **4**(2): p. 1–27.

49. McKee, M., et al., *Health policy-making in central and eastern Europe: lessons from the inaction on injuries?* Health Policy Plan, 2000. **15**(3): p. 263–9.
50. Stuckler, D., L. King, and M. McKee, *Mass privatisation and the post-communist mortality crisis: a cross-national analysis.* Lancet, 2009. **373**(9661): p. 399–407.
51. Kaldor, M. and I. Vejvoda, *Democratization in Central and East European Countries.* International Affairs (Royal Institute of International Affairs 1944-), 1997. **73**(1): p. 59–82.
52. Sung, H.E., *Democracy and political corruption: A cross-national comparison*. Crime, Law and Social Change, 2004. **41**(2): p. 179–193.
53. Transparency International, *Global Corruption Report 2006. Special Focus: Corruption and Health*. 2006, Pluto Press: London.
54. Figueras, J., et al., eds. *Health systems in transition: learning from experience*. 2004, World Health Organization, European Observatory on Health Systems and Policies.
55. Stockholm Network Research Team, *Health Care Reform in Central and Eastern Europe: Setting the Stage for Discussion*. 2008, Stockholm Network: Stockholm.
56. Nolte, E., R. Scholz, and M. McKee, *Progress in health care, progress in health? Patterns of amenable mortality in central and eastern Europe before and after political transition.* Demographic Research, 2004. **SPECIAL COLLECTION 2, ARTICLE 6**: p. 139–162.
57. Koupilova, I., M. McKee, and J. Holcik, *Neonatal mortality in the Czech Republic during the transition.* Health Policy, 1998. **46**(1): p. 43–52.
58. Bruce, S., *Did Protestantism Create Democracy?* Democratization, 2004. **11**(4): p. 3–20.
59. Aghion, P., A. Alesina, and F. Trebbi, *Endogenous political institutions*, in *NBER Working Paper Series*. 2002, National Bureau for Economic Research: Cambridge, MA.
60. Powell-Jackson, T., et al., *Democracy and growth in divided societies: A health-inequality trap?* Soc Sci Med, 2011. **73**(1): p. 33–41.
61. Haber, S. and V. Menaldo, *Do Natural Resources Fuel Authoritarianism? A Reappraisal of the Resource Curse.* American Political Science Review, 2011. **105**(1): p. 1–26.
62. Ross, M., *Does Oil Hinder Democracy?* World Politics, 2001. **53**(April): p. 325–61.
63. Papaioannou, E. and G. Siourounis, *Economic and Social Factors Driving the Third Wave of Democratization*. Dartmouth College, Economics Department: Hanover, NH
64. Glaeser, E.L., G. Ponzetto, and A. Shleifer, *Why democracy needs education?* NBER Working Paper Series. Vol. Working Paper 12128. 2006, Cambridge MA: National Bureau of Economic Research.
65. Lipset, S.M., *Some Social Requisites of Democracy: Economic Development and Political Legitimacy.* The American Political Science Review, 1959. **53**(1): p. 69–105.
66. Acemoglu, D., et al., *Income and democracy.* American Economic Review, 2008. **98**(3): p. 808–842.
67. Inglehart, R., *Modernization and postmodernization. Cultural, economic and political change in 43 societies.* 1997, Princeton: Princeton University Press.
68. Przeworksi, A., *Democracy and the market. Political and economic reforms in Eastern Europe and Latin America*. 1991, Cambridge: Cambridge Univerity Press.
69. Melzyg, R., P. Pleva, and R. Sprout, *Economic reforms, democracy and growth in Eastern Europe and Eurasia.*, in *Working Paper Series on the Transition Countries*. 2005, Program Office, Bureau for Europe & Eurasia, U.S. Agency for International Development: Washington, DC.
70. Preston, *The changing relation between mortality and level of economic development.* Popul Stud (Camb), 1975. **29**(2): p. 231–248.
71. Maddison, A., *The world economy: historical statistics.* . 2003, Paris: Organisation of Economic Cooperation and Development.

72. Caldwell, J.C., *Routes to Low Mortality in Poor Countries*. Population and Development Review, 1986. **12**(2): p. 171–220.

73. Gwartney, J., J. Hall, and R. Lawson, *2011 Economic Freedom Dataset, published in Economic Freedom of the World: 2010 Annual Report*, E.F. Network, Editor. 2011.

APPENDIX

Further details on control variables

In the analysis we will control for potential time-invariant confounders, such as religion [58], ethnic fractionalization [59, 60] and reliance on natural resources [61, 62], by using country fixed effects models (see Analysis section). We will explicitly control for the following time-variant control variables:

i. National income. According to a popular modernization theory, democracy is more likely to emerge and consolidate in economically developed societies [63–66]. A higher level of economic development often leads to cultural changes promoting democratization, partly through a more highly educated public and a larger middle class [2], partly through shifts in value systems [67] which create a demand for democracy [39]. On the other hand, economic crises may undermine the legitimacy of authoritarian regimes and precipitate democratization [2], and may also undermine the support for democratic institutions when these are no longer able to deliver gains to everyone [68]. In analyses of democratic reforms in Central and Eastern Europe, a mutual association has been found with economic reforms, with economic reforms increasing the likelihood of democratic reforms and vice versa [69]. Economic development is also a determinant of life expectancy [70]. Data on national income (in $1000s) per head of population were extracted from a dataset compiled by Maddison (http://www.ggdc.net/MADDISON/oriindex.htm). The national estimates of Gross Domestic Product (GDP) per head of population in this dataset are based on extensive harmonization efforts and on a conversion into 1990 International Geary-Khamis dollars using multilateral Purchasing Power Parities [71]. In order to allow for non-linear relations between national income and life expectancy we will use both GDP and GDP-squared as control variables.
ii. Education. This too is a determinant of both democracy [64] and life expectancy [72]. Average years of schooling (among those aged 25 years and over) were extracted from the Barro-Lee Educational Attainment dataset (http://www.barrolee.com/). This dataset has been constructed on the basis of school attainment figures collected in national censuses and surveys, as compiled by UNESCO, Eurostat, and other sources, using extensive harmonization methods. Because this dataset only presents data at 5-year intervals, we have imputed the data for other calendar-years by linear interpolation.
iii. Transition to independence. Cyprus only gained independence in 1960, and Malta in 1964. In the Soviet Union, the political changes after 1989 precipitated a dissolution into a large number of successor states some of which became formally independent in 1990 (e.g. the Baltic Republics), others in 1991 (e.g. Republic of Moldova). Similarly, the successor republics of Yugoslavia and Czechoslovakia obtained independence in 1990 (Slovenia), 1991 (e.g. Croatia), 1992 (e.g. Bosnia and Herzegovina) and 1993 (Czech and Slovak Republics). The disruption of governance structures accompanying transition to independence may temporarily depress life expectancy. Transition

to independence was coded as dummy variables (0 for no transition, 1 for the year of independence and the three subsequent years).

iv. Armed conflict. This has accompanied the political changes in Central and Eastern Europe after 1989, particularly after the dissolution of Yugoslavia and the Soviet Union, and may have affected population health either directly (by war casualties) or indirectly (by directing resources away from health care and other sectors important for population health). We used data on conflict location as collected by the Department of Peace and Conflict Research of Uppsala University (http://www.pcr.uu.se/research/ucdp/datasets/ucdp_prio_armed_conflict_dataset/), made available in the Quality of Government dataset [33] (http://www.qog.pol.gu.se/data/). Armed conflict was coded as dummy variables (ranging from 0 for no conflict to 3 for more than 1000 battle deaths per year).

v. Economic freedom. Previous studies have suggested that rapid economic liberalization has contributed to increases in mortality during the transition to democracy in Central and Eastern Europe [50]. We will indicate economic liberalization with the Economic Freedom of the World index, that indicates the presence of economic freedom on a 0–10 scale, based on size of government, security of property rights, access to sound money, freedom to trade internationally, and regulation of credit, labor and business. We used the chain-linked summary index as contained in the Economic Freedom of the World 2011 dataset (http://www.freetheworld.com/2011/2011/Dataset.xls)[73], and imputed some missing data for earlier periods using the 2002 dataset update (http://www.freetheworld.com/release_2002.html). This variable was not available before the 1980s.

Table A1 Cause-of-death codes
a. 1960–1990.

	ICD8[a]
All heart diseases	80–84
Ischemic heart disease	83
Cerebrovascular disease	85
All cancers	45–60
Cancer of lung[b]	50, 51
Cancer of breast	54
Pneumonia	91, 92
Liver cirrhosis	102
All external causes	138–148
Motor vehicle accidents	138
Suicide	147
Signs, symptoms and ill-defined	136, 137

a. International Classification of Diseases, 8th revision, Basic tabulation list
b. Includes cancer of other respiratory organs

b. 1987–2008.

	ICD10[a]
All circulatory diseases	I00-I09
Ischemic heart disease	I20-I25
Cerebrovascular disease	I60-I69
All cancers	C00-C97
Cancer of lung[b]	C33, C34
Cancer of breast	C50
All infectious diseases	A00-A99, B00-B99
Chronic liver disease and cirrhosis	K70, K73, K74
All external causes	V00-V99, W00-W99, X00-X99, Y00-Y99
Motor vehicle accidents	V02-V04, V09, V12-V14, V20-V79, V82-V87, V89
Suicide	X60-X84
Signs, symptoms and ill-defined	R00-R53, R55-R99

a. International Classification of Diseases, 10th revision
b. Includes cancer of other respiratory organs

Table A2 Descriptive statistics of the variables used in the analysis
a. 1960–1990.

VARIABLES	N	mean	sd	min	max
Current democracy (di)	887	13.60	8.110	1	20
Cumulative democracy (dic1)	837	18.98	14.02	0.750	45
National income (gdp)	858	9.374	4.898	1.451	30.571
Education (edu)	806	7.242	1.712	2.750	10.87
Armed conflict (war)	899	0.0612	0.333	0	3
Independence (inde)	899	0.00890	0.0940	0	1
Year	899	1,975	8.949	1,960	1,990
Country	899	15	8.371	1	29
Life expectancy (m)	819	69.18	2.807	59.94	76.23
Life expectancy (f)	819	75.34	2.906	62.80	81.83
All causes (m)	689	1,299	187.8	903.2	1,735
All heart diseases (m)	689	371.5	94.14	157.1	652.0
Ischemic heart disease (m)	689	258.9	109.6	55.47	560.3
Cerebrovascular disease (m)	689	150.4	54.47	49.40	343.9
All cancers (m)	689	246.0	45.14	136.2	367.4
Cancer of lung (m)	689	69.66	25.20	18.29	127.9
Pneumonia (m)	689	44.05	23.22	6.390	136.1
Liver cirrhosis (m)	681	24.41	16.77	2.790	77.10
All external causes (m)	641	93.36	28.76	21.24	182.6
Motor vehicle accidents (m)	632	27.18	10.29	7.990	64.24
Suicide (m)	629	23.86	13.30	3.150	70.31
Signs, symptoms and ill-defined (m)	678	67.77	78.80	0.230	424.9
All causes (f)	689	854.0	151.2	502	1,228
All heart diseases (f)	689	234.3	64.38	98.18	472.2
Ischemic heart disease (f)	689	138.7	63.39	33.81	303.1
Cerebrovascular disease (f)	689	131.1	46.78	46.88	277.8
All cancers (f)	689	154.2	26.34	99.85	212.4
Cancer of lung (f)	689	10.17	5.495	3.200	34.73
Cancer of breast (f)	689	25.82	8.002	6.700	42.93
Pneumonia (f)	689	31.60	18.27	4.800	93.36
Liver cirrhosis (f)	681	9.444	5.918	1.380	29.19
All external causes (f)	641	39.28	12.14	6.120	74.46
Motor vehicle accidents (f)	632	8.086	2.773	1.730	17.06
Suicide (f)	629	8.841	5.110	0.570	26.49
Signs, symptoms and ill-defined (f)	678	54.42	67.17	0.140	378.5

b. 1987–2008.

VARIABLES	N	mean	sd	min	max
Current democracy (di)	932	16.67	5.494	1	20
Cumulative democracy (dic1)	932	17.96	10.76	0.900	39
National income (gdp)	934	12.136	7.997	1.632	53.962
Education (edu)	814	9.632	1.323	5.735	13.09
Armed conflict (war)	946	0.122	0.513	0	3
Independence (inde)	946	0.0793	0.270	0	1
Economic liberalization (free)	724	6.609	1.219	2.775	8.390
Year	946	1,998	6.348	1,987	2,008
Country	946	22	12.42	1	43
Life expectancy (m)	856	71.11	4.819	57.12	80.14
Life expectancy (f)	856	78.22	3.386	58.30	85.01
All causes (m)	848	1,181	354.6	595.1	2,382
All circulatory diseases (m)	825	554.1	236.0	165.7	1,184
Ischemic heart disease (m)	824	276.8	168.4	53.28	737.6
Cerebrovascular disease (m)	825	137.2	75.13	30.92	379.4
All cancers (m)	825	183.0	33.33	75.22	278.0
Cancer of lung (m)	824	69.46	20.80	16.95	120.4
All infectious diseases (m)	824	13.37	9.832	1.170	61.07
Chronic liver disease and cirrhosis (m)	748	26.97	22.97	0	128.4
All external causes (m)	825	110.6	77.88	26.48	459.9
Motor vehicle accidents (m)	793	20.14	10.51	1.430	68.79
Suicide (m)	818	26.10	17.55	1.160	89.27
Signs, symptoms and ill-defined (m)	799	36.08	42.95	0.340	346.7
All causes (f)	848	697.5	181.7	380.5	1,358
All circulatory diseases (f)	825	362.7	160.5	96.43	758.8
Ischemic heart disease (f)	824	150.5	107.1	12.80	517.4
Cerebrovascular disease (f)	825	109.7	57.98	23.42	277.0
All cancers (f)	825	249.3	52.22	93.85	391.1
Cancer of lung (f)	824	14.51	8.448	3.240	44.54
Cancer of breast (f)	825	25.95	6.601	6.020	53.99
All infectious diseases (f)	824	5.657	2.725	0.300	19.62
Chronic liver disease and cirrhosis (f)	748	11.59	14.87	0	105.1
All external causes (f)	825	34.21	25.95	7.890	611.1
Motor vehicle accidents (f)	793	5.762	2.703	0	16.11
Suicide (f)	818	6.842	3.835	0	23.18
Signs, symptoms and ill-defined (f)	799	23.52	34.57	0.210	291.5

Table A3 Results of regression analyses with alternative model specifications
a. 1960–1990.

		Current democracy (di)		Cumulative democracy (dic1)		
Model		Male life expec	Female life exp	Male life expec	Female life exp	Notes
1	Confounders, country fixed effects	0.0928***	0.0918***	0.0690	0.0113	As in table 2
		(0.0311)	(0.0296)	(0.0523)	(0.0487)	
2	Confounders only	0.123***	0.0614***	0.0940***	0.0466***	As listed in table 2
		(0.0164)	(0.0114)	(0.0113)	(0.00905)	
3	No controls at all	0.208***	0.218***	0.144***	0.160***	Bivariate association
		(0.00979)	(0.0102)	(0.00466)	(0.00472)	
4	Feasible General Least Square (FGLS)	0.0314**	0.0427***	0.0618*	0.0100	Serial autocorrelation
		(0.0133)	(0.0119)	(0.0316)	(0.0214)	
5	Driscoll-Kraay standard errors	0.0928***	0.0690*	0.0918***	0.0113	Spatial autocorrelation
		(0.0313)	(0.0345)	(0.0228)	(0.0267)	
6	Two-way fixed effects	0.0866**	0.0414	0.0960***	0.00400	With year dummies
		(0.0350)	(0.0605)	(0.0302)	(0.0540)	
7	Democracy dichotomized	1.129**	1.160**	0.0743	0.0242	di<10=0, di>10=1
		(0.531)	(0.521)	(0.0438)	(0.0402)	
8	Democracy lagged 1 year	0.106***	0.101***			Current democracy, year t-1
		(0.0336)	(0.0323)			
9	Democracy lagged 3 years	0.0975**	0.0898**			Current democracy, year t-3
		(0.0373)	(0.0329)			
10	Democracy lagged 10 years	0.0360	0.0410			Current democracy, year t-10
		(0.0306)	(0.0253)			

b. 1987–2008.

		Current democracy (di)		Cumulative democracy (dic1)		
Model		Male life expec	Female life exp	Male life expec	Female life exp	Notes
1	Confounders, country fixed effects	-0.0761**	-0.00286	0.674**	0.248	As in table 2
		(0.0337)	(0.0206)	(0.295)	(0.195)	
2	Confounders only	0.0967**	0.0859***	0.399***	0.125***	As listed in table 2
		(0.0396)	(0.0203)	(0.0263)	(0.0144)	
3	No controls at all	0.455***	0.353***	0.366***	0.263***	Bivariate association
		(0.0257)	(0.0198)	(0.00790)	(0.00592)	
4	Feasible General Least Square (FGLS)	-0.000551	0.0279*	0.312***	0.201***	Serial autocorrelation
		(0.0192)	(0.0167)	(0.0354)	(0.0303)	
5	Driscoll-Kraay standard errors	-0.0761***	-0.00286	0.674***	0.248***	Spatial autocorrelation
		(0.0149)	(0.00945)	(0.0639)	(0.0584)	
6	Two-way fixed effects	-0.0655*	-0.00195	0.687**	0.267	With year dummies
		(0.0342)	(0.0219)	(0.323)	(0.216)	
7	Democracy dichotomized	-1.598***	-0.432	0.315***	0.186***	di<10=0, di>10=1
		(0.491)	(0.283)	(0.0774)	(0.0235)	
8	Democracy lagged 1 year	-0.0580	0.00135			Current democracy, year t-1
		(0.0368)	(0.0216)			
9	Democracy lagged 3 years	-0.0305	0.000283			Current democracy, year t-3
		(0.0267)	(0.0175)			
10	Democracy lagged 10 years	-0.00434	-0.00241			Current democracy, year t-10
		(0.0127)	(0.00824)			

Notes: We have performed various robustness checks of our main results for life expectancy, by comparing these with the results obtained with alternative model specifications.
In models with less extensive controls for confounding, such as a model without country fixed effects (model 2) or a model without any control for confounding (model 3) we find much stronger positive effects of democracy, but these are likely to be upwardly biased.
A Wald-test suggested by Wooldridge (2002) indicated that first-order serial autocorrelation exists in our dataset. We have therefore compared our main results with those of a Feasible General Least Square (FGLS) model, which corrects for panel-specific autocorrelation (model 4). However, this does not materially change our conclusions based on a model with clustered standard errors, and as the FGLS model has been criticized for its underestimation of standard errors (Beck and Katz, 1995), we prefer the results of model 1. We also applied the fixed effect regression model with Driscoll-Kraay standard errors to deal with potential heteroskedasticity and various forms of autocorrelation including spatial autocorrelation (model 5) (Hoechle, 2007). Again, these alternative model specifications do not materially change our conclusions.
In our main model, we control for time trends using a continuous variable representing calendar-years. This assumes a linear relationship between life expectancy and time. This assumption was relaxed in a two-way fixed effects model with dummies for each calendar-year (model 6). Again, these alternative model specifications do not materially change our conclusions.
The same applies to the use of dichotomized democracy scores (in which positive values of the original Polity2 index were scored as "1", and negative values as "0") (model 7).
Finally, we looked at delayed effects of current democracy, using lag-times of 1, 3 and 10 years. To this end, we replaced the current democracy by 1 year's, 3 years' and 10 years' lagged terms of democracy respectively in the fixed effects model. In the 1960–1990 period, positive associations between current democracy and life expectancy diminish with longer lag-times, suggesting that the associations with current democracy as found in our main model indeed indicate immediate or short-term effects. The same applies to the negative associations between current democracy and life expectancy in the 1987–2008 period (models 8, 9 and 10).
Regression coefficients for current democracy as presented in appendix table 3 have not been multiplied by 20, as in tables 2 and 3.

REFERENCES:

Wooldridge, J. M. 2002. Econometric Analysis of Cross Section and Panel Data. Cambridge, MA: MIT Press.

Beck, N., and J. N. Katz. 1995. What to do (and not to do) with time-series cross-section data. American Political Science Review 89: 634-647.

Hoechle, D. 2007. Robust standard errors for panel regressions with cross-sectional dependence. The Stata Journal,7 No.3, 281-312.

Chapter 3

Income inequality, life expectancy and mortality in Europe, 1987–2008

Yannan Hu, Frank J. van Lenthe, Johan P. Mackenbach.

Income inequality, life expectancy and cause-specific mortality in 43 European countries, 1987–2008. European Journal of Epidemiology, 2015. 30: p. 615–625.

ABSTRACT

Whether income inequality is related to population health is still open to debate. We aimed to critically assess the relationship between income inequality and mortality in 43 European countries using comparable data between 1987 and 2008, controlling for time-invariant and time-variant country-level confounding factors.

Annual data on income inequality, expressed as Gini index based on net household income, were extracted from the Standardizing the World Income Inequality Database (SWIID). Data on life expectancy at birth and age-standardized mortality by cause of death were obtained from the Human Lifetable Database and the World Health Organization European Health for All Database. Data on infant mortality were obtained from the United Nations World Population Prospects database. The relationships between income inequality and mortality indicators were studied using country fixed effects models, adjusted for time trends and country characteristics.

Significant associations between income inequality and many mortality indicators were found in pooled cross-sectional regressions, indicating higher mortality in countries with larger income inequalities. Once the country fixed effects were added, all associations between income inequality and mortality indicators became insignificant, except for mortality from external causes and homicide among men, and cancers among women. The significant results for homicide and cancers disappeared after further adjustment for indicators of democracy, education, transition to national independence, armed conflicts, and economic freedom.

Cross-sectional associations between income inequality and mortality seem to reflect the confounding effects of other country characteristics. In a European context, national levels of income inequality do not have an independent effect on mortality.

INTRODUCTION

Whether income inequality harms population health is still open to debate. Since Wilkinson [1] postulated the hypothesis that income inequality was not simply a summary of the balance of income between the rich and poor, but is a health risk in its own right [2], a wide array of studies, including multilevel studies within countries and cross-country ecological studies examined the link between income distribution and population health [3, 4]. However, no agreement has yet been reached because of discrepancies between the results of different studies.

International comparative studies linking income inequality to mortality suffer from limited comparability of the income inequality measures between countries and over time [5–7]. The Luxembourg Income Study (LIS) [8], regarded as the "gold standard", is the first choice for many studies [1, 5, 9, 10] because of its high quality and comparability. It covers, however, only a limited set of country-year observations, which may be the reason why many studies using this database performed a cross-sectional analysis. The Deininger and Squire database (1996) is often chosen as an alternative source [6, 11, 12] and provides more observations, but at a substantial loss of comparability. The World Income Inequality Database (WIID) covers the most comprehensive set of income inequality statistics. It incorporates several data sources, and enables researchers to maximize comparability by choosing data based on the criteria of comparability, but potentially leads to a risk of not piecing together the information in a meaningful way [13, 14]. The more recently developed Standardizing the World Income Inequality Database (SWIID) maximizes comparability for the broadest available set of country-year observations, and as such is better suited than other income inequality datasets for cross-country comparative research [14].

With some exceptions [9, 12], cross-sectional studies found significantly worse population health at higher levels of income inequality [1, 15–17]. However, these associations sometimes diminished after adjustment for observed country characteristics [3, 4, 6, 11, 18], suggesting there is a substantial risk of confounding. Fixed effect models, which require longitudinal data, are able to adjust for unobservable time-invariant confounding variables, by linking changes in income inequality to changes in health. Studies using fixed effects models to study the effect of income inequality on population health often reported insignificant results [6, 7, 11, 19–21]. However, these studies pooled men and women together [6, 7, 10, 11, 19, 22, 23], used relatively old data [6, 11, 19], restricted the outcome to infant mortality [21], or ignored some potential time-variant confounders [6, 10, 19, 20, 23, 24]. Only few studies investigated disease-specific outcomes, which would help to interpret findings on the basis of existing knowledge on determinants of population health and could point towards potential pathways through which income inequality may harm population health [7, 11, 19]. Studies specifically assessing the association between income inequality and mortality in a European context, which would be important for policy makers in Europe, are also limited in number [25, 26].

Using the SWIID data, we therefore aimed to refine and extend previous studies by critically investigating the relationships between income inequality and a set of disease-specific mortality indicators by gender in fixed effects models for 43 European countries over the period 1987–2008 [21].

DATA AND METHODS

Data

For income inequality, we made use of a new dataset called Standardizing the World Income Inequality Database (SWIID). Using the Gini index as measure, SWIID took version 2.0c of the WIID [27] as the starting-point and standardized it based on the inequality observations from the LIS [8]. Standardizing procedures were applied to account for differences in (a) population coverage (e.g. whether data cover all or nearly all of a country's population), (b) income reference units (e.g. household per capita, household adult equivalent, or household without adjustment of number of people), and (c) the definition of income (e.g. net income, gross income, expenditures or unidentified income). Finally, missing observations were imputed based on proximate years using a custom multiple-imputation algorithm [14]. In this study, we extracted information on the Gini index based on net household income (post-tax post-transfer) from SWIID version 4.0 covering 43 European countries with 879 country-year observations.

Data on life expectancy at birth and age-standardized mortality by cause of death at all ages (further referred to as "mortality indicators") were extracted from the Human Lifetable Database (www.lifetable.de) and the World Health Organization European Health for All Database (http://data.euro.who.int/hfadb/). Data on infant mortality, measured as infant deaths per 1000 live births, were obtained from the United Nations World Population Prospects database (http://esa.un.org/wpp/Excel-Data/mortality.htm). All mortality rates are log-transformed for normalization. ICD-code numbers are reported in a previous paper [28].

A key variable potentially confounding the relationship between income inequality and mortality, and for which the majority of existing studies controlled, is national income, which was measured by Gross Domestic Product (GDP) per head of population (in $1000s, extracted from a dataset compiled by Maddison http://www.ggdc.net/MADDISON/oriindex.htm). Besides GDP, a number of other potential confounders were added where appropriate, including indicators of democracy (the Policy2 index ranging from "strongly democratic (+10)" to "strongly autocratic (-10)", extracted from the Quality of Government dataset, http://www.qog.pol.gu.se/data/), average years of schooling (extracted from the Barro-Lee Educational Attainment dataset, http://www.barrolee.com/, made into annual data by linear interpolation), transition to national independence (0 for no transition, 1 for the year of independence and the three subsequent years), armed conflict (ranging from 0 for "no conflict" to 3 for "more than 1000 battle deaths per year", constructed using data on conflict location extracted from the Quality of Government dataset, http://www.qog.pol.gu.se/data/), and economic freedom (ranging from 0 to 10, extracted from the Economic Freedom of the World 2011 dataset and imputed some missing data for earlier periods using the 2002 dataset update, http://www.freetheworld.com/). Indicators of democracy, education and economic freedom were chosen because they represent changes in the underlying political, social and financial conditions prevailing in each country. Indicators of transition to national independence and armed conflicts were chosen to account for the disruption of governance structures and political changes in Central and Eastern Europe, and some Mediterranean and Western Balkan countries during the study-period. All these variables have been documented to affect population health in Europe [28] and potentially correlate with but may not be seen as causally resulting from income inequality [29–31].

Analytical approach

To align with previous studies, we first explored the pooled cross-sectional relation between income inequality and mortality. We adjusted analyses for year dummies and GDP per capita, and subsequently included more confounding variables. Robust standard errors were used to account for heteroskedasticity [32].

The model can be written as:

$$healthoutcome_{ij} = \alpha + \beta_1 Gini_{ij} + \beta_2 lngdp_{ij} + T_j + \gamma C_{ij}$$

where $healthoutcome_{ij}$ is the life expectancy or logarithmic form of the age-adjusted mortality rates for country i in year j; α is a constant; $Gini_{ij}$ represents the Gini index; $lngdp_{ij}$ is the logarithmic form of GDP per head, accounting for the potential non-linear relationship between national income and health; T_j is a vector of year dummies controlling the shared time trend in mortality during the study period; C_{ij} represents other potential time-variant confounders including indicators of democracy, years of schooling, independence, armed conflict and economic freedom, which were added subsequently.

As the next step, we applied fixed effects models, which allowed to control for unobserved time-invariant country heterogeneity, such as cultural, social, historical, geographic and other conditions that remained relatively constant within the study period. The results of the fixed effects models can be interpreted as the relationship between annual changes in income inequality and annual changes in health outcomes. Clustered sandwich estimators were used to allow for within-country correlation between error terms [32].

The model can be written as:

$$healthoutcome_{ij} = \alpha + \beta_1 Gini_{ij} + \beta_2 lngdp_{ij} + X_i + T_j + \gamma C_{ij}$$

where X_i is a vector of country fixed effects.

In the online supplementary material, supplementary analyses checking robustness of the results are reported, which include a) models allowing a different linear time trend and different effects of the country characteristics for the former Soviet countries (Supplementary Table 2); b) models allowing country-specific linear time trends by including the interaction terms between country dummies and year (Supplementary Table 2); c) analyses restricted to high-income countries within Europe (Supplementary Table 3); d) analyses using the Gini index based on gross instead of net household income (Supplementary Table 4); e) analyses sequentially replacing contemporaneous Gini indexes by indexes up to 10 years before the mortality outcomes ("lagged terms") (Supplementary Table 5); f) analyses using the Gini index from the LIS instead of the SWIID (Supplementary Table 6).

All regression analyses were performed in Stata 13.1.

RESULTS

Table 1 reports descriptive information on the Gini index, life expectancy and infant mortality in the period between 1987 and 2008 for each country (other outcomes are described in the Supplementary Table 1). In some countries, information on the Gini index only became available in more recent periods (e.g. Iceland, Cyprus, Albania, Bosnia Hercegovina, Malta, Montenegro and Serbia). The Gini index ranges from 15.77 (Slovakia in 1989) to 47.94 (Azerbaijan in 1997). The mean Gini index over the whole period was generally lower in the Nordic region, and higher in Britain and Ireland, and in countries of the former Soviet Union. Higher standard deviations of the Gini index reflect higher within-country variations over time, which particularly occurred in countries of the former Soviet Union and the Western Balkans. The mean life expectancy was lowest in former Soviet countries, followed by countries of the Western Balkans and Central and Eastern Europe. They were relatively similar to each other in the other four European regions. Infant mortality was highest in former Soviet countries, followed by countries of the Western Balkans and Central and Eastern Europe. Life expectancy of women was higher than life expectancy of men in all countries.

Figure 1 shows the trends of the Gini index experienced by the seven European regions over this period. Britain and Ireland and the Mediterranean countries maintained a high level of income inequality over time, while Nordic countries maintained a relatively low level of income inequality. Most regions experienced an increasing trend of income inequality between 1987 and 2008, with a substantial increase in former Soviet countries in the early 1990s.

Table 2 shows the results of linear regression analyses linking income inequality to life expectancy or cause-specific mortality by gender when all country-year observations were pooled together. Income inequality was significantly and negatively related to life expectancy for both men and women, indicating shorter life expectancies with larger income inequalities.

Table 1 Descriptive information on income inequality (GINI), life expectancy and infant mortality between 1987 and 2008 in 43 European countries

Country	Gini index[a]			Life expectancy at birth				All-cause mortality				Infant mortality		
	Years	N	Mean (sd)	Years	N	Mean [men]	Mean [women]	Years	N	Mean [men]	Mean [women]	Years	N	Mean
Nordic countries														
Denmark	1987–2008	22	23.37 (0.28)	1987–2006	20	73.81	78.90	1987–2006	20	993.57	656.05	1987–2008	22	5.55
Finland	1987–2008	22	23.25 (0.46)	1987–2008	22	73.55	80.98	1987–2008	22	996.25	552.38	1987–2008	22	4.15
Iceland	1992–2008	17	24.77 (0.60)	1987–2008	22	77.47	81.76	1987–2008	22	740.79	513.66	1987–2008	22	3.77
Norway	1987–2008	22	23.59 (0.20)	1987–2008	22	75.66	81.36	1987–2008	22	865.36	530.08	1987–2008	22	4.82
Sweden	1987–2008	22	23.14 (0.29)	1987–2008	22	76.82	81.89	1987–2008	22	795.45	506.83	1987–2008	22	4.15
Britain and Ireland														
Ireland	1987–2008	22	31.98 (0.32)	1987–2008	22	74.05	79.40	1987–2008	22	1005.51	640.92	1987–2008	22	6.13
United Kingdom	1987–2008	22	33.93 (0.29)	1987–2008	22	74.93	79.95	1987–2008	22	912.91	599.70	1987–2008	22	6.30
Continental Europe														
Austria	1987–2008	22	26.93 (0.45)	1987–2008	22	74.62	80.81	1987–2008	22	919.51	558.54	1987–2008	22	5.82
Belgium	1987–2008	22	25.00 (0.41)	1987–2005	15	73.68	80.31	1987–2005	15	991.75	573.11	1987–2008	22	5.97
Luxembourg	1987–2008	22	25.60 (0.35)	1987–2008	22	74.20	80.76	1987–2008	22	957.18	560.91	1987–2008	22	5.66
Germany (UnitedNations)	1987–2008	22	27.15 (0.20)	1987–2008	22	74.44	80.50	1990–2006	17	933.29	565.48	1987–2008	22	5.21
Switzerland	1987–2008	22	29.34 (0.44)	1987–2007	21	76.40	82.55	1987–2007	21	794.44	469.89	1987–2008	22	5.08
Netherlands	1987–2008	22	25.47 (0.28)	1987–2008	22	75.49	80.95	1987–2008	22	895.89	547.52	1987–2008	22	5.56

Table 1 Descriptive information on income inequality (GINI), life expectancy and infant mortality between 1987 and 2008 in 43 European countries (continued)

Country	Gini index[a]			Life expectancy at birth				All-cause mortality				Infant mortality		
Mediterranean countries														
Cyprus	1990–2008	19	25.77 (1.15)	1987–2008	14	76.36	80.70	1999–2008	10	752.23	519.37	1987–2008	22	7.83
France	1987–2008	22	28.45 (0.19)	1987–2008	22	75.13	82.93	1987–2008	22	850.94	453.08	1987–2008	22	5.27
Greece	1987–2008	22	33.40 (0.32)	1987–2008	22	75.70	80.64	1987–2008	22	836.24	576.05	1987–2008	22	6.97
Italy	1987–2008	22	32.91 (0.33)	1987–2008	20	75.66	81.92	1987–2008	20	849.39	501.79	1987–2008	22	6.01
Malta	2000–2008	9	27.97 (0.68)	1987–2008	22	75.27	79.87	1987–2008	22	911.89	619.11	1987–2008	22	8.87
Portugal	1987–2008	22	33.76 (0.69)	1987–2008	20	72.42	79.44	1987–2008	20	1046.94	623.32	1987–2008	22	7.60
Spain	1987–2008	22	32.26 (0.34)	1987–2008	22	75.34	82.37	1987–2008	22	848.94	478.96	1987–2008	22	5.56
Western Balkans														
Albania	1996–2008	13	30.25 (1.03)	1987–2004	17	71.46	77.89	1987–2004	17	1131.47	665.42	1987–2008	22	28.41
Bosnia-Hercegovina	1991–2007	17	32.72 (1.41)	1987–1991	5	69.51	75.19	-	-	-	-	1987–2008	22	16.79
Croatia	1987–2008	22	28.71 (0.69)	1987–2008	22	69.64	77.30	1987–2008	22	1326.29	775.23	1987–2008	22	8.62
Montenegro	2000–2008	9	31.22 (1.03)	1987–2008	20	72.24	78.16	1987–2008	20	1050.29	695.24	1987–2008	22	13.89
Serbia	2002–2008	7	31.14 (1.04)	1994–2008	13	69.98	75.28	1998–2008	11	1303.84	913.33	1987–2008	22	16.47
Slovenia	1987–2008	22	21.38 (0.56)	1987–2008	22	71.54	79.33	1987–2008	22	1153.23	639.76	1987–2008	22	6.00
TFYR Macedonia	1989–2008	20	33.56 (0.90)	1987–2003	14	70.30	75.04	1991–2003	13	1224.10	864.51	1987–2008	22	22.61
Central and Eastern Europe														
Bulgaria	1987–2008	22	26.18 (0.83)	1987–2008	22	68.34	75.30	1987–2008	22	1409.00	894.99	1987–2008	22	13.83
Czech Republic	1987–2008	22	23.96 (0.59)	1987–2008	22	70.75	77.72	1987–2008	22	1264.35	749.67	1987–2008	22	6.27
Hungary	1987–2008	22	28.30 (0.51)	1987–2008	22	66.94	75.64	1987–2008	22	1517.29	844.49	1987–2008	22	10.37
Poland	1987–2008	22	28.69 (0.41)	1987–2008	22	68.76	77.43	1987–2008	22	1346.58	738.32	1987–2008	22	11.30
Romania	1989–2008	20	27.68 (0.86)	1987–2008	22	67.10	74.22	1987–2008	22	1426.36	926.30	1987–2008	22	20.24
Slovakia	1987–2008	22	22.91 (0.82)	1987–2008	22	68.92	77.27	1987–2008	22	1353.97	753.51	1987–2008	22	9.50

Table 1 Descriptive information on income inequality (GINI), life expectancy and infant mortality between 1987 and 2008 in 43 European countries (continued)

Country	Gini index[a]			Life expectancy at birth				All-cause mortality				Infant mortality		
Former Soviet Union														
Armenia	1987–2008	22	34.60 (1.98)	1987–2008	21	67.07	73.31	1987–2008	18	1306.84	886.44	1987–2008	22	37.10
Azerbaijan	1987–2008	22	32.82 (2.74)	2007	1	71.33	76.25	1987–2007	19	1316.38	834.21	1987–2008	22	64.45
Belarus	1987–2008	22	25.69 (0.84)	1987–2007	20	64.09	75.02	1987–2007	20	1745.40	886.54	1987–2008	22	12.51
Estonia	1987–2008	22	32.37 (1.13)	1987–2008	22	65.05	76.06	1987–2008	22	1625.19	806.74	1987–2008	22	11.73
Georgia	1987–2008	22	39.73 (2.04)	1987–2006	20	66.56	74.10	1987–2001	14	1217.00	734.36	1987–2008	22	36.48
Latvia	1987–2008	22	29.94 (1.13)	1987–2008	22	64.14	75.18	1987–2008	22	1687.41	850.63	1987–2008	22	11.74
Lithuania	1987–2008	22	30.67 (0.95)	1987–2008	22	65.58	76.58	1987–2008	22	1516.67	760.50	1987–2008	22	10.49
Republic of Moldova	1987–2008	22	35.74 (1.53)	1987–2008	22	64.24	71.57	1987–2008	22	1771.57	1141.94	1987–2008	22	23.84
Russian Federation	1987–2008	22	37.36 (1.53)	1987–2008	22	60.60	72.85	1987–2008	22	1991.94	983.64	1987–2008	22	19.29
Ukraine	1987–2008	22	29.14 (1.69)	1987–2008	21	63.08	73.20	1987–2008	21	1792.63	959.86	1987–2008	22	15.62

Note:
a. N: number of available observations in the study period; SD: standard error of the Gini index within each country over time

Chapter 3

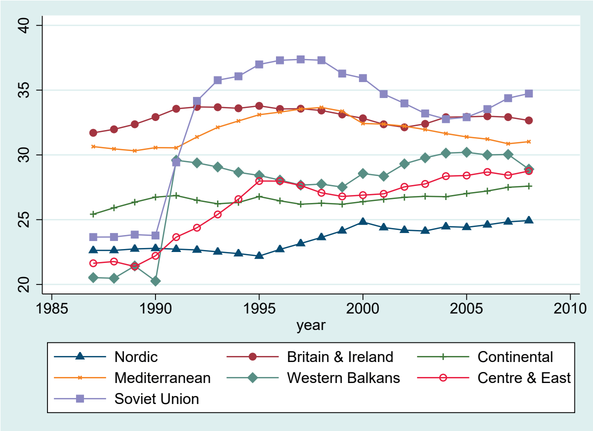


Figure 1 Trends in income inequality (mean GINI) for 7 European regions[a] (41 countries[b]), 1987–2008 (N=840), the SWIID database

Notes:

a. After 1990, United Kingdom experienced an increase of income inequality and Ireland experienced a decrease of income inequality. The line for Britain&Ireland is an average trend of these two countries.

b. 41 European countries were included in this graph, where Cyprus and TFYR Macedonia were excluded. This is because the Gini index of Cyprus was much lower than that of other Mediterranean countries and it was available from 1990. The inclusion of Cyprus would cause a sudden decrease of the Mediterranean average Gini index at the point of 1990. Similarly, the Gini index of TFYR Macedonia was much higher than that of other Western Balkan countries and it was available from 1989. The inclusion of TFYR Macedonia would cause a sudden increase of the Western Balkan average Gini index at the point of 1989.

This negative association was stronger for men than for women, and remained significant after additional adjustments for indicators of education, democracy, independence, armed conflicts and economic freedom. For cause-specific mortality, adjusted for GDP and time, income inequality was positively related to mortality from cerebrovascular disease (among women), all infectious disease, signs, symptoms and ill-defined conditions, and homicide (among men), and income inequality was also negatively related to some causes of mortality, such as all cancers and suicide. After further adjustment for indicators of education, democracy, independence, armed conflicts and economic freedom, positive associations with income inequality were found for almost all causes of mortality (except deaths from lung cancer among women, breast cancer, chronic liver diseases and suicide). The inverse association between income inequality and infant mortality was significant and remained significant after adjustment for the potential confounding variables.

Table 3 presents the results from fixed effects models linking changes in income inequality to changes in life expectancy or cause-specific mortality. All associations between income inequality and mortality indicators became insignificant, except for death from external causes and homicide among men, and all cancers among women. The significant results of homicide and all cancers disappeared after further adjustment of more country characteristics. In further analyses (Supplementary Table 2), the positive relation between income

Table 2 Linear regression coefficients of GINI from pooled cross-sectional analyses linking income inequality and population health measured by life expectancy and mortality, pooled 43 European countries 1987–2008

Outcomes	Men		Women	
	Model 1[a]	Model 2[b]	Model 1	Model 2
Life expectancy	**-0.0820**	**-0.2590**	**-0.0371**	**-0.0901**
	(-0.130, -0.034)	**(-0.317, -0.201)**	**(-0.065, -0.010)**	**(-0.122, -0.058)**
All causes	0.0001	**0.0128**	-0.0014	**0.0056**
	(-0.003, 0.004)	**(0.009, 0.017)**	(-0.004, 0.001)	**(0.003, 0.008)**
All circulatory disease	-0.0030	**0.0109**	-0.0020	**0.0110**
	(-0.007, 0.001)	**(0.005, 0.016)**	(-0.006, 0.002)	**(0.006, 0.016)**
Ischemic heart disease	0.0022	**0.0156**	0.0006	**0.0119**
	(-0.005, 0.009)	**(0.007, 0.024)**	(-0.007, 0.009)	**(0.003, 0.021)**
Cerebrovascular disease	0.0056	**0.0203**	**0.0098**	**0.0235**
	(-0.001, 0.012)	**(0.013, 0.027)**	**(0.004, 0.016)**	**(0.017, 0.030)**
All cancers	**-0.0066**	**0.0045**	**-0.0055**	**0.0102**
	(-0.010, -0.003)	**(0.002, 0.007)**	**(-0.010, -0.001)**	**(0.007, 0.014)**
Cancer of lung	-0.0058	**0.0148**	**-0.0197**	**-0.0213**
	(-0.012, 0.00002)	**(0.010, 0.020)**	**(-0.026, -0.013)**	**(-0.030, -0.012)**
Cancer of breast			0.0015	0.0032
			(-0.003, 0.006)	(-0.001, 0.007)
All infectious disease	**0.0349**	**0.0463**	**0.0205**	**0.0130**
	(0.026, 0.043)	**(0.036, 0.057)**	**(0.014, 0.027)**	**(0.006, 0.020)**
Chronic liver disease and cirrhosis	0.0003	0.0006	0.0119	0.0034
	(-0.013, 0.014)	(-0.021, 0.022)	(-0.001, 0.025)	(-0.014, 0.021)
All external causes	-0.0017	**0.0301**	**-0.0086**	**0.0162**
	(-0.011, 0.008)	**(0.020, 0.041)**	**(-0.017, -0.001)**	**(0.008, 0.024)**
Motor vehicle accidents	0.0053	**0.0381**	0.0010	**0.0293**
	(-0.004, 0.015)	**(0.029, 0.048)**	(-0.008, 0.010)	**(0.020, 0.038)**
Suicide	**-0.0375**	0.0070	**-0.0443**	**-0.0139**
	(-0.053, -0.022)	(-0.006, 0.020)	**(-0.057, -0.031)**	**(-0.026, -0.002)**
Signs, symptoms and ill-defined	**0.0249**	**0.0455**	**0.0234**	**0.0377**
	(0.008, 0.042)	**(0.21, 0.070)**	**(0.004, 0.042)**	**(0.011, 0.065)**
Homicide	**0.0175**	**0.0584**	-0.0089	**0.0256**
	(0.001, 0.034)	**(0.039, 0.078)**	(-0.024, 0.006)	**(0.008, 0.043)**
	Infant			
Infant mortality	**0.0216**	**0.0166**		
	(0.015, 0.028)	**(0.012, 0.021)**		

Notes: Significant results at the 95% confidence level are shown in bold.

a. Model 1 includes Gini index (95% confidence interval in parentheses, based on robust standard errors), year dummies and log(gdp).

b. Model 2 additionally adds democracy index, education, independence, armed conflict and economic freedom.

Chapter 3

Table 3 Linear regression coefficients of GINI from fixed effects models linking income inequality and population health measured by life expectancy and mortality, 43 European countries 1987–2008

Outcomes	**Men**		**Women**	
	Model 1[a]	Model 2[b]	Model 1	Model 2
Life expectancy	-0.0754	-0.0811	-0.0049	-0.0361
	(-0.178, 0.027)	(-0.174, 0.012)	(-0.071, 0.061)	(-0.077, 0.005)
All causes	0.0054	0.0044	0.00002	0.0020
	(-0.0005, 0.011)	(-0.001, 0.010)	(-0.005, 0.005)	(-0.002, 0.006)
All circulatory diseases	0.0063	0.0022	0.0022	0.0010
	(-0.003, 0.015)	(-0.005, 0.010)	(-0.006, 0.011)	(-0.004, 0.006)
Ischemic heart disease	0.0060	-0.0009	0.0019	-0.0044
	(-0.003, 0.015)	(-0.011, 0.009)	(-0.007, 0.011)	(-0.015, 0.006)
Cerebrovascular disease	0.0061	0.0036	0.0029	0.0034
	(-0.004, 0.017)	(-0.005, 0.013)	(-0.008, 0.014)	(-0.005, 0.011)
All cancers	0.0035	0.0032	**0.0051**	0.0036
	(-0.001, 0.008)	(-0.0001, 0.007)	**(0.0001, 0.010)**	(-0.001, 0.009)
Cancer of lung	0.0030	0.0018	-0.0068	-0.0052
	(-0.002, 0.008)	(-0.006, 0.009)	(-0.016, 0.003)	(-0.016, 0.005)
Cancer of breast			0.0061	0.0035
			(-0.003, 0.015)	(-0.003, 0.010)
All infectious diseases	0.0068	0.0235	0.00005	0.0126
	(-0.008, 0.022)	(-0.009, 0.056)	(-0.015, 0.015)	(-0.013, 0.038)
Chronic liver disease and cirrhosis	0.0153	0.0086	0.0153	0.0015
	(-0.004, 0.035)	(-0.015, 0.032)	(-0.007, 0.037)	(-0.021, 0.024)
All external causes	**0.0143**	**0.0143**	0.0078	**0.0158**
	(0.0002, 0.028)	**(0.001, 0.028)**	(-0.007, 0.023)	**(0.001, 0.031)**
Motor vehicle accidents	0.0091	0.0048	0.0109	0.0054
	(-0.007, 0.025)	(-0.009, 0.018)	(-0.004, 0.026)	(-0.009, 0.020)
Suicide	0.0051	0.0104	-0.0009	0.0127
	(-0.007, 0.017)	(-0.0003, 0.021)	(-0.015, 0.013)	(-0.001, 0.027)
Signs, symptoms and ill-defined	0.0219	0.0338	0.0078	0.0290
	(-0.011, 0.055)	(-0.003, 0.070)	(-0.029, 0.045)	(-0.007, 0.065)
Homicide	**0.0285**	0.0171	0.0147	0.0066
	(0.008, 0.049)	(-0.002, 0.036)	(-0.004, 0.033)	(-0.009, 0.022)
	Infant			
Infant mortality	-0.0023	-0.0498		
	(-0.007, 0.002)	(-0.130, 0.031)		

Notes: Significant results at the 95% confidence level are shown in bold.
a. Model 1 includes Gini index (95% confidence interval in parentheses, based on clustered standard errors), year dummies, log(gdp) and country fixed effects.
b. Model 2 additionally adds democracy index, education, independence, armed conflict and economic freedom.

inequality and death from external causes appeared not robust to variations in model specifications.

Essentially similar results were obtained in supplementary analyses for robustness checks: allowing a different time trend and interactive effects of the country characteristics for the former Soviet countries, allowing country-specific linear time trends, restricting the analyses to high-income countries, and using gross income Gini index. Moreover, we also introduced up to 10-year lags between the income inequality and life expectancy or infant mortality into the fixed effects models. None of the lagged terms of Gini index was significant. Simultaneously controlling all preceding income inequalities [33] gave essentially similar results (available upon request). Similar analyses were conducted using interpolated LIS data, which also have good quality but much less country-year observations. Again, statistically significant associations between income inequality and mortality were rare. Additionally controlling for the unemployment rate did not change our main findings (results not shown). These checks indicate that our main findings are robust against different model specifications, sample changes, using a Gini index based on gross household income and using another dataset for income inequality.

DISCUSSION

Summary of main findings

Significant associations between income inequality and many mortality indicators were found in pooled cross-sectional regressions, indicating higher mortality in countries with larger income inequalities. However, once the country fixed effects were added, all associations between income inequality and mortality indicators became insignificant, except for all external causes and homicide among men, and all cancers among women. The significant results of homicide and all cancers disappeared after further adjustment for indicators of democracy, average years of schooling, transition to national independence, armed conflicts, and economic freedom.

Study limitations

To identify the link between income inequality and mortality, we adjusted for an array of country characteristics. In order to be confounding variables, these factors should be related to both income inequality and mortality. To the extent however, that these indicators result from income inequality, and therefore should be considered as mediating variables, we may have over-controlled the analyses and thereby removed part of the association between income inequality on mortality. Whether the country characteristic should be seen as a confounder or mediator is not easy to determine, especially for education [34]. On the one hand, investing in education could be a strategy to reduce income inequality within a country [30], making it a potential confounder. On the other hand, high levels of income inequality and the associated underinvestment in public resources might in the long run lead to lower levels of education [35], which makes it a potential mediator. However, most associations between income inequality and mortality indicators were insignificant in the fixed effects models even without controlling for education and other country time-variant characteristics. Thus, the general conclusions are not threatened by the potential problem of "over-controlling".

In the main analyses, we related annual changes in income inequality to simultaneous annual changes in mortality. We also investigated the effects of income inequality in a specific year on mortality up to 10 years later, which produced similar results. This approach may be insufficient to fully capture the cumulative impact of a history of large income inequalities on mortality [21, 22, 36]. Future research, using data over even longer time-periods than available in our study, are necessary to explore the effects of long-term exposure to income inequality.

Our analysis was limited to European countries. This reduced the potential for confounding as country-level confounding variables can be expected to be more similar in Europe than in a global context. It also produced results that are relevant for policy makers in European countries, who are likely to be more concerned with the effect of income inequality as observed within the range of variation prevailing in a European context, than with the more extreme values of income inequality observed elsewhere in the world. However, it is important to note that our results cannot be generalized beyond this smaller range of variation, and that analyses including countries outside Europe with larger increases in income inequality over time may lead to different conclusions.

Interpretation

Our findings are in line with previous studies that found negative cross-sectional associations between income inequality and population health [1, 15, 16], and with some existing studies using fixed effects models where no significant effects were found [6, 19, 21]. We further strengthened the evidence however, because our results were obtained in a study in which we focused on a large array of European countries over a relatively long period of time, used better data on income inequalities and a set of disease-specific mortality indicators, considered a larger set of potential confounding variables, and used a country fixed effects approach. The only significant result in our fixed effects analysis, after adjustment for country characteristics, was the association between income inequality and external causes mortality. This cause of death group is strongly related to individual socioeconomic status [37, 38]. However, the results for all external causes mortality were not robust to variations in model specifications in further analyses (Supplementary Table 2).

Differences between results from pooled cross-sectional analyses and fixed effects models indicate that the observed association between income inequality and mortality is likely to result from confounding, and that income inequality as such is not a driving force of poor population health. We can only speculate which country characteristics might be responsible for the disappearance of the effect, and suspect that these are historical, social or cultural factors that are associated with both the hierarchical nature of societies, as indicated by income inequality, and the health of their populations. Unfortunately, many of these factors are not available in international databases. Further research is necessary to find appropriate measures for the relevant country characteristics and test their effects on the association between income inequality and mortality.

One underlying factor determining both income inequality and mortality could be social and health policies that vary across countries and are persistent over time. For example, poverty reduction policies such as minimum wage, disability allowances and return to work programs can reduce income inequality and simultaneously improve average population health by improving health of the poorest part of the population. Besides these, health care programs such as smoking cessation strategies, maternal education programs and cancer screening may also play roles since they tend to cluster in countries with strong redistribution policies, although without having a direct impact on income inequality [21]. The implementation of all these policies has varied between European countries [39], which could have produced a "spurious" association between income inequality and mortality. Other responsible factors could be some cultural and historical elements of a country, e.g. egalitarianism (importance of transcending self-interest and promoting the welfare of others), power distance (extent to which the less powerful accept that power is distributed unequally) and ethnic heterogeneity, which are potentially important determinants of population health [40, 41], and at the same time could be related to income inequality [42, 43]. The disappearance of the association between income inequality and health when moving to fixed effects models could be the result of controlling for these country heterogeneities.

It has been noted that the most consistent evidence for an adverse effect of income inequality on population health derives from within-country differences in the United States or some other countries with comparable or even larger income inequalities [3, 36]. The studies from countries having more equal income distribution outside of Europe, e.g. Canada, Australia and Japan, often produced insignificant findings [44, 45]. Therefore, one possible explanation for not finding significant relationships in our study is that most European countries in our sample are more egalitarian than the United States. It has been suggested that there is a "threshold effect", implying the existence of a threshold of income inequality above which adverse impacts on health begin to emerge [46]. However, this appears to only partly explain our findings, since restricting the analysis to the 18 countries with a mean Gini index larger than 30 (a potential threshold value suggested in the literature [46]) still produced insignificant results (results not shown). Another possible explanation for not finding significant relationships is that income inequality may be less strongly associated with the social distribution of major risk factors in Europe. For example, the delay of the epidemiologic transition and the more egalitarian social distribution of healthy "Mediterranean diets" in southern Europe make lower income less a risk factor for cardiovascular disease mortality in these countries [3, 47, 48]. Meanwhile, the well-developed welfare system in Europe, especially in some northern and continental European countries, may help to buffer the adverse effect on mortality of being poor [49]. A deeper exploration of why income inequality does not have more effect in Europe is needed.

Conclusions

Within Europe, cross-sectional associations between income inequality and mortality probably result from confounding. Fixed effect models which remove time-invariant country heterogeneity suggest that there is no statistically significant relation between income inequality and population health measured by life expectancy and cause-specific mortality in European countries between 1987 and 2008. Although reducing income inequality may be important for creating equality of opportunity and for the reduction of health inequalities, it has a limited role for reducing average mortality in Europe.

REFERENCES

1. Wilkinson, R.G., *Income distribution and life expectancy.* BMJ, 1992. **304**(6820): p. 165–8.
2. Deaton, A., *Commentary: The convoluted story of international studies of inequality and health.* Int J Epidemiol, 2002. **31**(3): p. 546–9.
3. Lynch, J.W., et al., *Is income inequality a determinant of population health? Part 1. A systematic review.* Milbank Q, 2004. **82**(1): p. 5–99.
4. Wilkinson, R.G. and K.E. Pickett, *Income inequality and population health: a review and explanation of the evidence.* Soc Sci Med, 2006. **62**(7): p. 1768–84.
5. Judge, K., J.A. Mulligan, and M. Benzeval, *Income inequality and population health.* Soc Sci Med, 1998. **46**(4–5): p. 567–79.
6. Beckfield, J., *Does income inequality harm health? New cross-national evidence.* J Health Soc Behav, 2004. **45**(3): p. 231–48.
7. Leigh, A. and C. Jencks, *Inequality and mortality: long-run evidence from a panel of countries.* J Health Econ, 2007. **26**(1): p. 1–24.
8. LIS, *Luxembourg Income Study (LIS) Data Access*. 2008: http://www.lisdatacenter.org/data-access/ (accessed 22 April, 2014).
9. Lynch, J.W., et al., *Income inequality, the psychosocial environment, and health: comparisons of wealthy nations.* Lancet, 2001. **358**(9277): p. 194–200.
10. Leigh, A., C. Jencks, and T.M. Smeeding, *Health and Economic Inequality.* In Salverda W, Nolan B, Smeeding TM, (Eds.), The Oxford Handbook of Economic Inequality. 2009, Oxford: Oxford University Press.
11. Mellor, J.M. and J. Milyo, *Reexamining the evidence of an ecological association between income inequality and health.* J Health Polit Policy Law, 2001. **26**(3): p. 487–522.
12. Gravelle, H., J. Wildman, and M. Sutton, *Income, income inequality and health: what can we learn from aggregate data?* Soc Sci Med, 2002. **54**(4): p. 577–89.
13. Atkinson, A.B. and A. Brandolini, *Promise and Pitfalls in the Use of `Secondary' Data-Sets: Income Inequality in OECD Countries as a Case Study.* Journal of Economic Literature, 2001. **39(3)**: p. 771–799.
14. Solt, F., *Standardizing the World Income Inequality Database.* Social Science Quarterly, 2009. **90**(2): p. 231–242.
15. De Vogli, R., et al., *Has the relation between income inequality and life expectancy disappeared? Evidence from Italy and top industrialised countries.* J Epidemiol Community Health, 2005. **59**(2): p. 158–62.
16. Rodgers, G.B., *Income and inequality as determinants of mortality: an international cross-section analysis.* Population Studies, 1979. **33**: p. 343–51.
17. Ram, R., *Further examination of the cross-country association between income inequality and population health.* Soc Sci Med, 2006. **62**(3): p. 779–91.
18. Macinko, J.A., et al., *Income inequality and health: a critical review of the literature.* Med Care Res Rev, 2003. **60**(4): p. 407–52.
19. Babones, S.J., *Income inequality and population health: correlation and causality.* Soc Sci Med, 2008. **66**(7): p. 1614–26.
20. Shkolnikov, V.M., et al., *Losses of expected lifetime in the United States and other developed countries: methods and empirical analyses.* Demography, 2011. **48**(1): p. 211–39.
21. Avendano, M., *Correlation or causation? Income inequality and infant mortality in fixed effects models in the period 1960–2008 in 34 OECD countries.* Soc Sci Med, 2012. **75**(4): p. 754–60.
22. Clarkwest, A., *Neo-materialist theory and the temporal relationship between income inequality and longevity change.* Soc Sci Med, 2008. **66**(9): p. 1871–81.

23. Pop, I.A., E. van Ingen, and W. van Oorschot, *Inequality, wealth and health: is decreasing income inequality the key to create healthier societies?* Social Indicators Research, 2013. **113**(3): p. 1025–1043.
24. Torre, R. and M. Myrskyla, *Income inequality and population health: an analysis of panel data for 21 developed countries, 1975–2006*. Popul Stud (Camb), 2014. **68**(1): p. 1–13.
25. Bobak, M., et al., *Societal characteristics and health in the former communist countries of Central and Eastern Europe and the former Soviet Union: a multilevel analysis.* Journal of Epidemiology and Community Health, 2007. **61**(11): p. 990–996.
26. Zagorski, K., et al., *Does National Income Inequality Affect Individuals' Quality of Life in Europe? Inequality, Happiness, Finances, and Health*. Social Indicators Research, 2014. **117**(3): p. 1089–1110.
27. UNU-WIDER, *World Income Inequality Database, Version 2.0c*. 2008: Available online at http://www.wider.unu.edu/research/Database/en_GB/database/ (accessed 17 March 2015).
28. Mackenbach, J.P., Y. Hu, and C.W. Looman, *Democratization and life expectancy in Europe, 1960–2008*. Soc Sci Med, 2013. **93**: p. 166–75.
29. Reuveny, R. and Q. Li, *Economic openness, democracy, and income inequality - An empirical analysis.* Comparative Political Studies, 2003. **36**(5): p. 575–601.
30. Sylwester, K., *Can education expenditures reduce income inequality?* Economics of Education Review, 2002. **21**(1): p. 43–52.
31. Carter, J.R., *An empirical note on economic freedom and income inequality.* Public Choice, 2007. **130**(1–2): p. 163–177.
32. Wooldridge, J.M., *Econometric analysis of cross section and panel data (second edition)*. 2010: MIT Press.
33. Zheng, H., *Do people die from income inequality of a decade ago?* Soc Sci Med, 2012. **75**(1): p. 36–45.
34. Mackenbach, J.P., *Income inequality and population health*. BMJ, 2002. **324**(7328): p. 1–2.
35. Lynch, J.W., et al., *Income inequality and mortality: importance to health of individual income, psychosocial environment, or material conditions.* BMJ, 2000. **320**(7243): p. 1200–4.
36. Kawachi, I. and S.V. Subramanian, *Income inequality*. In Berkman, L. F., Kawachi, I., Glymour, M. M. (Eds.), Social Epidemiology (pp. 126–151). 2014, United States of America: Oxford University Press.
37. Middelkoop, B.J., et al., *Urban cause-specific socioeconomic mortality differences. Which causes of death contribute most?* Int J Epidemiol, 2001. **30**(2): p. 240–7.
38. Leinsalu, M., et al., *Educational inequalities in mortality in four Eastern European countries: divergence in trends during the post-communist transition from 1990 to 2000.* Int J Epidemiol, 2009. **38**(2): p. 512–25.
39. Mackenbach, J.P. and M. McKee, *Successes and failures of health policy in Europe: Four decades of diverging trends and converging challenges*. 2013, Buckingham: Open University Press.
40. Alesina, A., R. Baqir, and W. Easterly, *Public goods and ethnic divisions.* Quarterly Journal of Economics, 1999. **114**(4): p. 1243–1284.
41. Mackenbach, J.P., *Cultural values and population health: a quantitative analysis of variations in cultural values, health behaviours and health outcomes among 42 European countries.* Health Place, 2014. **28**: p. 116–32.
42. Hofstede, G., *Dimensionlizing cultures: the Hofstede model in context*, in *Online Readings in Psychology and Culture*. 2011: http://scholarworks.gvsu.edu/cgi/viewcontent.cgi?article=1014&context=orpc (accessed 01 December 2014).
43. Kraal, K., J. Roosblad, and J. Wrench, *Equal opportunities and ethnic inequality in European labour markets: discrimination, gender and policies of diversity*. 2014: Amsterdam University Press.
44. Ross, N.A., et al., *Relation between income inequality and mortality in Canada and in the United States: cross sectional assessment using census data and vital statistics.* BMJ, 2000. **320**(7239): p. 898–902.
45. Ross, N.A., et al., *Metropolitan income inequality and working-age mortality: a cross-sectional analysis using comparable data from five countries.* J Urban Health, 2005. **82**(1): p. 101–10.

46. Kondo, N., et al., *Income inequality, mortality, and self rated health: meta-analysis of multilevel studies.* BMJ, 2009. **339**: p. b4471.
47. Kulhanova, I., et al., *Why does Spain have smaller inequalities in mortality? An exploration of potential explanations.* European Journal of Public Health, 2014. **24**(3): p. 370–7.
48. Mackenbach, J.P., et al., *Socioeconomic inequalities in health in 22 European countries.* N Engl J Med, 2008. **358**(23): p. 2468–81.
49. Zambon, A., et al., *Do welfare regimes mediate the effect of socioeconomic position on health in adolescence? A cross-national comparison in Europe, North America, and Israel.* International Journal of Health Services, 2006. **36**(2): p. 309–329.

SUPPLEMENTARY MATERIAL

Supplementary Table 1 Descriptive statistics of the other variables used in the analysis

	(1)	(2)	(3)	(4)	(5)
VARIABLES	N	mean	sd	min	max
democracy	932	6.668	5.494	-9	10
gdp	934	12,136	7,997	1,632	53,962
education	814	9.632	1.323	5.735	13.09
conflict	946	0.122	0.513	0	3
independence	946	0.079	0.270	0	1
economic free	724	6.609	1.219	2.775	8.390
male_all	848	1,181	354.6	595.1	2,382
male_cir	825	554.1	236.0	165.7	1,184
male_isch	824	276.8	168.4	53.28	737.6
male_cere	825	137.2	75.13	30.92	379.4
male_neo	825	183.0	33.33	75.22	278.0
male_lung	824	69.46	20.80	16.95	120.4
male_infe	824	13.37	9.832	1.170	61.07
male_liver	748	26.97	22.97	0	128.4
male_exte	825	110.6	77.88	26.48	459.9
male_traf	793	20.14	10.51	1.430	68.79
male_suic	818	26.10	17.55	1.160	89.27
male_sign	799	36.08	42.95	0.340	346.7
male_homi	827	6.393	10.74	0	106.4
fe_all	848	697.5	181.7	380.5	1,358
fe_cir	825	362.7	160.5	96.43	758.9
fe_isch	824	150.5	107.1	12.80	517.4
fe_cere	825	109.7	57.98	23.42	277.0
fe_neo	825	249.3	52.22	93.85	391.1
fe_resp	824	14.51	8.448	3.240	44.54
fe_brea	825	25.95	6.601	6.020	53.99
fe_infe	824	5.657	2.725	0.300	19.62
fe_liver	748	11.59	14.87	0	105.1
fe_exte	825	34.21	25.95	7.890	611.1
fe_traf	793	5.762	2.703	0	16.11
fe_suic	818	6.842	3.835	0	23.18
fe_sign	799	23.52	34.57	0.210	291.5
fe_homi	827	2.019	2.391	0	14.33

Supplementary Table 2 Linear regression coefficients of GINI from fixed effects models (different specifications) linking income inequality and population health measured by life expectancy and mortality, 43 European countries 1987–2008

Outcomes	Men			Women		
	Model 1[a]	Model 2[b]	Model 3[c]	Model 1	Model 2	Model 3
Life expectancy	-0.0194	-0.0371	-0.0281	0.0136	-0.0169	0.0021
	(-0.099, 0.060)	(-0.088, 0.014)	(-0.113, 0.057)	(-0.040, 0.067)	(-0.048, 0.014)	(-0.068, 0.072)
All causes	0.0027	0.0021	0.0008	-0.0018	0.0006	-0.0007
	(-0.002, 0.007)	(-0.001, 0.006)	(-0.003, 0.004)	(-0.005, 0.001)	(-0.002, 0.004)	(-0.004, 0.002)
All circulatory diseases	0.0020	0.0014	-0.0006	-0.001	0.0014	-0.0012
	(-0.004, 0.007)	(-0.004, 0.007)	(-0.003, 0.002)	(-0.006, 0.004)	(-0.003, 0.005)	(-0.004, 0.002)
Ischemic heart disease	0.0025	0.0002	-0.0005	0.0009	-0.0014	-0.0024
	(-0.006, 0.011)	(-0.009, 0.010)	(-0.005, 0.004)	(-0.008, 0.010)	(-0.012, 0.009)	(-0.009, 0.004)
Cerebrovascular disease	0.0021	0.0011	-0.0008	-0.0003	0.0015	-0.0009
	(-0.005, 0.009)	(-0.010, 0.012)	(-0.006, 0.004)	(-0.008, 0.008)	(-0.008, 0.011)	(-0.006, 0.004)
All cancers	0.0028	0.0036	-0.0019	0.0043	0.0044	-0.0018
	(-0.001, 0.007)	(-0.001, 0.008)	(-0.006, 0.003)	(-0.0002, 0.009)	(-0.001, 0.010)	(-0.006, 0.003)
Cancer of lung	0.0029	0.0039	-0.00317	-0.0027	-0.0016	-0.0045
	(-0.002, 0.008)	(-0.004, 0.012)	(-0.009, 0.003)	(-0.008, 0.003)	(-0.008, 0.005)	(-0.012, 0.003)
Cancer of breast				0.0033	0.0015	-0.0015
				(-0.002, 0.009)	(-0.006, 0.009)	(-0.008, 0.005)
All infectious diseases	0.0058	0.0228	-0.0018	0.0021	0.0119	-0.0008
	(-0.010, 0.021)	(-0.006, 0.052)	(-0.011, 0.007)	(-0.012, 0.017)	(-0.012, 0.036)	(-0.010, 0.009)
Chronic liver disease and cirrhosis	0.0092	0.0075	0.0043	0.0077	-0.0019	0.0056
	(-0.005, 0.024)	(-0.017, 0.032)	(-0.006, 0.015)	(-0.007, 0.022)	(-0.025, 0.022)	(-0.004, 0.016)

Supplementary Table 2 Linear regression coefficients of GINI from fixed effects models (different specifications) linking income inequality and population health measured by life expectancy and mortality, 43 European countries 1987–2008 (continued)

Outcomes		**Men**			**Women**	
All external causes	0.0122	0.0069	0.0065	0.0059	**0.0105**	0.0004
	(-0.003, 0.027)	(-0.002, 0.016)	(-0.006, 0.019)	(-0.009, 0.021)	**(0.002, 0.019)**	(-0.013, 0.013)
Motor vehicle accidents	0.0123	0.0046	0.0020	0.0109	0.0053	-0.0009
	(-0.005, 0.029)	(-0.009, 0.018)	(-0.011, 0.015)	(-0.006, 0.028)	(-0.008, 0.019)	(-0.013, 0.012)
Suicide	0.0040	0.0057	-0.0101	-0.0003	0.0107	-0.0147
	(-0.007, 0.015)	(-0.005, 0.016)	(-0.030, 0.010)	(-0.014, 0.014)	(-0.004, 0.025)	(-0.039, 0.009)
Signs, symptoms and ill-defined	0.0193	0.0251	0.0112	0.0082	0.0207	0.0146
	(-0.012, 0.050)	(-0.013, 0.063)	(-0.011. 0.033)	(-0.027, 0.044)	(-0.017, 0.059)	(-0.012, 0.041)
Homicide	**0.0257**	0.0087	0.0038	0.0130	0.0006	-0.0085
	(0.002, 0.049)	(-0.009, 0.026)	(-0.013, 0.020)	(-0.008, 0.034)	(-0.012, 0.014)	(-0.024, 0.007)
		Infant				
Infant mortality	-0.0027	-0.0043	0.0005			
	(-0.007, 0.002)	(-0.012, 0.004)	(-0.003, 0.004)			

Notes: Significant results at the 95% confidence level are shown in bold.

a. Model 1 includes Gini index (95% confidence interval in parentheses, based on clustered standard errors), year dummies, log(gdp), year*soviet dummy and country fixed effects. Soviet dummy is 1 for Armenia, Azerbaijan, Belarus, Estonia, Georgia, Latvia, Lithuania, Republic of Moldova, Russian Federation and Ukraine, and soviet dummy is 0 for all other countries.

b. Model 2 additionally adds democracy index, education, independence, armed conflict, economic freedom, and the interactions between the soviet dummy and the adjusted variables.

c. Model 3 includes Gini index, year dummies, country fixed effects, log(gdp) and country-specific linear time trends.

Supplementary Table 3 Linear regression coefficients of GINI from fixed effects models linking income inequality and population health measured by life expectancy and mortality, 19 high-income European countries[a] 1987–2008

Outcomes	Men		Women	
	Model 1[b]	Model 2[c]	Model 1	Model 2
Life expectancy	0.0055	0.0105	0.0414	0.0406
	(-0.051, 0.061)	(-0.038, 0.059)	(-0.011, 0.094)	(-0.014, 0.095)
All causes	-0.0004	-0.0008	-0.0038	-0.0038
	(-0.006, 0.005)	(-0.005, 0.003)	(-0.009, 0.002)	(-0.009, 0.002)
All circulatory diseases	-0.0005	-0.0008	-0.0007	-0.0006
	(-0.008, 0.007)	(-0.007, 0.005)	(-0.008, 0.006)	(-0.008, 0.006)
Ischemic heart disease	0.0034	0.0021	0.0036	0.0031
	(-0.013, 0.020)	(-0.016, 0.020)	(-0.016, 0.023)	(-0.018, 0.024)
Cerebrovascular disease	-0.0061	-0.0064	-0.0067	-0.0069
	(-0.022, 0.010)	(-0.023, 0.010)	(-0.021, 0.007)	(-0.022, 0.008)
All cancers	0.0032	0.0027	0.0045	0.0037
	(-0.006, 0.012)	(-0.004, 0.009)	(-0.007, 0.016)	(-0.004, 0.011)
Cancer of lung	0.0039	0.0033	-0.0079	-0.0058
	(-0.016, 0.024)	(-0.012, 0.018)	(-0.021, 0.005)	(-0.015, 0.004)
Cancer of breast			-0.0003	-0.0003
			(-0.016, 0.016)	(-0.015, 0.014)
All infectious diseases	0.0481	0.0483	0.0304	0.0317
	(-0.003, 0.099)	(-0.005, 0.102)	(-0.016, 0.077)	(-0.019, 0.082)
Chronic liver disease and cirrhosis	0.0049	0.0037	-0.0103	-0.0101
	(-0.042, 0.052)	(-0.039, 0.047)	(-0.057, 0.037)	(-0.054, 0.033)
All external causes	0.0058	0.0054	0.0148	**0.0148**
	(-0.009, 0.021)	(-0.008, 0.019)	(-0.0002, 0.030)	**(0.003, 0.027)**
Motor vehicle accidents	0.0134	0.0126	0.0142	0.0132
	(-0.012, 0.039)	(-0.013, 0.038)	(-0.011, 0.039)	(-0.012, 0.039)
Suicide	0.0023	0.0011	0.0139	0.0137
	(-0.019, 0.023)	(-0.019, 0.021)	(-0.020, 0.048)	(-0.019, 0.046)
Signs, symptoms and ill-defined	0.0074	0.0046	-0.0032	-0.0055
	(-0.051, 0.066)	(-0.058, 0.067)	(-0.062, 0.055)	(-0.067, 0.056)
Homicide	-0.0076	-0.0068	0.0023	0.0033
	(-0.037, 0.021)	(-0.035, 0.021)	(-0.021, 0.026)	(-0.022, 0.028)
	Infant			
Infant mortality	**-0.0170**	**-0.0172**		
	(-0.033, -0.002)	**(-0.031, -0.003)**		

Notes: Significant results at the 95% confidence level are shown in bold.
a. Countries included in the analysis are Austria, Belgium, Denmark, Finland, France, Germany, Greece, Iceland, Ireland, Italy, Luxembourg, Malta, Netherlands, Norway, Portugal, Spain, Sweden, Switzerland and United Kingdom.
b. Model 1 includes Gini index (95% confidence interval in parentheses, based on clustered standard errors), year dummies, log(gdp) and country fixed effects.
c. Model 2 additionally adds democracy index, education, independence, armed conflict and economic freedom.

Supplementary Table 4 Linear regression coefficients of GINI (based on gross income) from fixed effects models linking income inequality and population health measured by life expectancy and mortality, 43 European countries 1987–2008

Outcomes	Men		Women	
	Model 1[a]	Model 2[b]	Model 1	Model 2
Life expectancy	-0.0185	-0.0283	0.0055	-0.0143
	(-0.086, 0.049)	(-0.077, 0.021)	(-0.036, 0.047)	(-0.039, 0.010)
All causes	0.0017	0.0017	-0.0006	0.0011
	(-0.002, 0.006)	(-0.002, 0.005)	(-0.004, 0.002)	(-0.001, 0.004)
All circulatory diseases	0.0009	0.0002	-0.0006	0.0003
	(-0.005, 0.007)	(-0.005, 0.005)	(-0.006, 0.004)	(-0.003, 0.004)
Ischemic heart disease	0.0007	-0.0003	-0.0001	-0.0009
	(-0.006, 0.007)	(-0.007, 0.007)	(-0.007, 0.007)	(-0.009, 0.007)
Cerebrovascular disease	0.0015	0.0030	0.0006	0.0036
	(-0.006, 0.009)	(-0.004, 0.010)	(-0.007, 0.008)	(-0.002, 0.009)
All cancers	0.0017	0.0021	0.0026	0.0025
	(-0.001, 0.005)	(-0.001, 0.005)	(-0.001, 0.006)	(-0.0009, 0.006)
Cancer of lung	0.0028	0.0036	-0.0010	0.0006
	(-0.001, 0.007)	(-0.001, 0.008)	(-0.007, 0.005)	(-0.006, 0.007)
Cancer of breast			0.0026	0.0020
			(-0.003, 0.008)	(-0.002, 0.006)
All infectious diseases	0.0046	0.0117	0.0030	0.0076
	(-0.007, 0.017)	(-0.007, 0.030)	(-0.008, 0.014)	(-0.007, 0.022)
Chronic liver disease and cirrhosis	0.0085	0.0091	0.0052	0.0004
	(-0.004, 0.021)	(-0.007, 0.026)	(-0.007, 0.017)	(-0.013, 0.014)
All external causes	0.0066	0.0054	0.0046	**0.0089**
	(-0.004, 0.018)	(-0.002, 0.013)	(-0.006, 0.015)	**(0.0010, 0.017)**
Motor vehicle accidents	0.0043	0.0024	0.0046	0.0033
	(-0.008, 0.016)	(-0.007, 0.012)	(-0.007, 0.016)	(-0.006, 0.013)
Suicide	0.0030	0.0060	0.0030	**0.0110**
	(-0.006, 0.012)	(-0.002, 0.014)	(-0.008, 0.014)	**(0.0002, 0.022)**
Signs, symptoms and ill-defined	0.0150	0.0183	0.0071	0.0163
	(-0.012, 0.042)	(-0.011, 0.047)	(-0.023, 0.037)	(-0.014, 0.047)
Homicide	0.0161	0.0091	0.0095	0.0046
	(-0.003, 0.035)	(-0.005, 0.023)	(-0.007, 0.026)	(-0.006, 0.015)
	Infant			
Infant mortality	-0.0029	-0.0536		
	(-0.007, 0.001)	(-0.119, 0.012)		

Notes: Significant results at the 95% confidence level are shown in bold.
a. Model 1 includes Gini index based on gross income (95% confidence interval in parentheses, based on clustered standard errors), year dummies, log(gdp) and country fixed effects.
b. Model 2 additionally adds democracy index, education, independence, armed conflict and economic freedom.

Supplementary Table 5 Linear regression coefficients of lagged GINI from fixed effects models linking lagged terms of income inequality and population health measured by life expectancy and infant mortality, 43 European countries 1987–2008

	Male life expectancy		**Female life expectancy**		**Infant mortality**	
	Model 1[a]	Model 2[b]	Model 1	Model 2	Model 1	Model 2
lag-1 year	-0.0684	-0.0692	-0.0046	-0.0336	-0.0033	-0.0033
	(-0.167, 0.030)	(-0.164, 0.025)	(-0.070, 0.061)	(-0.077, 0.0099)	(-0.009, 0.002)	(-0.011, 0.004)
lag-2 years	-0.0613	-0.0493	-0.0120	-0.0250	-0.0035	-0.0032
	(-0.136, 0.014)	(-0.137, 0.038)	(-0.054, 0.030)	(-0.071, 0.0213)	(-0.009, 0.002)	(-0.011, 0.005)
lag-3 years	-0.0346	-0.00496	-0.0014	0.0028	-0.0031	-0.0033
	(-0.092, 0.023)	(-0.072, 0.062)	(-0.035, 0.032)	(-0.038, 0.044)	(-0.009, 0.0024)	(-0.011, 0.005)
lag-4 years	-0.0180	0.0366	0.00045	0.0246	-0.0026	-0.0032
	(-0.062, 0.026)	(-0.020, 0.093)	(-0.026, 0.026)	(-0.013, 0.063)	(-0.008, 0.003)	(-0.011, 0.004)
lag-5 years	-0.0102	0.0235	0.0007	0.0168	-0.0022	-0.0026
	(-0.050, 0.029)	(-0.024, 0.071)	(-0.022, 0.023)	(-0.019, 0.052)	(-0.007, 0.003)	(-0.010, 0.005)
lag-6 years	-0.0067	0.0020	0.0065	0.0084	-0.0021	-0.0027
	(-0.047, 0.033)	(-0.041, 0.045)	(-0.013, 0.026)	(-0.022, 0.038)	(-0.007, 0.003)	(-0.010, 0.005)
lag-7 years	-0.0155	-0.0214	0.0022	-0.0055	-0.0022	-0.0030
	(-0.055, 0.024)	(-0.067, 0.025)	(-0.020, 0.024)	(-0.038, 0.027)	(-0.007, 0.002)	(-0.010, 0.004)
lag-8 years	-0.0283	-0.0402	-0.0037	-0.0123	-0.0022	-0.0038
	(-0.074, 0.017)	(-0.100, 0.019)	(-0.032, 0.025)	(-0.056, 0.031)	(-0.007, 0.002)	(-0.011, 0.004)
lag-9 years	-0.0346	-0.0461	-0.0091	-0.0170	-0.0019	-0.0038
	(-0.086, 0.016)	(-0.110, 0.018)	(-0.041, 0.023)	(-0.066, 0.032)	(-0.006, 0.0025)	(-0.011, 0.0033)
lag-10 years	-0.0364	-0.0467	-0.0144	-0.0232	-0.0012	-0.0031
	(-0.089, 0.016)	(-0.108, 0.015)	(-0.048, 0.019)	(-0.071, 0.025)	(-0.006, 0.0035)	(-0.010, 0.004)

Notes: Significant results at the 95% confidence level are shown in bold.
a. Model 1 includes lagged terms of Gini index (95% confidence interval in parentheses, based on clustered standard errors), year dummies, log(gdp) and country fixed effects.
b. Model 2 additionally adds democracy index, education, independence, armed conflict and economic freedom.

Supplementary Table 6 Linear regression coefficients of GINI (from Luxembourg Income Study) from fixed effects models linking income inequality and population health measured by life expectancy and mortality, 24 European countries for the year 1990/1995/2000/2005[a]

Outcomes	Men		Women	
	Model 1[b]	Model 2[c]	Model 1	Model 2
Life expectancy	2.538	2.087	4.220	2.862
	(-5.123, 10.20)	(-5.559, 9.733)	(-3.650, 12.09)	(-5.576, 11.30)
All causes	-0.153	-0.402	-0.360	-0.405
	(-0.889, 0.583)	(-0.993, 0.189)	(-1.126, 0.407)	(-1.200, 0.391)
All circulatory diseases	0.0351	-0.668	0.386	-0.269
	(-1.306, 1.376)	(-1.657, 0.322)	(-1.043, 1.814)	(-1.438, 0.901)
Ischemic heart disease	0.0714	-1.424	1.111	-0.962
	(-1.987, 2.130)	(-4.021, 1.174)	(-1.273, 3.494)	(-4.118, 2.195)
Cerebrovascular disease	-0.582	-1.940	-0.717	**-2.033**
	(-2.359, 1.195)	(-3.937, 0.057)	(-2.554, 1.120)	**(-3.833, -0.234)**
All cancers	0.217	-0.244	0.412	-0.189
	(-0.645, 1.079)	(-1.135, 0.646)	(-0.616, 1.440)	(-1.207, 0.829)
Cancer of lung	-0.458	-1.060	-0.845	-0.768
	(-1.901, 0.984)	(-2.357, 0.238)	(-2.204, 0.515)	(-1.954, 0.419)
Cancer of breast			0.610	-0.0790
			(-1.008, 2.227)	(-2.087, 1.929)
All infectious diseases	3.125	7.182	0.133	3.589
	(-4.558, 10.81)	(-2.023, 16.39)	(-6.105, 6.371)	(-4.827, 12.01)
Chronic liver disease and cirrhosis	-0.166	0.0147	2.060	2.610
	(-3.381, 3.049)	(-4.075, 4.104)	(-1.366, 5.485)	(-1.375, 6.595)
All external causes	1.066	1.401	1.137	**2.120**
	(-0.142, 2.273)	(-0.071, 2.873)	(-0.333, 2.607)	**(0.834, 3.405)**
Motor vehicle accidents	**2.206**	2.519	**3.811**	**4.028**
	(0.063, 4.349)	(-0.023, 5.061)	**(2.322, 5.299)**	**(2.121, 5.934)**
Suicide	0.216	-0.237	-0.546	-0.648
	(-1.815, 2.247)	(-2.317, 1.843)	(-3.286, 2.195)	(-3.757, 2.461)
Signs, symptoms and ill-defined	5.066	2.255	2.185	-1.549
	(-3.517, 13.65)	(-8.746, 13.26)	(-6.739, 11.11)	(-13.51, 10.41)
Homicide	1.924	-1.015	-0.703	-2.792
	(-3.723, 7.570)	(-8.283, 6.254)	(-5.387, 3.980)	(-8.282, 2.699)
	Infant			
Infant mortality	-1.202	-0.289		
	(-2.633, 0.230)	(-1.534, 0.956)		

Notes: Significant results at the 95% confidence level are shown in bold.
a. Based on the available data from LIS, we constructed longitudinal data for 24 European countries with measurements for 1990/1995/2000/2005. If income inequality was not available for the year needed, Gini from the nearest year was used. If there were more than 1 neighboring year, the linear interpolation was used. The included countries were Austria, Belgium, Czech Republic, Denmark, Estonia, Finland, Germany, Greece, Hungary, Ireland, Italy, Luxembourg, Netherlands, Norway, Poland, Romania, Russia, Slovak Republic, Slovenia, Spain, Sweden, Switzerland and United Kingdom.
b. Model 1 includes Gini index (95% confidence interval in parentheses, based on clustered standard errors), year dummies, log(gdp) and country fixed effects.
c. Model 2 additionally adds democracy index, education, independence, armed conflict and economic freedom.

Part II

Social and political determinants of inequalities in health

Chapter 4

Trends in socioeconomic inequalities in self-assessed health in 17 European countries between 1990 and 2010

Yannan Hu, Frank J. van Lenthe, Gerard J. Borsboom, Caspar W.N. Looman, Matthias Bopp, Bo Burström, Dagmar Dzúrová, Ola Ekholm, Jurate Klumbiene, Eero Lahelma, Mall Leinsalu, Enrique Regidor, Paula Santana, Rianne de Gelder, Johan P. Mackenbach.
Trends in socioeconomic inequalities in self-assessed health in 17 European countries between 1990 and 2010. Journal of Epidemiology and Community Health, 2016. Published online first: doi:10.1136/jech-2015–206780

ABSTRACT

Background

Between the 1990s and 2000s, relative inequalities in all-cause mortality increased, whereas absolute inequalities decreased in many European countries. Whether similar trends can be observed for inequalities in other health outcomes is unknown. This paper aims to provide a comprehensive overview of trends in socioeconomic inequalities in self-assessed health (SAH) in Europe between 1990 and 2010.

Methods

Data were obtained from nationally representative surveys from 17 European countries for various years between 1990 and 2010. The age-standardized prevalence of less-than-good SAH was analysed by education and occupation among men and women aged 30–79 years. Socioeconomic inequalities were measured by means of absolute rate differences and relative rate ratios. Meta-analysis with random effects models was used to examine the trends of inequalities.

Results

We observed declining trends in the prevalence of less-than-good SAH in many countries, particularly in Southern and Eastern Europe and the Baltic States. In all countries, less-than-good SAH was more prevalent in lower educational and manual groups. For all countries together, absolute inequalities in SAH were mostly constant, whereas relative inequalities increased. Almost no country consistently experienced a significant decline in either absolute or relative inequalities.

Conclusions

Trends in inequalities in SAH in Europe were generally less favourable than those found for inequalities in mortality, and there was generally no correspondence between the two when we compared the trends within countries. In order to develop policies or interventions that effectively reduce inequalities in SAH, a better understanding of the causes of these inequalities is needed.

INTRODUCTION

Europe offers excellent opportunities for conducting between-country comparative research of socioeconomic inequalities in health, and this can help to identify successful strategies to reduce these inequalities [1, 2]. Between 1990 and 2010, many European countries have made the reduction of health inequalities an explicit aim of national health policies [3–5]. A comprehensive overview of trends in health inequalities among a large set of European countries can show whether there are differences between countries in the extent to which a reduction of health inequalities has actually been achieved.

A recent study assessed trends in inequalities in premature mortality between the 1990s and 2000s [6]. The study found that relative inequalities increased in most populations in Europe except for Southern Europe, while absolute inequalities decreased in many European countries. Whether these trends can be generalised to other domains of health, however, is unknown. One other health outcome for which trend data on inequalities is available is self-assessed health (SAH), which has been shown to be a reliable indicator of general health and well-being [7–9], and an independent predictor of mortality and survival [10–12] that can be used in population health monitoring.

SAH is strongly associated to indicators of socioeconomic position in all countries with available data [1, 13–16] but studies on trends in these inequalities often focus on one country or at most, on a small number of countries [17–26]. One notable exception is a study covering 10 European countries between the 1980s and 1990s that showed a high degree of stability of these inequalities [27]. A comprehensive overview of recent trends in SAH inequalities based on a larger set of European countries is still lacking.

This paper aims to provide such a comprehensive overview by analysing trends in socioeconomic inequalities in SAH among adults in 17 European countries between 1990 and 2010.

METHODS

Data

We obtained nationally representative health surveys from 17 countries (Table 1). All available and comparable surveys between 1990 and 2010 were used, which led to a different number of surveys for each country. The first and last observation years differed between countries, but all years were between 1990 and 2010. Data came from the same survey over time for most countries, except for the Netherlands, Austria and Italy. However, the chosen surveys within the three countries have a high comparability [28–31], and thus can be used for trend analysis. The age range used for most countries was 30–79 years. Younger respondents were excluded because many of them were still receiving full-time education. Older respondents were excluded to avoid the potential bias caused by the exclusion of

Table 1 Countries included in the analysis, sources of data and descriptive information on self-assessed health (SAH)

Country	Survey year	Survey name	Included age range (years)	Number of included respondents[a]	Prevalence of less-than-good SAH among men[b] (%)	Prevalence of less-than-good SAH among women[b] (%)
Finland	1993/1995/1997/1999/2001/ 2003/2005/2007/2009	Health Behaviour and Health	30–79	3582 ~ 4069	36 ~ 44	33 ~ 42
Sweden	1996–1997/2004–2005	Swedish Survey of Living Conditions	30–79	7414 ~ 8524	21 ~ 23	26 ~ 28
Denmark	1994/2000/2005	Danish Health and Morbidity Survey	30–79	3322 ~ 12373	19 ~ 22	23 ~ 27
England	1990/1996/2000/2005	General Household Survey	30–69	8589 ~ 13517	34 ~ 40	40 ~ 46
Scotland	1995/1998/2003	Scottish Health Survey	30–64	5036 ~5997	24 ~ 27	25~ 26
Ireland	1998/2002/2007	Survey of Lifestyle and Nutrition	30–74	4235 ~ 7638	15 ~ 17	12 ~ 16
Netherlands	1990	Ongoing Survey of Living Condition (DLO)	30–79	3472	26	30
	1997/2000/2005/2009	Permanent Survey on Living Conditions (POLS)	30–79	5078 ~ 6169	20 ~ 24	25 ~ 29
Belgium	1997/2001/2004/2008	Health Interview Survey	30–79	6230 ~ 7811	23 ~ 25	27 ~ 31
Switzerland	1992/1997/2002/2007	Swiss Health survey	30–79	8267 ~14075	12 ~ 15	14 ~ 19
Austria	1991/1999	Micro Census	30–79	27817 ~ 28817	28 ~ 31	29 ~ 33
	2006	Health Interview Survey	30–79	11421	26	25
Italy	1994/2000/2005	Health and Health Care Utilization	30–79	36446 ~ 87673	37 ~ 45	45 ~ 53
	2010	Aspects of Daily Living	30–79	31417	30	36
Spain	1993/2001/2006	National Health Survey	30–79	14187 ~ 23396	27 ~ 31	38 ~ 41
Portugal	1995–1996/1998–1999/2005–2006	National Health Survey	30–79	26091 ~ 30199	50 ~ 62	64 ~ 77
Czech Republic	1993/1999/2002	Sample Survey of the Health Status	30–79	1137 ~ 1731	41 ~ 70	47 ~ 73
Poland	1996/2004/2009	Polish Health Interview Survey	30–69	21353 ~ 29712	40 ~ 64	44 ~ 71
Lithuania	1994/2000/2006/2010	Health Behaviour	30–64	1396 ~ 1723	54 ~ 66	51 ~ 68
Estonia	1996/2006	Estonian Health Interview Survey	30–79	3696 ~ 4734	58 ~ 65	57 ~ 67

Notes: a. We provide the range of the number of respondents per year for each country.
b. We provide the range of the age-standardised prevalence of less-than-good SAH per year for each country by gender.

institutionalised population in many surveys. As shown in Table 1, some countries had upper age limits that were different from 79 years (England, Scotland, Ireland, Poland and Lithuania). Although these different age limits could influence the comparability of the level of inequalities in SAH between countries, the risk of bias was considered limited in our analysis of trends in inequalities over time within each country. The number of included respondents per year ranged from 1137 (Czech Republic, 1993) to 87,673 (Italy, 2000). The prevalence of less-than-good SAH ranged from 12% (Switzerland, 2002, 2007) to 70% (Czech Republic, 1993) among men, and ranged from 12% (Ireland, 2002) to 77% (Portugal, 1995–1996) among women.

To measure SAH, we used answers from a question which was framed similarly to "how is your health in general?" In all countries except for England (in which three answer categories were used, i.e. "good/fairly good/not good"), five answer categories were distinguished, which were normally "very good/good/fair/bad/very bad". The precise answer categories varied slightly in some countries, but the consistency over time was retained in all countries. We dichotomised the answers by collapsing those that reported a less-than-good SAH into one category.

Socioeconomic position was measured by education and occupation. Educational levels were recorded as the highest level of education completed or currently being attended by the respondent. It was harmonised on the basis of the International Standard Classification of Education (ISCED) and reclassified into three categories: levels 0–2 (no, primary or lower secondary education, considered "low-educated"), levels 3–4 (upper secondary and post-secondary non-tertiary education, considered "middle-educated"), levels 5–6 (tertiary education, considered "high-educated"). Occupational classes were classified as "manual" versus "non-manual". Respondents who were economically inactive, and who could not be classified on the basis of their last or main occupation were classified as missing. For trends in education-related inequalities, we included all 17 available countries with 62 country-year observations. For trends in occupation-related inequalities, we included 16 countries with 53 country-year observations, where Belgium was excluded due to a large number of missing values for occupation. Some recent years for the Netherlands and Finland were also excluded as information on occupation was not available.

Statistical methods

The prevalence rates of less-than-good SAH were calculated by country, year, sex, and education or occupation, and age-standardised to the European Standard Population [32] using the direct standardisation method.

Inequalities were measured by means of absolute prevalence rate differences (RD) and relative prevalence rate ratios (RR) of low versus high level of education or manual versus non-manual occupation. A bootstrap procedure with 1000 iterations was used to calculate 95% confidence intervals. We also calculated the slope index of inequality (SII) and relative index of inequality (RII) based on education, which took into account the distribution of the population by education [33]. In order to facilitate the comparison between education-related and occupation-related inequalities, however, we mainly used the RD and RR as the inequality measures, and provide the results for the RII and SII in the online supplementary figure S1. Survey weights were available for some countries or years. Unweighted results are reported in the results section. Analyses with available weighting factors are reported in the online supplement figure S2 and S3 as a sensitivity analysis.

To study the trends over time in each country and in the ensemble of countries as a whole, we employed meta-analysis with random-effects models, using the prevalence of less-than-good SAH, the RD and RR by education or occupation as the outcomes. In the analysis of the trends within each country, the year of data collection was used as the only moderator variable in the models. In the analysis of the trends in the ensemble of countries, we pooled all available observations and additionally added the country dummies into the regressions. In all models, the regression parameters for the year of data collection were taken as indicators of the linear time trends. The estimated country-specific linear time trends and their 95% confidence intervals were plotted in the form of forest plots together with the estimate for all countries as a whole (the "average" in the forest plots). A reference vertical line representing no change over time was plotted. The I^2 statistic measuring the proportion of total variability explained by heterogeneity and the p value testing the residual heterogeneity were reported. The meta-analysis was performed with the R (3.0.1.) package metafor [34].

RESULTS

Figure 1 shows the results of the meta-analysis for the trends in age-standardised prevalence of less-than-good SAH in the 17 European countries between 1990 and 2010. As shown in the forest plots, a statistically significant decline in the prevalence of less-than-good SAH was observed in a pooled analysis including all countries (represented by "average"), among males and females. These declines were mainly the result of significant declines in some, but not all countries, including Italy, Portugal, Czech Republic, Poland, Lithuania and Estonia.

To give an overview of the levels of and variation in inequalities among the 17 European countries, the absolute and relative inequalities in SAH by educational level (Figure 2) are presented, using the most recent year for which data were available in each country. Absolute education-related inequalities in SAH, as measured by RD, were found in all countries and these ranged between 0.08 and 0.35, indicating between 8% and 35% points difference in less-than-good SAH between the low and high educated, with no clear pattern emerging between different regions in Europe. Relative inequalities as measured by RR were also present in all countries, and these ranged between 1.26 and 4.14. These were particularly high in Scotland, Ireland, Switzerland (among males only) and Austria. Inequalities in SAH by occupational class are presented in the online supplementary figure S4. Absolute and

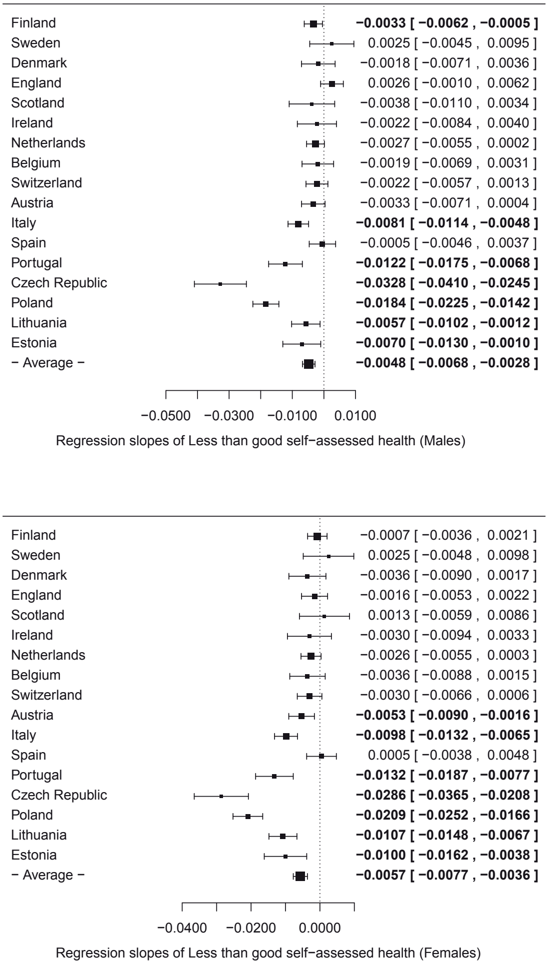


Figure 1 Meta-analysis for the trends in age-standardised prevalence of less-than-good self-assessed health (SAH) for men and women separately, 1990–2010

Note: The regression slope (with 95% confidence interval in brackets) indicates the estimated linear trend in age-standardised prevalence of less-than-good SAH within each country or in the ensemble of countries as a whole. For males, I^2=97.44%, test for residual heterogeneity p<0.0001. For females, I^2=98.04%, test for residual heterogeneity p<0.0001.

relative inequalities by occupation were observed in all countries. Again, no clear pattern was observed among different European regions.

Results from the meta-analysis for the trends in absolute and relative inequalities based on education are reported as forest plots (Figure 3a and 3b). Pooling all countries together, we found no statistically significant trend in absolute educational inequalities among males. Also, no significant trend was observed within most countries, except for an increasing

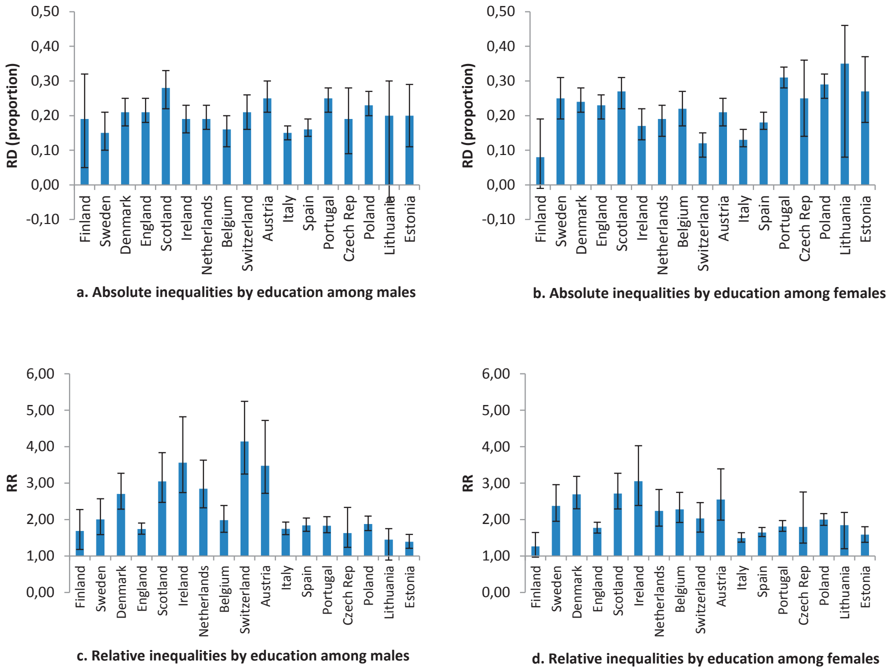
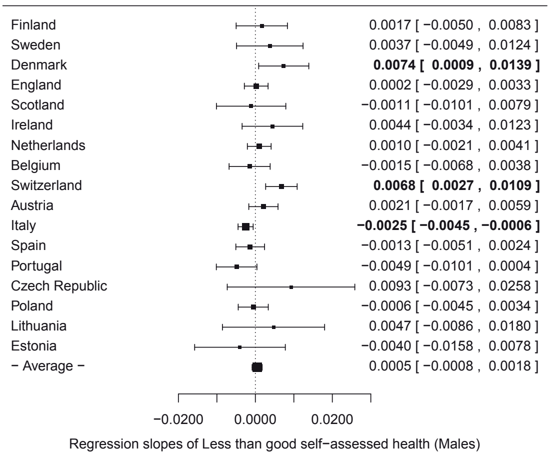


Figure 2 Absolute and relative inequalities in the prevalence of less-than-good self-assessed health (SAH) according to educational level for men and women separately, using the most recent year for each country (RD, rate difference; RR, rate ratio)

trend in inequalities in Denmark and Switzerland, and a declining trend in Italy. Among females, a significantly increasing trend was found in a pooled analysis of all countries and in England, Scotland, the Netherlands, Poland and Lithuania. The general picture became less favourable when relative inequalities were used. Pooling all countries together, we found significantly increasing trends in relative inequalities among males and females. Among males, significant increases in relative inequalities were detected in Switzerland, Austria and Poland. Among females, significant increases in relative inequalities were observed in England, the Netherlands, Switzerland, Portugal, Poland and Lithuania. No country showed a significantly decreasing trend in relative inequalities.

Results from the meta-analysis for trends in inequalities based on occupation are presented in Figure 4a and 4b. For males, the trends resembled those seen for education, with relatively stable trends for absolute inequalities in a pooled analysis, as well as in most countries. The stable trends were also found among females in a pooled analysis and in most countries. Again, the general picture became less favourable when relative inequalities were studied. When we pooled all countries together, significant increases in relative inequalities were found among males and females, as well as in a number of separate countries. Significant decline in relative inequalities was only observed in England among males.



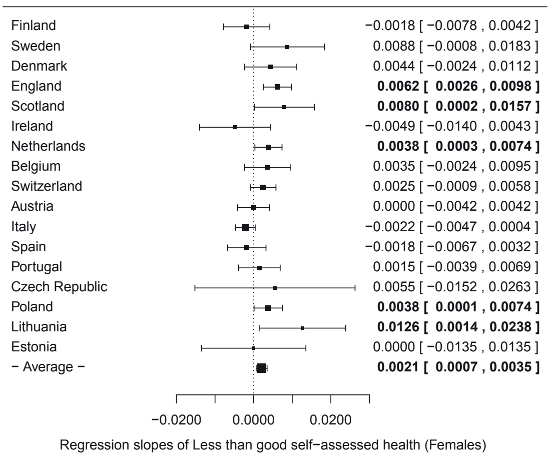
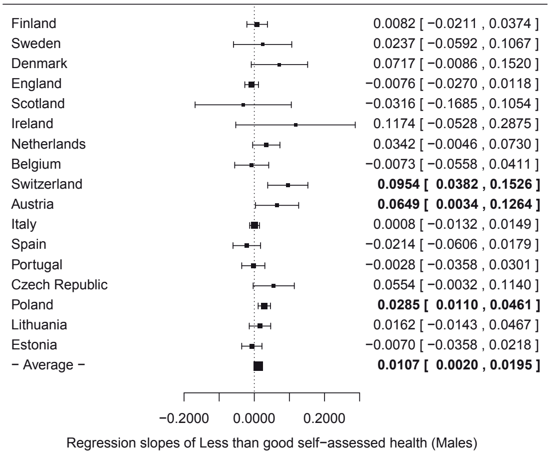


Figure 3a Meta-analysis for the trends in absolute inequalities in less-than-good self-assessed health (SAH) according to educational level for men and women separately, 1990–2010

Note: The regression slope (with 95% confidence interval in brackets) indicates the estimated linear trend in absolute inequalities in less-than-good SAH based on educational level within each country or in the ensemble of countries as a whole. For absolute inequalities among males, I^2=30.36%, test for residual heterogeneity p=0.0706. For absolute inequalities among females, I^2=32.34%, test for residual heterogeneity p=0.0074.

Chapter 4



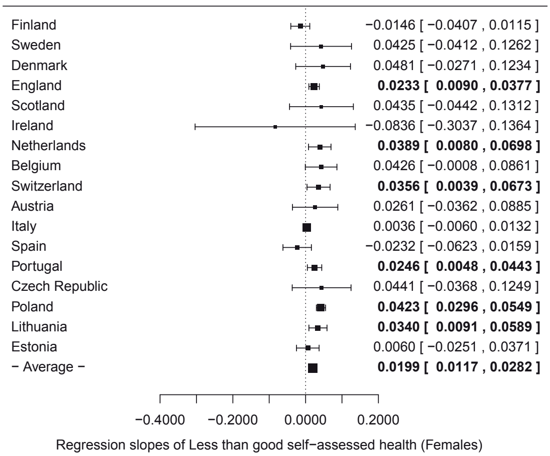
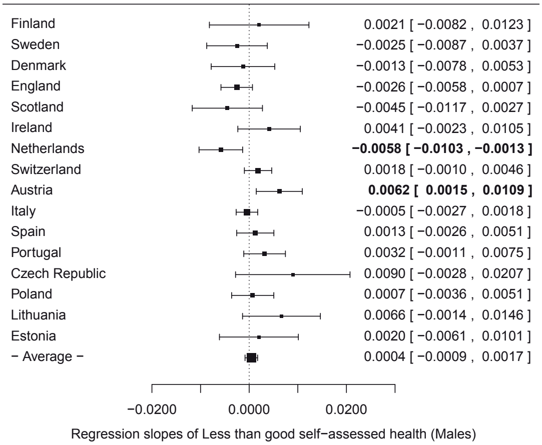


Figure 3b Meta-analysis for the trends in relative inequalities in less-than-good self-assessed health (SAH) according to educational level for men and women separately, 1990–2010

Note: The regression slope (with 95% confidence interval in brackets) indicates the estimated linear trend in relative inequalities in less-than-good SAH based on educational level within each country or in the ensemble of countries as a whole. For relative inequalities among males, I^2=24.28%, test for residual heterogeneity p=0.0160. For relative inequalities among females, I^2=38.27%, test for residual heterogeneity p=0.0043.



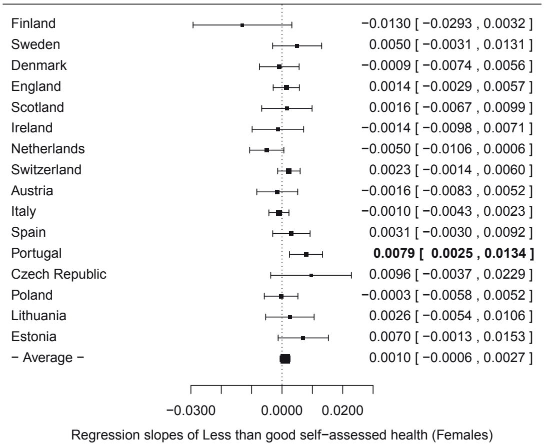
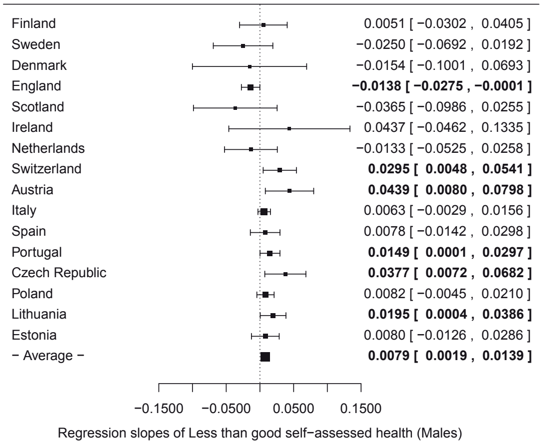


Figure 4a Meta-analysis for the trends in absolute inequalities in less-than-good self-assessed health (SAH) according to occupational class for men and women separately, 1990–2010

Note: The regression slope (with 95% confidence interval in brackets) indicates the estimated linear trend in absolute inequalities in less-than-good SAH based on occupational class within each country or in the ensemble of countries as a whole. For absolute inequalities among males, I^2=54.18%, test for residual heterogeneity p=0.0001. For absolute inequalities among females, I^2=62.96%, test for residual heterogeneity p<0.0001.



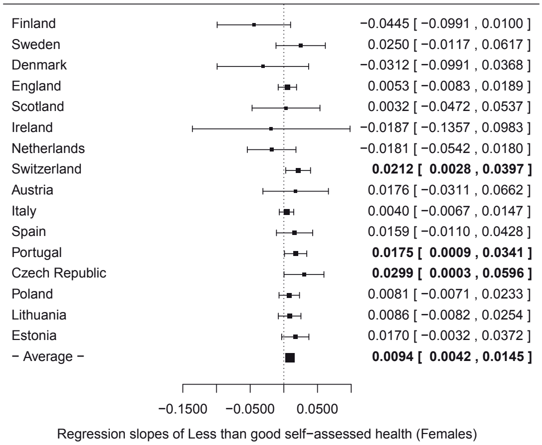


Figure 4b Meta-analysis for the trends in relative inequalities in less-than-good self-assessed health (SAH) according to occupational class for men and women separately, 1990–2010

Note: The regression slope (with 95% confidence interval in brackets) indicates the estimated linear trend in relative inequalities in less-than-good SAH based on occupational class within each country or in the ensemble of countries as a whole. For relative inequalities among males, I^2=56.12%, test for residual heterogeneity p=0.0005. For relative inequalities among females, I^2=56.17%, test for residual heterogeneity p<0.0001.

DISCUSSION

Summary of findings

We observed declining trends in the prevalence of less-than-good SAH in many countries, particularly in Southern and Eastern Europe and the Baltic States. In all countries, less-than-good SAH was more prevalent in lower educational and manual groups. For all countries together, absolute inequalities in SAH were mostly constant, whereas relative inequalities increased. Almost no country consistently experienced a statistically significant decline in either absolute or relative inequalities.

Interpretation

For all countries together, the prevalence of less-than-good SAH declined during the study period, while socioeconomic inequalities (mainly relative inequalities) in SAH generally increased. This can perhaps be partly explained by the commonly observed negative association between the prevalence of health problems and the magnitude of relative inequalities, which results from the fact that relative declines in prevalence tend to be larger in higher socioeconomic groups because of their lower prevalence [35]. However, this cannot explain the fact that among females absolute educational inequalities also increased in some countries.

We mainly used the RD and RR of low versus high education, in order to facilitate the comparison to the RD and RR by occupation. However, measures like the RD and RR cannot take into account the distribution of the population by socioeconomic group [33]. In many European countries, the proportion of individuals with a low level of education is decreasing over time [18, 27, 36]. Thus, the unfavourable trends in inequalities measured by RD and RR might reflect the fact that the shrinking low-educated group is increasingly composed of people who have been socially marginalised [37]. This is partly confirmed by the results based on the SII and RII by education, which adjust for these differences in population composition (see supplementary figure S1). This analysis finds stable trends in inequalities in all countries among males, and significantly increasing trends in a few countries and in the ensemble of countries as a whole only among females.

Although we included more countries and adopted a new technique to assess the trends in inequalities, it is worthwhile to compare our findings to the previous ones. Discrepancies between the inequalities in SAH and inequalities in mortality were found when we compared the magnitude of the inequalities [6], which is consistent with existing findings [1]. As for the trends in inequalities, our overall results about relative inequalities in SAH are consistent with those about relative inequalities in mortality between the 1990s and the 2000s [6], where increasing inequalities were found in many populations. However, on a country-specific basis, the two do not always correspond. Countries for which increasing relative inequalities in mortality and SAH were found include Switzerland and Lithuania, but Finland, Sweden, Belgium and Estonia showed stable relative inequalities in SAH and increasing relative inequalities in mortality. Reductions of absolute inequalities were commonly found for mortality (except for the Baltic States), but were not generally found for SAH. One potential explanation for the different trends in absolute inequalities is that the mortality among higher educated is reaching a level below which it is difficult to decline further without new breakthroughs in prevention or treatment [6], whereas this might be not true for SAH as it is a subjective self-reported measure of health. Our results suggest that the trends in inequalities seen for mortality cannot be generalised to other health outcomes such as SAH.

In agreement with other studies, socioeconomic inequalities in SAH were found to be persistent over time [17, 18, 20, 22, 24, 25]. Although the study periods were not exactly the same, our findings are generally consistent with those from previous studies using data from one country or a small number of countries, for example, stable inequalities in Nordic countries [18] and increased educational inequalities among Dutch women [21]. In an earlier study [27], stable trends in SAH inequalities between the 1980s and 1990s was found in many European countries. It is worthwhile to compare this to the more recent trends found in our study. England and the Netherlands, where stable education-related inequalities were reported between the 1980s and 1990s, showed increasing trends, particularly among females, in the recent years covered by our study. This is disappointing against the background of the increasing awareness among policy-makers of socioeconomic inequalities, and the implementation of policies to tackle inequalities. In contrast, significant widening of education-related inequalities in Spain and Italy was found between the 1980s and 1990s; however, this was not detected in our recent period. Stable trends in education-related inequalities in Finland and Sweden were consistently found in earlier and recent periods.

More generally, our results do not support the idea that countries with national policies to tackle health inequalities have fared better in terms of inequalities in SAH than countries without such policies. England is the first and only European country that pursued a systematic and well-resourced policy to reduce inequalities in health [38]. As indicated above, inequalities in SAH have not narrowed down in England during this period except for the relative inequalities by occupation among males; in fact, education-related inequalities among females increased. Scotland also pursued a coordinated action plan to tackle health inequalities [39], but inequalities in SAH were rather stable and even increased among females during the study period. The Netherlands is another country that has had some national activities to reduce health inequalities [40], but while absolute occupation-related inequalities among males declined, no positive changes were seen for education-related inequalities. Among all the countries, Italy shows the most encouraging trends as it had a significantly deceasing trend in prevalence of less-than-good SAH, and declines in absolute inequalities by education. However, efforts to reduce inequalities have not been stronger in Italy than in other countries [4]. Other structural developments may have undone the effects of policies on health inequalities, such as the economic recessions in some European countries in the early 1990s and the late 2000s [41], increases in income inequalities (e.g. England) [20, 42], and changes of working conditions (e.g. more work-related stress) [21]. Nevertheless, the observed trends of inequalities in SAH are not consistent with the amount of efforts made to reduce health inequalities in some countries.

Strengths and limitations

This is the largest study ever of inequalities in SAH in terms of the number of countries and years included. It is also for the first time that a meta-analysis has been used to systematically assess the trends in inequalities. Two indicators (education and occupation) were used to capture the multidimensional nature of the concept of socioeconomic position, and we included both absolute and relative inequalities.

One issue to consider in all international comparisons is data comparability. Despite great harmonisation efforts, we were not able to remove all differences between countries in data collection such as the framing of survey questions, the population coverage or the response rates. The tendency to report less-than-good SAH may differ between countries due to different cultural backgrounds [9, 43], for example, persons in Central and Eastern European countries may tend to report their health as less good than persons in other European countries [44, 45]. However, as we focused on the trends in inequalities in SAH, we retained comparability over time within each country, and therefore consider the risk of bias due to between-country variations in these aspects to be limited. Weighting factors were available for some countries or years as the aim was to compensate the survey design and make the sample representative of the population. Essentially similar results were obtained in a sensitivity analysis of trends in RD and RR, where weighting factors were incorporated when available (Supplementary Figure S2 and Figure S3).

Another concern is that the populations used to assess education-related and occupation-related inequalities differed. When we assessed the occupation-related inequalities, respondents who were economically inactive and who could not be classified on the basis of their last or main occupation were coded as missing, whereas they were included when education-related inequalities were assessed. This resulted in a larger percentage of missing values among people older than 65 years in occupation-based analysis (e.g. in Denmark and Austria), than in education-based analysis. It may have resulted in smaller occupation-related inequalities, since economically inactive people tend to have worse health as compared with employed people, and tend to originate from the lower occupational groups. Again, the impact on the comparison between trends in occupation-related and education-related inequalities is likely to be limited.

The broad age range used in the analysis might hide the potential heterogeneity in the trends in inequalities in SAH among different age groups. Therefore, we did a supplementary analysis using a smaller age range of 30–64 years, which could also facilitate the comparison between trends in occupation-related and education-related inequalities. We found that limiting the analysis to the age group 30–64 years did not essentially change our results (Supplementary Figure S5 and Figure S6).

Our meta-analysis assumed a linear trend in the outcome measures – an assumption that might not always hold. It cannot be excluded that results based on a non-linear trend assumption would have changed our conclusions. Nevertheless, this technique gives a useful "helicopter view" of the trends in inequalities in SAH in Europe. Future hypothesis-driven research should assess whether non-linear trends better fit the data.

Owing to lack of appropriate data, our analysis was mainly focused on recent trends in inequalities in SAH among adults. Future research should consider exploring the trends in SAH inequalities among adolescents.

Conclusions

For all countries together, relative socioeconomic inequalities in less-than-good SAH widened, whereas absolute inequalities were more stable. Trends in inequalities in SAH in Europe were generally less favourable than those found for inequalities in mortality over the same time period, and there was generally no correspondence between the two when we compared the trends within countries. In order to develop policies or interventions that effectively reduce inequalities in SAH, a better understanding of the causes of these inequalities is needed.

REFERENCES

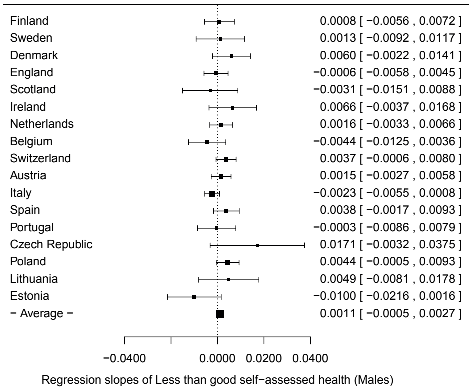
1. Mackenbach, J.P., et al., *Socioeconomic inequalities in health in 22 European countries.* N Engl J Med, 2008. **358**(23): p. 2468–81.
2. Kunst, A.E., V. Bos, and J.P. Mackenbach, *Guidelines for monitoring health inequalities in the European Union.* Rotterdam: the Netherlands: Department of Public Health, 2001.
3. Mackenbach, J.P., M.J. Bakker, and E.N.I. Pol, *Tackling socioeconomic inequalities in health: analysis of European experiences.* Lancet, 2003. **362**(9393): p. 1409–1414.
4. Judge, K., et al., *Health inequalities: a challenge for Europe.* An independent, expert report commissioned by the UK, Presidency of the EU (February 2006), 2006.
5. Crombie, I.K., et al., *Closing the health inequalities gap: an international perspective.* World Health Organization, 2005.
6. Mackenbach, J.P., et al., *Trends in inequalities in premature mortality: a study of 3.2 million deaths in 13 European countries.* J Epidemiol Community Health, 2014.
7. Manderbacka, K., *Examining what self-rated health question is understood to mean by respondents.* Scandinavian Journal of Social Medicine, 1998. **26**(2): p. 145–153.
8. Miilunpalo, S., et al., *Self-rated health status as a health measure: The predictive value of self-reported health status on the use of physician services and on mortality in the working-age population.* Journal of Clinical Epidemiology, 1997. **50**(5): p. 517–528.
9. Jylha, M., *What is self-rated health and why does it predict mortality? Towards a unified conceptual model.* Social Science & Medicine, 2009. **69**(3): p. 307–316.
10. Idler, E.L. and Y. Benyamini, *Self-rated health and mortality: a review of twenty-seven community studies.* J Health Soc Behav, 1997. **38**(1): p. 21–37.
11. Huisman, M., F. van Lenthe, and J. Mackenbach, *The predictive ability of self-assessed health for mortality in different educational groups.* Int J Epidemiol, 2007. **36**(6): p. 1207–13.
12. Bopp, M., et al., *Health risk or resource? Gradual and independent association between self-rated health and mortality persists over 30 years.* PLoS One, 2012. **7**(2): p. e30795.
13. Alvarez-Galvez, J., et al., *The Impact of Socio-Economic Status on Self-Rated Health: Study of 29 Countries Using European Social Surveys (2002–2008).* International Journal of Environmental Research and Public Health, 2013. **10**(3): p. 747–761.
14. Eikemo, T.A., et al., *Health inequalities according to educational level in different welfare regimes: a comparison of 23 European countries.* Sociology of Health & Illness, 2008. **30**(4): p. 565–582.
15. von dem Knesebeck, O., P.E. Verde, and N. Dragano, *Education and health in 22 European countries.* Social Science & Medicine, 2006. **63**(5): p. 1344–1351.
16. Eikemo, T.A., et al., *Class-related health inequalities are not larger in the East: a comparison of four European regions using the new European socioeconomic classification.* Journal of Epidemiology and Community Health, 2008. **62**(12): p. 1072–1078.
17. Helasoja, V., et al., *Trends in the magnitude of educational inequalities in health in Estonia, Latvia, Lithuania and Finland during 1994–2004.* Public Health, 2006. **120**(9): p. 841–853.
18. Lahelma, E., et al., *Analysing changes of health inequalities in the Nordic welfare states.* Social Science & Medicine, 2002. **55**(4): p. 609–625.
19. Lahelma, E., et al., *Widening or narrowing inequalities in health? Comparing Britain and Finland from the 1980s to the 1990s.* Sociology of Health & Illness, 2000. **22**(1): p. 110–136.
20. Gravelle, H. and M. Sutton, *Income related inequalities in self assessed health in Britain: 1979–1995.* Journal of Epidemiology and Community Health, 2003. **57**(2): p. 125–129.

21. Dalstra, J.A.A., et al., *Trends in socioeconomic health inequalities in the Netherlands, 1981–1999*. Journal of Epidemiology and Community Health, 2002. **56**(12): p. 927–934.
22. Daponte-Codina, A., et al., *Area deprivation and trends in inequalities in self-rated health in Spain, 1987–2001*. Scandinavian Journal of Public Health, 2008. **36**(5): p. 504–515.
23. Heistaro, S., E. Vartiainen, and P. Puska, *Trends in self-rated health in Finland 1972–1992*. Preventive Medicine, 1996. **25**(5): p. 625–632.
24. Granstrom, F., et al., *Exploring trends in and determinants of educational inequalities in self-rated health*. Scand J Public Health, 2015. **43**(7): p. 677–86.
25. Maheswaran, H., E. Kupek, and S. Petrou, *Self-reported health and socio-economic inequalities in England, 1996–2009: Repeated national cross-sectional study*. Soc Sci Med, 2015. **136–137**: p. 135–46.
26. Pfortner, T.K. and F.J. Elgar, *Widening inequalities in self-rated health by material deprivation? A trend analysis between 2001 and 2011 in Germany*. J Epidemiol Community Health, 2015.
27. Kunst, A.E., et al., *Trends in socioeconomic inequalities in self-assessed health in 10 European countries*. Int J Epidemiol, 2005. **34**(2): p. 295–305.
28. Großschadl, F. and W.J. Stronegger, *Long-term trends in obesity among Austrian adults and its relation with the social gradient: 1973–2007*. European Journal of Public Health, 2013. **23**(2): p. 306–12.
29. European Foundation for the Improvements of Living and Working Conditions, *Trends in quality of work in the Nethelands*. 2007: http://www.eurofound.europa.eu/publications/htmlfiles/ef05167.htm (accessed 20 Oct 2014).
30. Italian National Institute of Statistics (ISTAT), *Multipurpose survey on households: aspects of daily life*. Italian National Institute of Statistics, http://siqual.istat.it/SIQual/visualizza.do?id=0058000 (accessed 12 December 2014).
31. Italian National Institute of Statistics (ISTAT), *Condizioni di salute e ricorso ai servizi sanitari [Health and Health Care Utilization]*. Italian National Institute of Statistics, http://www.istat.it/it/archivio/5471 (accessed 12 December 2014).
32. Ahmad, O.B., et al., *Age Standardization of Rates: a New WHO Standard.* . GBP Discussion Paper Series No. 31. World Health Organization http://www.who.int/healthinfo/paper31.pdf (accessed 10 October 2014), 2001.
33. Mackenbach, J.P. and A.E. Kunst, *Measuring the magnitude of socio-economic inequalities in health: an overview of available measures illustrated with two examples from Europe*. Social Science & Medicine, 1997. **44**(6): p. 757–71.
34. Viechtbauer, W., *Conducting Meta-Analyses in R with the metafor Package*. Journal of Statistical Software, 2010. **36**(3): p. 1–48.
35. Eikemo, T.A., V. Skalicka, and M. Avendano, *Variations in relative health inequalities: are they a mathematical artefact?* International Journal for Equity in Health, 2009. **8**.
36. Silventoinen, K. and E. Lahelma, *Health inequalities by education and age in four Nordic countries, 1986 and 1994*. J Epidemiol Community Health, 2002. **56**(4): p. 253–8.
37. Bronnum-Hansen, H. and M. Baadsgaard, *Increase in social inequality in health expectancy in Denmark*. Scand J Public Health, 2008. **36**(1): p. 44–51.
38. Mackenbach, J.P., *Can we reduce health inequalities? An analysis of the English strategy (1997–2010)*. J Epidemiol Community Health, 2011. **65**(7): p. 568–75.
39. Scottish Executive, *Closing the opportunity gap*. 2004: http://www.scotland.gov.uk/resource/doc/46997/0024934.pdf (accessed 10 October 2014).
40. Mackenbach, J.P. and K. Stronks, *A strategy for tackling health inequalities in the Netherlands*. BMJ, 2002. **325**(7371): p. 1029–32.

41. Bacigalupe, A. and A. Escolar-Pujolar, *The impact of economic crises on social inequalities in health: what do we know so far?* Int J Equity Health, 2014. **13**: p. 52.
42. Jutz, R., *The role of income inequality and social policies on income-related health inequalities in Europe.* Int J Equity Health, 2015. **14**(1): p. 117.
43. Jurges, H., *True health vs response styles: Exploring cross-country differences in self-reported health.* Health Economics, 2007. **16**(2): p. 163–178.
44. Babones, S.J., *The consistency of self-rated health in comparative perspective.* Public Health, 2009. **123**(2): p. 199–201.
45. Reile, R., et al., *The recent economic recession and self-rated health in Estonia, Lithuania and Finland: a comparative cross-sectional study in 2004–2010.* J Epidemiol Community Health, 2014. **68**(11): p. 1072–9.

Chapter 4

Online Supplementary Material



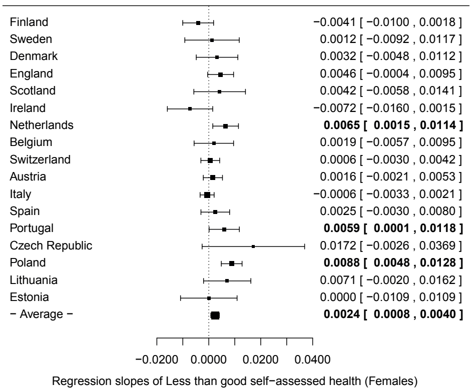
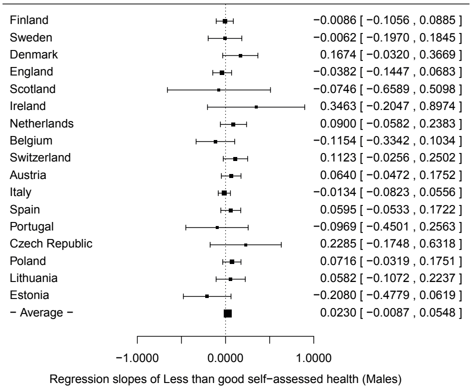


Figure S1.a Meta-analysis for the trends in slope index of inequality in less-than-good self-assessed health (SAH) according to educational level for men and women separately, 1990–2010

Note: The regression slope (with 95% confidence interval in brackets) indicates the estimated linear trend in slope index of inequality in less-than-good SAH based on education within each country or in the ensemble of countries as a whole. For slope index of inequalities among males, I^2=32.59%, test for residual heterogeneity p=0.0146. For slope index of inequalities among females, I^2=38.04%, test for residual heterogeneity p=0.0024.



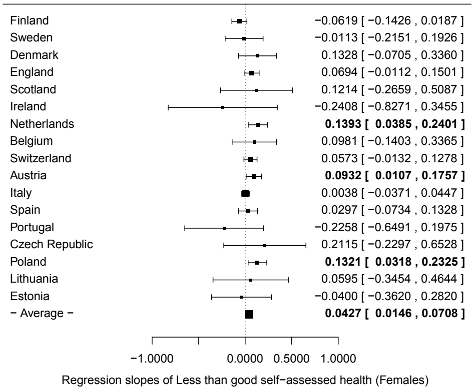
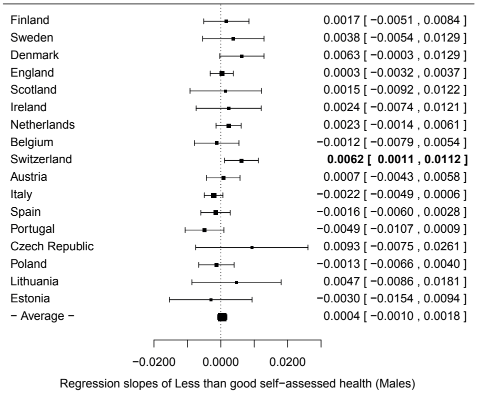


Figure S1.b Meta-analysis for the trends in relative index of inequality in less-than-good self-assessed health (SAH) according to educational level for men and women separately, 1990–2010

Note: The regression slope (with 95% confidence interval in brackets) indicates the estimated linear trend in relative index of inequality in less-than-good SAH based on education within each country or in the ensemble of countries as a whole. For relative index of inequalities among males, I^2=25.78%, test for residual heterogeneity p=0.0182. For relative index of inequalities among females, I^2=23.61%, test for residual heterogeneity p=0.0512.



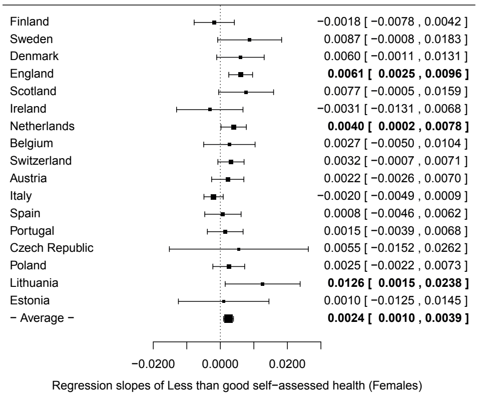
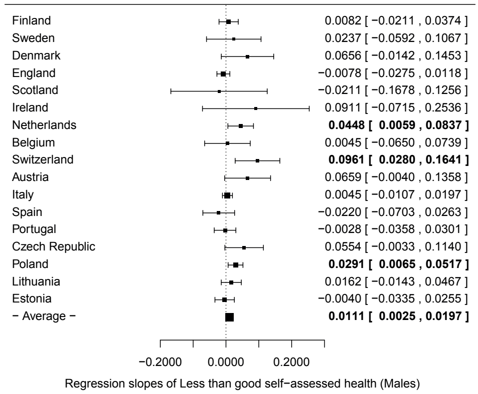


Figure S2.a Meta-analysis for the trends in absolute inequalities in less-than-good self-assessed health (SAH) according to educational level for men and women separately, 1990–2010, weighted
Note: The regression slope (with 95% confidence interval in brackets) indicates the estimated linear trend in absolute inequalities in less-than-good SAH based on educational level within each country or in the ensemble of countries as a whole. For absolute inequalities among males, I^2=24.83%, test for residual heterogeneity p=0.1571. For absolute inequalities among females, I^2=25.68%, test for residual heterogeneity p=0.0339.



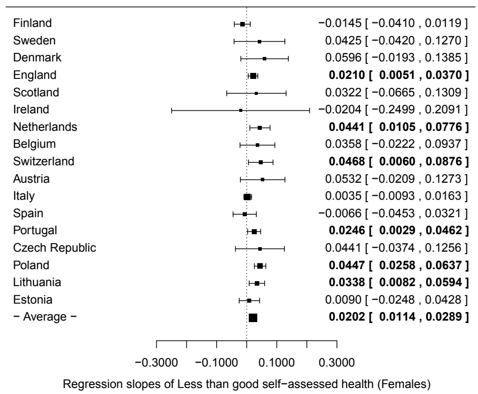
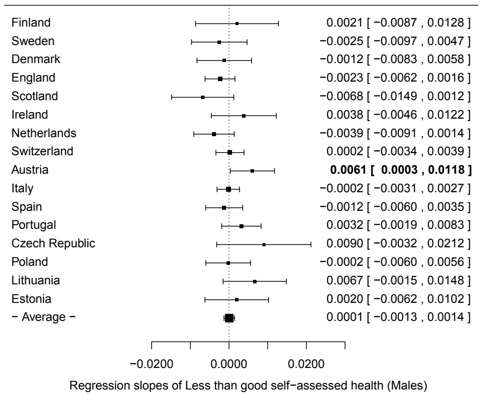


Figure S2.b Meta-analysis for the trends in relative inequalities in less-than-good self-assessed health (SAH) according to educational level for men and women separately, 1990–2010, weighted

Note: The regression slope (with 95% confidence interval in brackets) indicates the estimated linear trend in relative inequalities in less-than-good SAH based on educational level within each country or in the ensemble of countries as a whole. For relative inequalities among males, I^2=12.47%, test for residual heterogeneity p=0.1224. For relative inequalities among females, I^2=33.65%, test for residual heterogeneity p=0.0191.



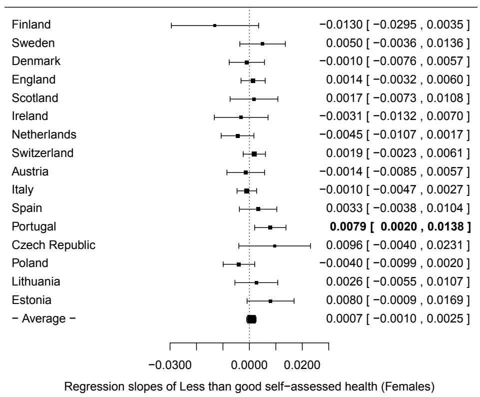
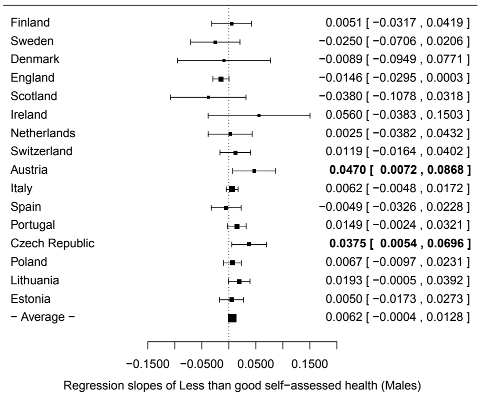


Figure S3.a Meta-analysis for the trends in absolute inequalities in less-than-good self-assessed health (SAH) according to occupational class for men and women separately, 1990–2010, weighted

Note: The regression slope (with 95% confidence interval in brackets) indicates the estimated linear trend in absolute inequalities in less-than-good SAH based on occupational class within each country or in the ensemble of countries as a whole. For absolute inequalities among males, I^2=45.42%, test for residual heterogeneity p=0.0020. For absolute inequalities among females, I^2=63.53%, test for residual heterogeneity p<0.0001.



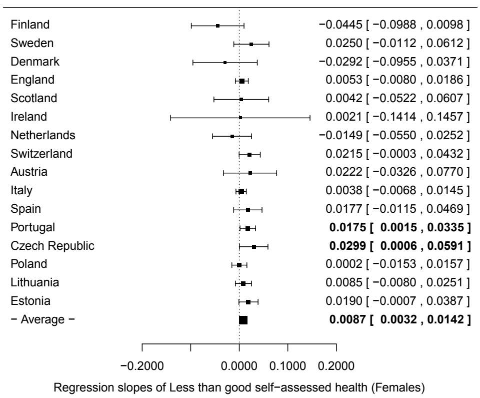


Figure S3.b Meta-analysis for the trends in relative inequalities in less-than-good self-assessed health (SAH) according to occupational class for men and women separately, 1990–2010, weighted

Note: The regression slope (with 95% confidence interval in brackets) indicates the estimated linear trend in relative inequalities in less-than-good SAH based on occupational class within each country or in the ensemble of countries as a whole. For relative inequalities among males, I^2=57.27%, test for residual heterogeneity p=0.0003. For relative inequalities among females, I^2=53.75%, test for residual heterogeneity p<0.0001.

Chapter 4

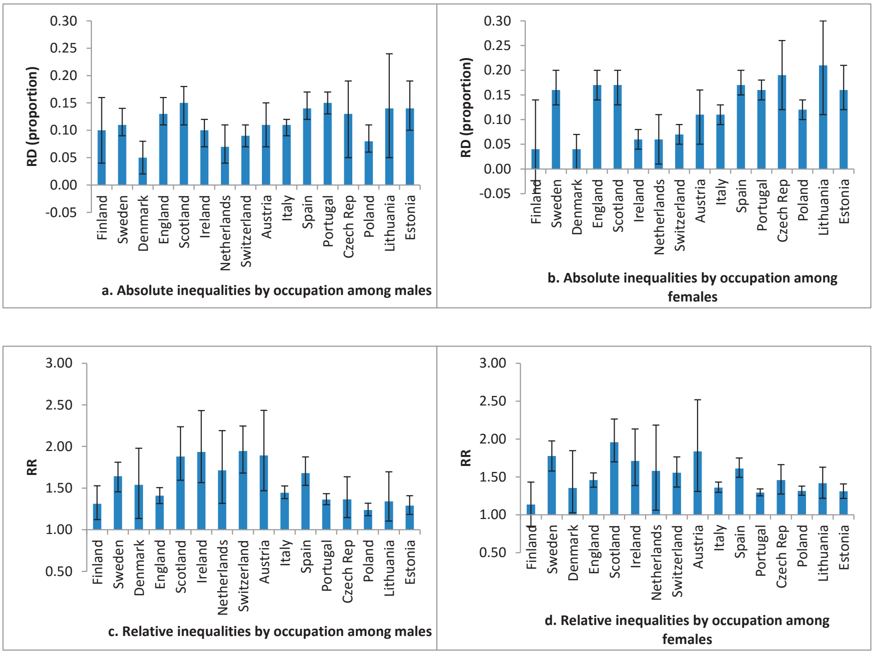


Figure S4 Absolute and relative inequalities in the prevalence of less-than-good SAH according to occupational class for men and women separately, using the most recent year for each country

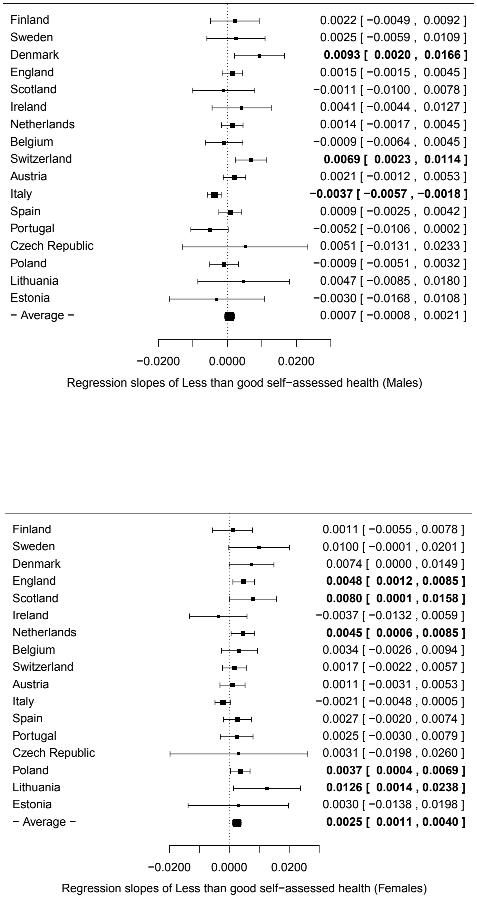
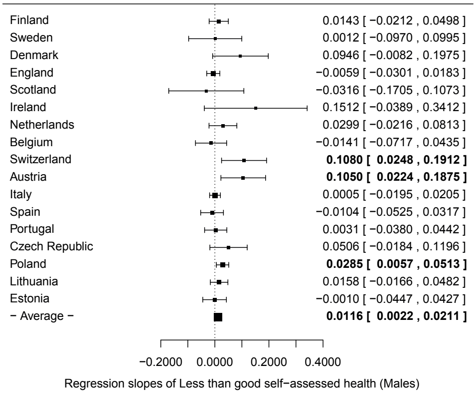


Figure S5.a Meta-analysis for the trends in absolute inequalities in less-than-good self-assessed health (SAH) according to educational level for men and women separately, 1990–2010, using the age group 30–64

Note: The regression slope (with 95% confidence interval in brackets) indicates the estimated linear trend in absolute inequalities in less-than-good SAH based on educational level within each country or in the ensemble of countries as a whole. For absolute inequalities among males, I^2=36.15%, test for residual heterogeneity p=0.0180. For absolute inequalities among females, I^2=29.56%, test for residual heterogeneity p=0.0212.



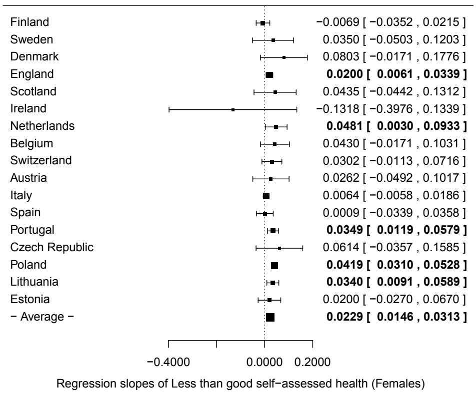
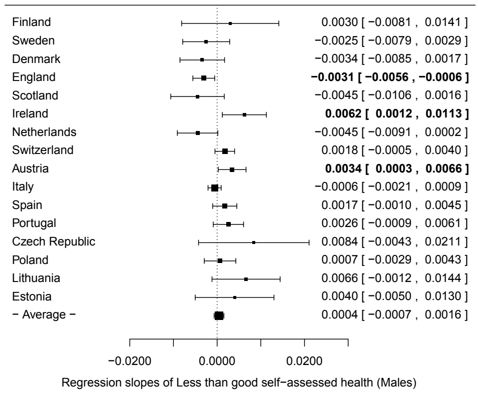


Figure S5.b Meta-analysis for the trends in relative inequalities in less-than-good self-assessed health (SAH) according to educational level for men and women separately, 1990–2010, using the age group 30–64
Note: The regression slope (with 95% confidence interval in brackets) indicates the estimated linear trend in relative inequalities in less-than-good SAH based on educational level within each country or in the ensemble of countries as a whole. For relative inequalities among males, I^2=18.62%, test for residual heterogeneity p=0.0117. For relative inequalities among females, I^2=27.85%, test for residual heterogeneity p=0.0386.



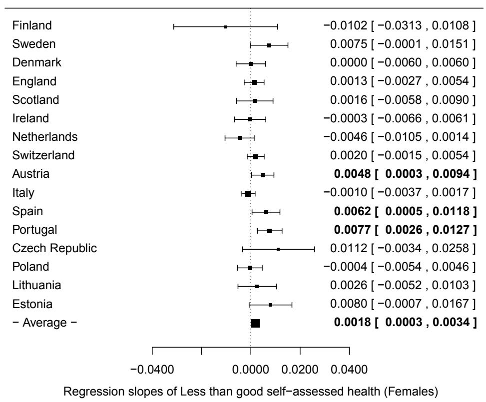
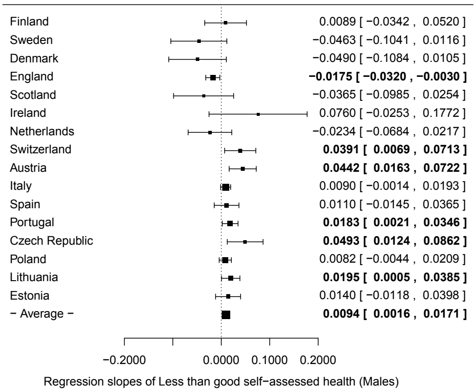


Figure S6.a Meta-analysis for the trends in absolute inequalities in less-than-good self-assessed health (SAH) according to occupational class for men and women separately, 1990–2010, using the age group 30–64
Note: The regression slope (with 95% confidence interval in brackets) indicates the estimated linear trend in absolute inequalities in less-than-good SAH based on occupational class within each country or in the ensemble of countries as a whole. For absolute inequalities among males, I^2=35.77%, test for residual heterogeneity p=0.0143. For absolute inequalities among females, I^2=57.41%, test for residual heterogeneity p<0.0001.

Chapter 4



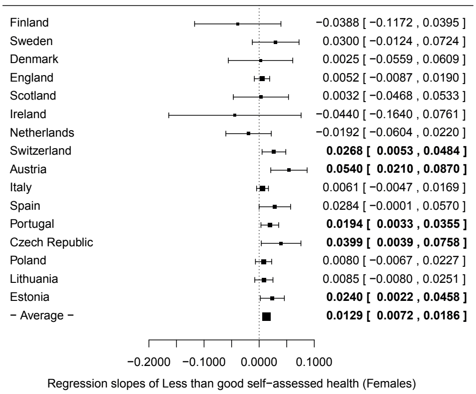


Figure S6.b Meta-analysis for the trends in relative inequalities in less-than-good self-assessed health (SAH) according to occupational class for men and women separately, 1990–2010, using the age group 30–64
Note: The regression slope (with 95% confidence interval in brackets) indicates the estimated linear trend in relative inequalities in less-than-good SAH based on occupational class within each country or in the ensemble of countries as a whole. For relative inequalities among males, I^2=62.53%, test for residual heterogeneity p<0.0001. For relative inequalities among females, I^2=54.69%, test for residual heterogeneity p=0.0001.

Chapter 5

Assessing the impact of natural policy experiments on socioeconomic inequalities in health: how to apply commonly used quantitative analytical methods?

Yannan Hu, Frank J van Lenthe, Rasmus Hoffmann, Karen van Hedel, Johan P. Mackenbach. Assessing the impact of natural policy experiments on socioeconomic inequalities in health: how to apply commonly used quantitative analytical methods? (submitted)

ABSTRACT

Background

The scientific evidence-base for policies to tackle health inequalities is limited. Natural policy experiments (NPE) have drawn increasing attention as a means to evaluating the effects of policies on health. Several analytical methods can be used to evaluate the outcomes of NPEs in terms of average population health, but it is unclear whether they can also be used to assess the outcomes of NPEs in terms of health inequalities. The aim of this study therefore was to assess whether, and to demonstrate how, a number of commonly used analytical methods for the evaluation of NPEs can be applied to quantify the effect of policies on health inequalities.

Methods

We identified seven quantitative analytical methods for the evaluation of NPEs: regression adjustment, propensity score matching, difference-in-differences analysis, fixed effects analysis, instrumental variable analysis, regression discontinuity and interrupted time-series. We assessed whether these methods can be used to quantify the effect of policies on the magnitude of health inequalities either by conducting a stratified analysis or by including an interaction term, and illustrated both approaches in a fictitious numerical example.

Results

All seven methods can be used to quantify the equity impact of policies on absolute and relative inequalities in health by conducting an analysis stratified by socioeconomic position, and all but one (propensity score matching) can be used to quantify equity impacts by inclusion of an interaction term between socioeconomic position and policy exposure.

Conclusion

Methods commonly used in economics and econometrics for the evaluation of NPEs can also be applied to assess the equity impact of policies, and our illustrations provide guidance on how to do this appropriately. The low external validity of results from instrumental variable analysis and regression discontinuity makes these methods less desirable for assessing policy effects on population-level health inequalities. Increased use of the methods in social epidemiology will help to build an evidence base to support policy making in the area of health inequalities.

INTRODUCTION

There is overwhelming evidence for the existence of socioeconomic inequalities in health in many countries [1–3]. Improvements in understanding their underlying mechanisms have reached a point where several entry-points have been identified for interventions and policies aimed at reducing health inequalities [2,4]. The latter has often been made a priority in national and local health policy [2,5–9]. Yet, the scientific evidence-base for interventions and policies to tackle health inequalities is still very limited, and mostly applies to the proximal determinants of health inequalities such as smoking and working conditions [10–14]. Policies that address the social and economic conditions in which people live probably have the greatest potential to reduce health inequalities, but these are the hardest to evaluate [15].

Randomized controlled trials (RCTs) are regarded as the "gold standard" in the effect evaluation of clinical studies. The limitations of RCT's in evaluating policies in public health, however, have been clearly recognized [16,17]. For policies aimed at tackling health inequalities, an obvious limitation is that policies to improve material and psychosocial living conditions, access to essential (health care) services, and health-related behaviours often cannot be randomized.

Natural policy experiments (NPEs), defined as "policies that are not under the control of the researchers, but which are amenable to research using the variation in exposure that they generate to analyse their impact" have been advocated as a promising alternative [18,19]. In NPEs, researchers exploit the fact that often not all (groups of) individuals are exposed to the policy, e.g. because some individuals are purposefully assigned to the policy and others are not, or because the policy is implemented in some geographical units but not in others. For example, a policy to improve housing conditions in neighbourhoods might be implemented in neighbourhoods where the need to do so is largest, or some cities may decide to implement the policy and others not. Of course, in these cases those in the intervention and control group are likely to differ in many other factors than exposure to the policy, and analytical methods will have to adequately control for confounding in order to allow reliable causal inference.

The application of methods for the evaluation of NPEs, such as difference-in-differences and regression discontinuity, is reasonably well advanced in economics and econometrics. While these methods have also entered the field of public health [20,21], and have been applied occasionally to study policy impacts on health inequalities [22,23], there is as yet no general understanding of whether and how each of these methods can be applied to assess the impact of policies on the magnitude of socioeconomic inequalities in health. If they can, however, they can help to extend the evidence-base in this area substantially.

The main aim of this study therefore is to assess whether, and to demonstrate how, a number of commonly used analytical methods for the evaluation of NPEs can be applied to quantify the impact of policies on health inequalities. In doing so, we will also pay attention to two issues that may complicate assessing the impact of policies on socioeconomic inequalities in health. Firstly, socioeconomic inequalities in health can be measured in different ways. Secondly, policies may reduce health inequalities in different ways.

With regard to the measurement of health inequalities, it is important to distinguish relative and absolute inequalities. Relative inequalities in health are usually measured by taking the ratio of the morbidity or mortality rate in lower socioeconomic groups relative to those in higher socioeconomic groups, e.g. an odds ratio (OR), a rate ratio (RR), or a relative index of inequality [24]. Absolute inequalities in health are usually measured by taking the difference between the morbidity or mortality rates of lower and higher socioeconomic groups, e.g. a simple rate difference or the more complex slope index of inequality [24]. Relative and absolute inequalities both are considered important, although it is sometimes argued that a reduction in absolute inequalities is a more relevant policy outcome than a reduction in relative inequalities, because it is the absolute excess morbidity or mortality in lower socioeconomic groups that ultimately matters most for individuals. Nevertheless, quantitative methods used for the evaluation of policies should be able to measure the impact on both absolute and relative inequalities in health.

With regard to the second issue, there are two ways through which a policy can reduce socioeconomic inequalities in health: (1) the policy has a larger effect on exposed people in lower socioeconomic groups, or (2) more people in lower socioeconomic groups are exposed to it. Clearly, both can also occur simultaneously; raising the tax on tobacco may affect individuals with lower incomes more than those with higher incomes, and given the higher prevalence of smokers in low income groups also affects more smokers in low than high income groups. In fact, changes in aggregated health outcomes collected for a country or region (e.g. mortality rates or the prevalence of self-assessed health) after the introduction of a policy are the result of an effect among the exposed as well as the proportion of exposed persons. For the ultimate goal to assess whether a reduction in health inequalities in the population occurred this is less relevant – one could argue that eventually only the end result counts, that is a change in the magnitude of socioeconomic inequalities in health. Many statistical techniques, however, 'only' provide the effect of the policy among the exposed; they do not take into account the proportion of persons exposed to a policy. In order to be able to quantify the impact of a policy on socioeconomic inequalities in health in a population, an additional step is then needed: the policy effect should be combined with information about the proportion of exposed persons in higher and lower socioeconomic groups.

The structure of this paper is as follows. We first describe a fictitious data example that allowed us to assess the applicability of seven commonly used analytic methods techniques for evaluating NPEs, which we also briefly describe. We then demonstrate the use of these methods for assessing the impacts of policies on the magnitude of health inequalities in our fictitious dataset. Finally, we discuss the advantages and disadvantages of the various methods.

DATA AND METHODS

A fictitious data example

We generated a fictitious dataset of 20,000 residents of a city. In this city, half of the residents were low educated, and within each educational group there were 50% males. The health outcome that we used was self-assessed health, dichotomized into either 'poor' or 'good'. The numbers (shown in Box 1) were chosen such that the proportion of persons with poor health before the introduction of the policy was higher among the low educated persons (20%) than among the high educated persons (10%). At one point in time, the city council introduced a free medical care service in a number of neighborhoods, most of which were deprived. Thus, relatively more low educated people were exposed to the policy (50%) as compared to high educated people (25%). At the same time, more women than men used the free health care within each educational group, because women care more about their health. Because women already had better health before the introduction of the policy and

Box 1 Numbers of residents in a city: a fictitious dataset

Education (n)	Sex (n)	Policy allocation (n)	Self-assessed health	Before the policy (Health t_1)	After the policy (Health t_2)
Low (10000)	Male (5000)	Exposed[1] (1250)	Poor	333	221
			Good	917	1029
		Unexposed (3750)	Poor	1000	950
			Good	2750	2800
	Female (5000)	Exposed (3750)	Poor	500	333
			Good	3250	3417
		Unexposed (1250)	Poor	167	159
			Good	1083	1091
High (10000)	male (5000)	Exposed (625)	Poor	83	46
			Good	542	579
		Unexposed (4375)	Poor	584	467
			Good	3791	3908
	female (5000)	Exposed (1875)	Poor	125	70
			Good	1750	1805
		Unexposed (3125)	Poor	208	166
			Good	2917	2959

[1] exposure was defined as actually using the free medical care service

tended to be more exposed to the intervention, gender was a confounder in the association between the policy exposure and self-assessed health.

We assumed that the effect of the policy was a reduction of the prevalence (or probability) of poor health among the exposed of 30%, regardless of their education level. Moreover, we imposed a naturally occurring recovery from poor to good health: even without the intervention, high educated people had a 20% chance of reverting to good health and low educated people had a 5% chance of reverting to good health. This could be due to spontaneous recovery or to external conditions such as other policies or changes in macro-economic factors, which were not directly related to the policy introduced. As a result, and for example, the number of low educated men who had poor health and who were exposed to the policy declined from 333 before the policy was implemented to 221 (333*0.70*0.95) after the policy was implemented (see Box 1). As those with good health were assumed not to change to poor health, the number of low educated persons exposed to the policy with good health became 1029 (917+(333–221)). Similarly, and as another example, the number of high educated women unexposed to the policy with good health after the introduction of the policy became 2959 (2917 + 208*0.2).

Compared to men, a smaller proportion of women reported poor health before the policy, and more women were exposed to the policy: the proportion of poor health before the policy was 20% (2000/10000) among men and 10% (1000/10000) among women, and the proportion of persons exposed to the policy was 56.25% for women (5625/10000) and only 18.75% for men (1875/10000). Gender thus was a confounder of the relation between policy exposure and health.

Quantitative methods for the evaluation of natural policy experiments

To identify potentially relevant quantitative methods for the evaluation of NPEs, we started by reviewing the classical econometric literature [20, 25–31]. Seven quantitative methods were identified as potentially suitable for the evaluation of NPE's (Table 1): (1) regression adjustment, (2) propensity score matching, (3) difference-in-differences analysis, (4) fixed effects analysis, (5) instrumental variable analysis, (6) regression discontinuity and (7) interrupted time-series. We will not elaborate upon the general application of these methods – for this we refer the reader to existing textbooks and papers [20,25,31,32]. Nevertheless, a basic understanding of the concepts behind these techniques is important for our purposes.

1. *Regression adjustment:* Standard multivariate regression techniques allow investigating the effect of a policy by adjusting the association between policy exposure and health outcomes for observed differences between those exposed and unexposed to the policy in the prevalence of confounding factors. Theoretically, if all possible confounders can be controlled for, the estimated policy effect will be unbiased. It is unrealistic to assume, however, that all possible confounders can be measured.

Table 1: Concepts, limitations and applications of statistical approaches for the evaluation of natural policy experiments

Method	Main concept	Minimum data requirement	Adjustment for confounders	Main limitations	Application to the evaluation of policies on health inequalities
Regression adjustment	Adjustment for confounders, i.e. factors related to both intervention allocation and health outcomes.	Cross-sectional	Observed confounders	Vulnerable to unobserved confounders	[49]
Propensity score matching	For a given propensity score, exposure to the intervention is random. The intervention effect is therefore the average difference in the outcomes between the exposed and the matched unexposed units with the same propensity scores.	Cross-sectional	Observed confounders	Vulnerable to unobserved confounders	[50]
Difference-in-differences	As long as the naturally occurring changes over time in the intervention and control group are the same, the difference in the change in the outcome between both groups can be interpreted as the intervention effect.	Repeated cross-sectional	Observed and time-invariant unobserved confounders	Vulnerable to violation of the common trend assumption	[22]
Fixed effects	Multiple observations within units are compared, such as repeated measurements over time within individuals. Effects of unobserved confounders that differ between units but remain constant over time are eliminated.	Longitudinal	Observed and time-invariant unobserved confounders	Vulnerable to unobserved time-variant confounders; Knocks out all cross-sectional variations between units; Susceptible to measurement errors over time;	[51,52]
Instrumental variable approach	An instrument creates variation in exposure to the intervention, without being directly related to the outcome itself.	Cross-sectional	Observed and unobserved confounders	Difficult to find good instrumental variables; Exogeneity of instruments cannot be easily tested; Weak instruments and finite samples might result in bias; Local average treatment effect problem;	[53]

Chapter 5

Table 1: Concepts, limitations and applications of statistical approaches for the evaluation of natural policy experiments (continued)

Method	Main concept	Minimum data requirement	Adjustment for confounders	Main limitations	Application to the evaluation of policies on health inequalities
Regression discontinuity	As long as the association between a variable and an outcome is smooth, any discontinuity in the outcome after a cut-off point of this variable can be regarded as an intervention effect.	Cross-sectional	Observed and unobserved confounders	Low external validity; Local average treatment effect problem in a fuzzy design;	[23]
Interrupted time-series	Identification of a sudden change in level of the health outcome (a change of intercept) or a more sustained change in trend of the health outcome (a change of slope) around the time of the implementation of the intervention.	Repeated measures	Observed confounders	Difficult to evaluate the interventions implemented slowly, or need unpredictable time to be effective; Vulnerable to other external interventions or shocks within the period;	[54, 55,56]

2. *Propensity score matching:* Propensity score matching involves estimating the 'propensity' or likelihood that each person or group has of being exposed to the policy, based on a number of known characteristics, and then matching exposed to unexposed individuals based on similar levels of the propensity score. Propensity score matching assumes that for a given propensity score, exposure to the policy is random. It is similar to regression analysis with control for confounding in that it aims to reduce bias due to observed confounding variables.

3. *Difference-in-differences analysis:* Difference-in-differences analysis compares the change in outcome for an exposed group between a moment before and a moment after the implementation of a policy to the change in outcome over the same time period for a non-exposed group. The two groups may have different levels of the outcome before the policy, but as long as any 'naturally occurring' changes over time can be expected to be the same for both, the difference in the change in outcome between the exposed and non-exposed groups will be an unbiased estimate of the policy effect.

4. *Fixed effects analysis*: Fixed effects analysis compares multiple observations within the same individuals or groups over time, and reveals the average change in the outcome due to the policy. Because each individual or group is compared with itself over time, differences between individuals or groups that remain constant over time – even if unmeasured – are eliminated and cannot confound the results.

5. *Instrumental variable analysis:* Instrumental variable analysis involves identifying a variable predictive of exposure to the policy, which in itself has no direct relationship with the outcome except through its effects on policy exposure or through other variables which have been adjusted in the regression. The technique uses the variation in outcome generated by this 'instrument' to test whether exposure to the policy is related to the outcome.

6. *Regression discontinuity:* Regression discontinuity is a form of analysis that can be used when areas or individuals are assigned to a policy depending on a cut-off point of a continuous measure. The basic idea is that, conditional on the relationship between the assignment variable and the outcome, the exposure to the policy at the cut-off point is as good as random, comparing health outcomes of those just below and just above the cut-off point provides an estimate of the effect of the policy.

7. *Interrupted time-series:* Where time-series data are available and there is a clear-cut change in policy at a specific point in time, interrupted time-series analysis can be used to estimate the policy effect. Regression analysis is used to detect any sudden change in level of the health outcome (in regression terms: a change of intercept) or a more sustained change in the trend of the health outcome (in regression terms: a

change of slope) around the time the policy is implemented. The analysis estimates the policy effect by comparing the health outcomes before and after policy implementation.

Statistical assessment of the impact of NPE in terms of socioeconomic inequalities in health

Analytically, assessing to what extent a policy does have an effect in lower and higher socioeconomic groups can be done in two ways. The first is to conduct a stratified analysis, using socioeconomic position as a stratification variable, resulting in policy effects for both lower and higher socioeconomic groups. The second is to include an interaction term between the variable for policy exposure and the indicator of socioeconomic position. For the latter, if the confounding effects of other covariates differ between socioeconomic groups, interaction terms between the indicator of socioeconomic position and these covariates also need to be added. If all interactions are included, the policy effects derived from an analysis stratified by socioeconomic position and from an analysis with interaction terms will be the same. For illustrative purpose, we included all the interactions in our analysis so that the results from interaction terms and stratified analysis were the same.

Most of the techniques described above require a regression analysis. Whereas a linear regression analysis results in an absolute effect of the policy, a logistic regression analysis results in a relative policy effect. Propensity score matching uses a pair-matched difference in the outcome to quantify the policy effect.

For those techniques resulting in a policy effect among the exposed only (all techniques described above, except interrupted time series), we then need to combine these effects with the proportion of exposed persons in higher and lower socioeconomic groups, in order to calculate the impact of policy on absolute and relative inequalities among the whole population. Currently, there is no prescribed statistical procedure to do this. Our approach is to calculate the prevalence of people having poor health in each educational group after the policy (an observed prevalence) and the predicted prevalence of people having poor health in absence of the policy (a predicted prevalence). The latter can be calculated by excluding the coefficient for the policy assignment from the equation, while keeping all other coefficients in the model the same. With the observed and predicted prevalence rates, absolute rate differences and relative rate ratios can be calculated. The differences in the absolute rate differences or the relative rate ratios with and without the policy then show the impact of the policy on the magnitude of health inequalities. Bootstrapping was used to calculate the confidence intervals around the estimated impact of a policy. All analyses were performed in Stata 13.1.

RESULTS

Regression adjustment

We illustrate this method using data obtained after the policy only ($Health_{t2}$), because this method is often applied in situations where data obtained before the policy are not available. In a stratified analysis, the effect of the policy can be modeled for those in higher and lower educational groups separately, adjusting for gender as a confounder:

$$Health_{it2} = \beta_0 + \beta_1 policy_i + \beta_2 gender_i + \mu_i$$

where $health_{it2}$ is the health of individual *i* in year *t2*, β_0 is the intercept, β_1 and β_2 are regression coefficients and μ_i is the error term.

If we use logistic regression, which is appropriate in situations with a binary health outcome as in our example, the odds ratio for the policy effect can be calculated from β_1 and represents the higher or lower odds of having poor health after the policy for those exposed to the policy as compared to those unexposed to the policy. Because gender in this example is the only confounder, and because we were able to measure and adjust for it, the odds ratio can be interpreted as the policy effect. Table 2 shows these policy effects for the low and high educated. The policy effect is essentially similar for low (OR=0.647, 95% CI [0.570, 0.734]) and high educated people (OR=0.679, 95% CI [0.550, 0.839]). Please note that this analysis gives us estimates of relative rather than absolute policy effects. The discrepancy between the estimated odds ratios for the policy effect (0.647 and 0.679) on the one hand and the policy effects that we imposed in the dataset (a reduction of the probability of poor health among the exposed as compared to the unexposed of 30% for both high and low educated groups) on the other hand is due to the logistic transformation.

Regression analysis also allows us to introduce an interaction term between (low) education and exposure to the policy ("low-edu*policy"):

$$Health_{it2} = \beta_0 + \beta_1 policy_i + \beta_2 gender_i + \beta_3 low\text{-}edu_i + \beta_4 (low\text{-}edu_i * policy_i) + \beta_5 (low\text{-}edu_i * gender_i) + \mu_i$$

where β_0 is the intercept, $\beta_1, \beta_2, \beta_3, \beta_4$ and β_5 are regression coefficients and μ is the error term.

As shown in table 2, the interaction term between the policy and education (β_4) was not statistically significant (0.953, 95% CI [0.745, 1.218]). This indicates that we cannot show that the policy effects for low- and high-educated people are different, which is in line with the findings from the stratified analysis.

These results only represent the relative policy effect for people exposed as compared to those unexposed to the policy; they do not take into account the proportion of exposed

Table 2 Policy effects derived from the seven methods based on education-stratified analysis and the inclusion of interaction terms

Method	Specification	low-educated [95% CI]	high-educated [95% CI]	interaction term [95% CI]
Regression adjustment	Logistic regression, adjusted for gender	0.647 [0.570, 0.734] (odds ratio)	0.679 [0.550, 0.839] (odds ratio)	0.953 [0.745, 1.218] (odds ratio)
Propensity score matching	Matched on gender	-0.048 [-0.065, -0.031] (probability difference)	-0.020 [-0.031, -0.009] (probability difference)	not applicable
Difference-in-differences	Logistic regression	0.666 [0.574, 0.773] (odds ratio)	0.687 [0.530, 0.890] (odds ratio)	0.970 [0.719, 1.307] (odds ratio)
Fixed effects	Linear regression, adjusted for time	-0.044 [-0.051, -0.037] (probability difference)	-0.016 [-0.023, -0.009] (probability difference)	-0.029 [-0.039, -0.019] (probability difference)
Instrumental variable	Probit regression	-0.050 [-0.063, -0.037] (probability difference)	-0.020 [-0.029, -0.011] (probability difference)	-0.036 [-0.057, -0.015] (probability difference)
Regression discontinuity	Logistic regression around the income threshold	0.678 [0.495, 0.929] (odds ratio)	0.687 [0.483, 0.977] (odds ratio)	0.987 [0.615, 1.583] (odds ratio)
Interrupted time-series	Linear regression	-0.023 [-0.027,-0.020] (probability difference)	-0.005 [-0.008, -0.002] (probability difference)	-0.019 [-0.023, -0.014] (probability difference)

people in each educational group. To do so, we had to apply an extra step. Using the stratified analyses, we calculated the predicted prevalence of poor health if the policy would have not been implemented (please note that we could have also used the analysis with the interaction terms; if all interactions are included this will provide exactly the same results). This was done by leaving out the term for the policy, keeping all other coefficients in the regression equations, and computing the predicted prevalence of poor health. Subsequently, we calculated the rate difference between high and low educated people using the observed prevalence (for the situation in which the policy was implemented), and the predicted prevalence (for the situation in which the policy would not have been implemented) (Table 3). For example, the rate difference in the situation with the policy effect was 9.14% (16.63–7.49) and was 11.11% without the policy. In a similar way, the rate ratios were also calculated for both situations. The impact of the policy on health inequality could now be measured (1) as the change in absolute inequality (e.g., as the change in the rate difference) or (2) as a change in relative inequality (e.g., as a change in the rate ratio). In our example, the change in the rate difference is 1.97% points (11.11% - 9.14%) which means that the policy reduced the absolute inequality between low- and high-educated people by almost 2 percent points (Table 4). Further, the change in the rate ratio was 12.2% ((2.39 – 2.22)/(2.39 – 1)). This means that the policy reduced the relative inequality by more than 12%. We have also calculated the confidence intervals of these estimates (Table 4).

Table 3 Observed and predicted prevalence of poor health, rate difference and rate ratio for low and high educated groups with and without the implementation of the policy, as obtained using the seven methods

	low-educated (%)	high-educated (%)	rate difference	rate ratio
Observed prevalence with policy effect	16.63	7.49	9.14	2.22
Predicted prevalence without the policy effect[1]				
Regression adjustment	19.11	8.00	11.11	2.39
Propensity score matching	19.03	7.99	11.04	2.38
Difference-in-differences analysis	18.97	7.98	10.99	2.38
Fixed effects models	18.84	7.88	10.96	2.39
Instrumental variable analysis	19.15	7.99	11.16	2.40
Regression discontinuity	Not comparable	Not comparable	Not comparable	Not comparable
Interrupted time-series	18.96	7.97	10.99	2.38

[1] As derived from the stratified analyses, reported as proportion of individuals with poor health (or, equivalently, individual probability of having poor health)

Table 4 Summary table of the policy effect on absolute and relative inequalities in health.

Method	**Estimated policy effect on absolute health inequality [a] (reduced rate difference in % points, [95% CI])**	**Estimated policy effect on relative health inequality [b] (reduced rate ratio, in %, [95% CI])**
1. Regression adjustment	1.97 [1.19, 2.76]	12.20 [4.49, 19.90]
2. Matching	1.89 [1.77, 2.02]	11.60 [8.99, 14.20]
3. Difference-in-differences	1.85 [0.88, 2.82]	11.33 [1.37, 21.29]
4. Fixed effects	1.82 [1.28, 2.36]	12.26 [5.45, 19.08]
5. Instrumental variable	2.02 [1.34, 2.69]	12.62 [6.07, 19.17]
6. Regression discontinuity	not comparable	not comparable
7. Interrupted time-series	1.85 [1.45, 2.26]	11.53 [6.05, 17.00]
Real policy effect	1.86	11.25
Simple before-and-after comparison	0.86	-22.03

[a.] We calculated the prevalence of people having poor health in each educational group following the real policy implementation and the predicted prevalence if leaving out the term for the policy effect (when there was no policy). The reported numbers represent the absolute reduction of the rate difference that can be attributed to the policy.
[b.] The reported numbers represent the relative reduction of the rate ratio (RR) calculated as follows: $(RR_{without\ policy} - RR_{with\ policy})/(RR_{without\ policy} - 1)*100$

Propensity score matching

We illustrate this method also with data obtained after the policy (Health_{t2}), because this method is often applied in situations where data before the policy are not available. In order to obtain an estimate for the effect of the policy on health inequalities we conducted a stratified analysis, i.e. we applied propensity score matching within the group of low and high educated individuals separately. The first step in the analysis was to calculate the "propensity" of being exposed to the policy. Logistic regression analysis, with being exposed or not as the binary outcome and gender as the predictor, was used to calculate the propensity of being exposed. Individuals with the same propensity who were indeed exposed to the policy could then be matched with individuals with almost the same ("the nearest neighbor") propensity who were not exposed to the policy.

The policy effect was estimated as the average of the differences in the probability of poor health within matched pairs of exposed and unexposed individuals. This produces an absolute measure of the policy effect. Table 2 lists the results obtained from the propensity score matching analysis for low and high educated people separately. For low and high educated people, the policy reduced the probability of having poor health among exposed individuals by almost 5 percentage points (-0.048) and 2 percentage points (-0.020), respectively. Although we imposed the same relative policy effect regardless of the education level in the data, the absolute effect of the policy was larger for low than for high educated people, because the prevalence of poor health before the policy was higher among the low educated.

To calculate the absolute decrease of the prevalence of poor health, the effect of the policy for low and high educated persons should be multiplied with the proportion of low and high educated persons in the population exposed to the policy. Among all the low educated, regardless of whether they were exposed or not to the policy, the probability of having poor health declined by 2.5 percentage points ((-0.048)*(5000/10000)=-0.024). Among all the high educated, the probability of having poor health declined by 0.5 percentage points ((-0.020)*(2500/10000)=-0.005).

In order to estimate the effect of the policy on the magnitude of health inequalities, we need the rate difference and rate ratio in a scenario with and in a scenario without the policy effect. In a scenario without the policy effect, the predicted prevalence of having poor health for the lower educated is the observed prevalence (16.63%) plus the reduction as a result of the policy (2.4%), which is then 19.03%. For high educated people, the prevalence is 7.99% (7.49% +0.5%).

The rate difference in the scenario with the policy was 9.14% (16.63–7.49); in the scenario without the policy it was 11.04% (19.03–7.99). This means that the policy reduced the absolute inequality in poor health by almost 2%. The rate ratio in the scenario with the policy was

2.22 (16.63/7.49); in the scenario without the policy it was 2.38 (19.03/7.99). This means that the policy reduced the relative inequality of poor health by almost 12%.

In propensity score matching, the policy effect is indicated by the average difference between the exposed and the matched unexposed individuals. There is no regression equation in the matching process, and therefore it was considered impossible to use an interaction term in a propensity matching analysis.

Difference-in-differences analysis

In order to illustrate this technique, we had to slightly modify our data example. Thus far, we only used data after the implementation of the policy. For the difference-in-differences analysis, we assumed that the data in our example had been collected in a repeated cross-sectional design.

In the analysis, we modeled health (measured both before and after implementation of the policy) as a function of exposure to the policy, time, and an interaction between exposure to the policy and time. By allowing levels of health to be different between exposed and unexposed before the policy, the technique accounts for unobservable confounding by time-invariant characteristics that differ in their prevalence between the exposed and unexposed. In our example 'gender' was not controlled for, and therefore acted as an unobservable confounder.

In a stratified analysis, the model to be used for low and high educated groups separately is:

$$Health_{it} = \beta_0 + \beta_1 policy_i + \beta_2 year_t + \beta_3 (policy_i * year_t) + \mu_{it}$$

where $health_{it}$ is the health of individual i in year t, β_0 is the intercept, β_1, β_2 and β_3 are regression coefficients and μ_{it} is the error term.

If we again use logistic regression, the coefficient for the variable "*policy*" (β_1) now measures the relatively higher or lower odds of having poor health for those exposed as compared to those unexposed to the policy before the implementation of the policy (which in our example was driven by the fact that women were in better health before the implementation and more exposed to the policy). The coefficient for the variable "*year*" (β_2) represents the naturally occurring change in health over time among the unexposed. The coefficient for the interaction term "*policy*year*" (β_3) indicates the policy effect, i.e. the difference in change of poor health over time between the unexposed and exposed. Table 2 shows the policy effects for low and high educated persons. The relative policy effect is essentially similar for low educated people (OR=0.666, 95% CI [0.574; 0.773]) and for high educated people (OR=0.687, 95% CI [0.530; 0.890]). This is again in line with what we imposed in the data.

In a difference-in-difference analysis, we can also introduce a three-way interaction term between policy, year and low education:

$$Health_{it} = \beta_0 + \beta_1 policy_i + \beta_2 year_t + \beta_3 (policy_i{*}year_t) + \beta_4 (low\text{-}edu_i) + \beta_5 (low\text{-}edu_i{*}policy_i) + \beta_6 (low\text{-}edu_i{*}year_t) + \beta_7 (low\text{-}edu_i{*}policy_i{*}year_t) + \mu_{it}$$

where $health_{it}$ is the health of individual i in year t, β_0 is the intercept, β_1 - β_7 are regression coefficients and μ_{it} is the error term.

The three-way interaction labeled "$low\text{-}edu_i{*}policy_i{*}year_t$" ($\beta_7$) indicates the differential policy effect for low and high educated persons. As shown in Table 2, this interaction term was not statistically significant (OR=0.970, 95% CI [0.719; 1.307]). Thus, the policy effect was not significantly different for low and high educated people, which corresponds to what we have imposed in the data example.

Using a similar approach as for the regression adjustment, and again based on the stratified analyses, we subsequently calculated the predicted prevalence of poor health if the policy would not have been implemented. It allowed us to calculate the rate differences between high and low educated people based on the predicted prevalence (if the policy would not have been implemented) as well as the rate ratios. As shown in Table 3, the policy effect on absolute health inequalities (e.g. the change in the rate differences) was 1.85% (10.99–9.14). This means that the policy reduced the rate difference between low- and high-educated people by almost 2 percentage points. Similarly, we can calculate the policy effect on relative health inequality as the change in the rate ratio, resulting in the finding that the policy reduced the relative excess risk of the lower educated by more than 11%.

Fixed effects model

In order to illustrate the fixed effects analysis, we considered our fictitious dataset to be a longitudinal dataset with repeated measures of self-assessed health before and after the implementation of the policy. With a binary outcome, one could use logistic regression analysis. However, in fixed effects logistic regression analysis, observations with the same health status in two (or more) periods will be excluded from the analysis; only the within-unit variations over time will be used. Therefore, a large part of the observations in our dataset would be excluded. While logistic regression analysis would produce valid estimates, it would lead to results that cannot be compared to those obtained from the other methods. For reasons of comparability, we used linear probability regressions with fixed effects, which also produces valid estimates.[1] Again, we treated 'gender' as an unobserved confounder.

Linear probability regression was used, in which the coefficient for the policy (β_1 in the formula below) represented an absolute change in the probability of having poor health as a result of exposure to the policy. In a stratified analysis, this can be modeled as follows for those in higher and lower educational groups separately:

$$Health_{it} = \beta_0 + \beta_1 policy_{it} + \beta_2 year_t + \mathrm{d}_i + \mu_{it}$$

where $health_{it}$ is health of individual i in year t, β_0 is the intercept, β_1 and β_2 are regression coefficients, d_i is a set of individual dummy variables and μ_{it} is the error term.

Table 2 shows that the absolute policy effect is larger among the low-educated ($\beta_{1=}$-0.044, 95% CI [-0.051; -0.037]) than among the high-educated ($\beta_{1=}$-0.016, 95% CI [-0.023; -0.009]).

Fixed effects analysis also allows us to introduce an interaction term between (low) education and exposure to the policy:

$$Health_{it} = \beta_0 + \beta_1 policy_{it} + \beta_2 year_t + \beta_3 (low\text{-}edu_{it} * policy_{it}) + \beta_4 (low\text{-}edu_{it} * year_t) + d_i + \mu_{it}$$

where $health_{it}$ is health of individual i in year t, β_0 is the intercept, $\beta_1 - \beta_4$ are regression coefficients, d_i is a set of individual dummy variables and μ_{it} is the error term.

As shown in Table 2, the interaction term for low education and policy ("*low-edu*policy*") is statistically significant (β_3=-0.029, 95% CI [-0.039; -0.019]), which indicates that the policy effect is indeed different between low- and high-educated people. The negative sign of the coefficient for the interaction term indicates that the absolute policy effect is larger among the lower educated, as was also found in the stratified analysis.

1 There was one disadvantage of using linear probability regressions with the fixed effects model in our dataset. Although the relative policy effect on health was independent of gender, there were some interaction effects between gender and policy if we use linear regression to assess the absolute policy effect. The absolute policy effect is lower among women, which is caused by a relatively lower prevalence of poor health among women. Strictly speaking, this makes the effect for women no longer "fixed" in an absolute setting. Women had a lower effect on health through the policy effect in the second period. The interaction effect however, was rather limited in our data, and did not lead to large discrepancies of the results from fixed effects models. Therefore, we decided to ignore these limited interaction effects between "female" and "policy", and assumed that the variable "female" was still a "fixed effect" that can be eliminated by a fixed effects model.

Again we can use the fitted values to estimate the policy effect on health inequalities. Based on the results in Table 3, we can calculate the policy effect on absolute health inequality as the change in the rate differences: 10.96–9.14=1.82. This means that the policy has reduced the rate difference between low- and high-educated people by almost 2 percentage points. Similarly, we can calculate the policy effect on relative health inequality, which then results in the finding that the policy reduced the relative excess risk of low as compared to high educated people by more than 12%.

Instrumental variable analysis

We illustrate the instrumental variable approach with the cross-sectional data obtained after the policy. Again, we used gender as an unobserved confounder. In a straightforward regression analysis, exposure to the policy would be endogenous (as gender would determine exposure to the policy to some extent, and is now included in the error term), and as a consequence the estimated effect of policy on health would be biased. We therefore used an instrument, e.g. the "distance from the house of the respondent to the closest free medical service". For this to be a valid instrument, it should be clearly predictive of exposure to the policy, related to health via the policy (use of the free medical service) only, and not be related to any unmeasured confounder (information about the construction of the instrumental variable used in our analyses is available upon request).

The instrumental variable analysis was conducted in a two-stage least squares regression. The basic idea of this analysis in our example was to first regress the policy exposure on the instrumental variable in order to capture the variation in policy exposure induced by the instrument, and to subsequently regress the health outcome on the predicted values for policy exposure. The instrumental variable analysis with logistic regression cannot be easily conducted in Stata, and therefore we used probit regression (specifically "ivprobit"). The coefficients from the probit regressions were transformed into marginal effects to make them comparable to those from linear regressions.

While the approach is intuitively easy if stratified by education, it is more complicated for an analysis using the interaction between policy and education. Because exposure to the policy is endogenous, the interaction between education and policy exposure is endogenous as well. This requires an instrument for the interaction terms as well; we used the interaction between education and distance from home to the closest free medical service for this purpose. In the first step of the two stage regression, both instruments predict exposure to the policy as well as the interaction between education and exposure to the policy. The predicted values are then used in the second stage of regression, resulting in unbiased effects of exposure to the policy and the interaction between policy exposure and education on health.

Table 2 shows that the absolute policy effect is larger among the low educated (β=-0.050, 95% CI [-0.063; -0.037]) than among the high-educated (β=-0.020, 95% CI [-0.029; -0.011]).

The interaction term for low education and policy was statistically significant (β=-0.036; 95% CI [-0.057; -0.015]), which indicated that the policy effect was different between low- and high-educated people indeed.

As for the other methods, we used the predicted values from the regression analysis (in this case, the second stage of the analysis) to estimate the policy effect on health inequality (Table 3). Using the values in Table 3, we calculated the policy effect on absolute health inequalities as the change of the rate difference: 11.16–9.14=2.02%. This means that the policy reduced the rate difference between low- and high-educated people by 2 percentage points. Similarly, we calculated the policy effect on relative health inequalities as the change of the rate ratio and found that the policy reduced the relative excess risk of poor health among the low educated by almost 13%.

Regression discontinuity

To illustrate the application of regression discontinuity, we had to create a new dataset. The main reason was the need to create a "threshold", and thereby to introduce a new variable, distinguishing persons who could receive the policy from those who were not eligible for it. For this purpose, we used income: those with a household income of less than 2000 euros per month could receive the free medical care, whereas those with higher incomes were not eligible to receive the free medical care. We assumed that the sharp threshold of 2000 euro resulted in a 'sharp' regression discontinuity, without changing the effect of income on health. Because low educated people generally tend to have lower household incomes, more low educated people were exposed to the policy. The imposed policy effect was still a reduction of the prevalence of poor health by 30% regardless of education level. The dataset created contained cross-sectional data after the implementation of the policy. Details about the generation of the data for the regression discontinuity are available upon request.

In a stratified analysis, this was modeled as follows for individuals in higher and lower educational groups separately:

$$Health_i = \beta_0 + \beta_1 (income_i - 2000) + \beta_2 policy_i + \mu_i$$

where $health_i$ is health of individual i, β_0 is the intercept, β_1 and β_2 are regression coefficients, and μ_i is the error term.

Individual-level health was still the health outcome. The value for the variable "*policy*" was 1 if the individual's monthly income was equal to or less than 2000 euro per month. The regression coefficient β_1 reflects the average effect of income on health. The regression coefficient β_2 reflects the discontinuity in health which was caused by the implementation of the policy. The analysis was done among individuals whose monthly income is around 2000 (e.g. individuals whose monthly income is between 1800 and 2200). Using logistic regression, the odds ratio resulting from the coefficient for the variable "policy" (β_2) measured the higher or lower odds of having poor health between those exposed to the policy and those not exposed to the policy. Table 2 shows that the relative policy effect is similar for low educated people (OR = 0.678, 95% CI [0.495; 0.929]) and for high educated people (OR = 0.687, 95% CI [0.483; 0.977]). Approximately, this is the 30% chance of reversing from poor to good health regardless of education level, as imposed in the data.

Regression discontinuity analysis also allows us to introduce interaction terms:

$$Health_i = \beta_0 + \beta_1 (income_i\text{-}2000) + \beta_2 policy_i + \beta_3 low\text{-}edu_i + \beta_4 (low\text{-}edu_i{*}(income_i\text{-}2000)) + \beta_5 (low\text{-}edu_i{*}policy_i) + \mu_i$$

where *health*$_i$ is health of individual i, β_0 is the intercept, $\beta_1 - \beta_5$ are regression coefficients, and μ_i is the error term.

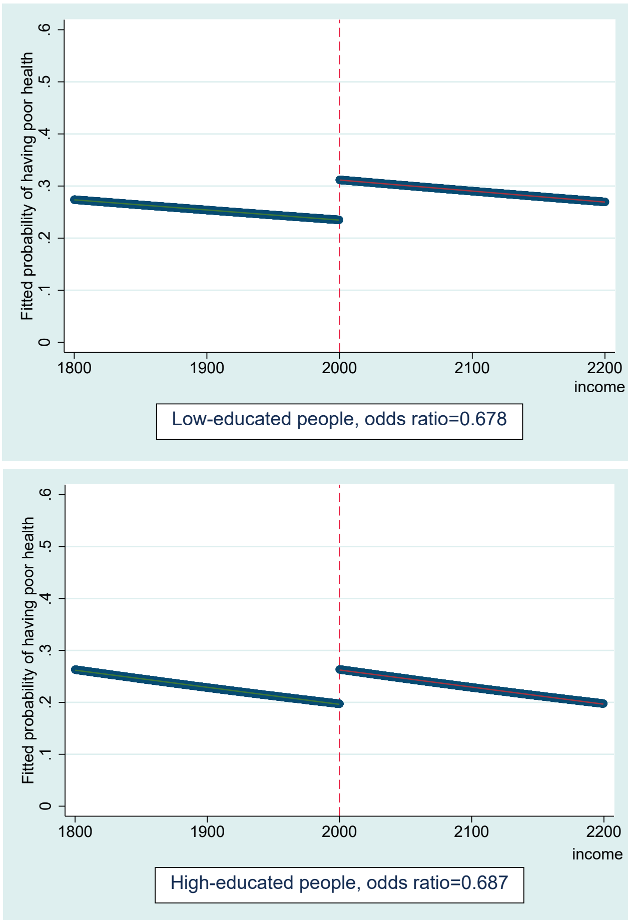


Figure 1 Results from regression discontinuity by education

As shown in Table 2, the interaction term for low education and policy ("*low-edu*policy*") is statistically insignificant (OR=0.987, 95% CI [0.615; 1.583]), which indicates that the policy effect is not statistically different between low- and high-educated people.

Results from the regression discontinuity analysis are also reported in a graphical way (Figure 1). In Figure 1, similar discontinuities around the income level of 2000 euro per month were observed among low- and high-educated people. This indicates similar instant policy effects. In our example, although the policy effects were independent of educational level, more low-educated people were exposed to the policy, leading to a decreased health inequality. However, this cannot be shown in the figure.

Again we can use the fitted values to estimate the policy effect on health inequalities and follow the same steps to calculate the changes in absolute and relative inequalities as a result of the policy. However, as the analysis was only performed based on the observations around the cutoff point of 2000 euro per month (e.g. 1800 – 2200 euro per month), we could not produce the policy effect on health inequalities among the whole population. This is a characteristic of the regression discontinuity method, and should not be seen as a failure of the example. Given that we generated a different setting for this method and the estimated policy effects only represented the policy effects among a part of the whole population, the calculated policy effects on health inequalities were not comparable to those from other methods and we therefore did not present them in Table 3 and 4.

Interrupted time series

To illustrate this method, we generated a time-series dataset which contained 40 years of observations. The quantitative characteristics of the dataset are similar to those used in the other calculation examples. Because this method (in our example) uses aggregate data, we could consider our health outcome, the prevalence of poor health, to be continuous (as opposed to binary in the other examples). For low educated people, the prevalence of poor health decreased by around 0.1 percentage points each year before the policy. For high educated people, the prevalence of poor health decreased by around 0.2 percentage points each year before the policy. The policy was implemented half way during the period of observation (i.e. year 20). For reasons of simplicity, we assumed that the policy affected the level of health (the intercept) immediately after its implementation. Details about the way of generating the data are available upon request.

In a stratified analysis, the model used for individuals in higher and lower educational groups separately was:

$$Health_t = \beta_0 + \beta_1\, year_t + \beta_2\, policy_t + \beta_3\, (year\ after\ policy)_t + \mu_t$$

where $health_t$ is the prevalence of self-assessed health, β_0 is the intercept, β_1 and β_2 are regression coefficients, and μ_t is the error term.

The variable "*year*" represented the calendar years and ranged from 1 to 40. The variable "*policy*" was a dummy variable with value 1 if it was larger than 20, and value 0 if it is smaller or equal to 20. The variable "year after policy" was the number of years after the implementation of the policy. In the regression, the coefficient of "*year*" (β_1) indicated the natural trend before the policy. The coefficients for "*policy*" and "*year after policy*" represented the change in the intercept and the change in the slope due to the policy.

Figure 2 presents the results of the interrupted time-series analysis, stratified by education. As mentioned above, aggregated data were used, which already incorporated the effect of

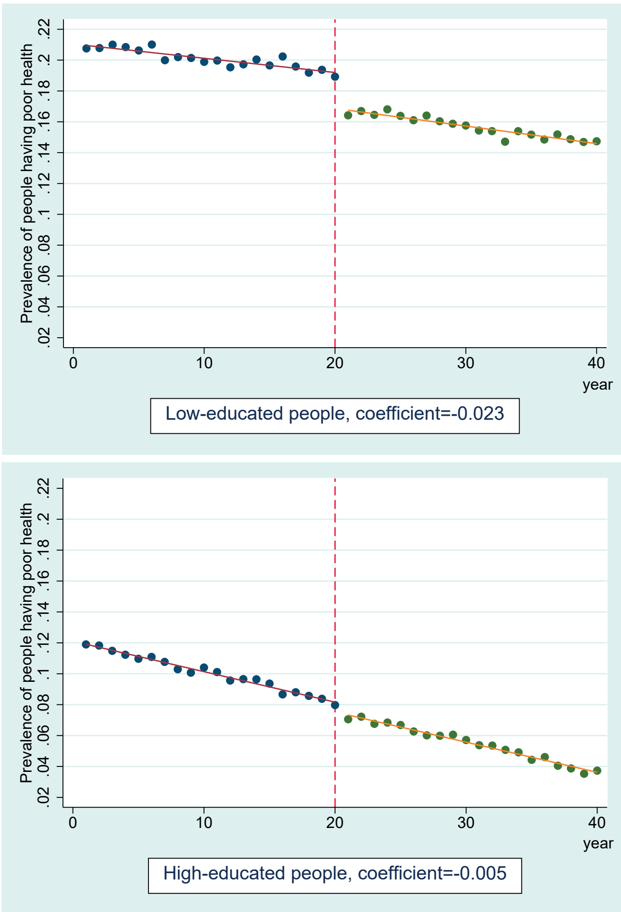


Figure 2 Results from interrupted time-series by education

both the real policy effect on the exposed people and the proportion of exposed high and low educated. Since more low educated people were exposed, an instant effect on reducing health inequalities was observed, indicated by a larger drop in the prevalence of poor health among the low-educated people directly after the implementation of the intervention in year 20. As shown in table 2, the policy reduced the prevalence of poor health for low educated people immediately by 2.3% points and it reduced the prevalence of poor health for high educated people immediately by 0.5% points.

Interrupted time-series analysis also allows us to introduce interaction terms:

$$Health_t = \beta_0 + \beta_1 year_t + \beta_2 policy_t + \beta_3 (year\ after\ policy)_t + \beta_4 low\text{-}edu_t + \beta_5 (low\text{-}edu_t{*}year_t) + \beta_6 (low\text{-}edu_t{*}policy_t) + \beta_7 (low\text{-}edu_t{*}(year\ after\ policy)_t) + \mu_t$$

where $health_t$ is the prevalence of self-assessed health, β_0 is the intercept, $\beta_1 - \beta_7$ are regression coefficients, and μ_t is the error term.

The coefficients for "*low-edu*policy*" represent the change in the intercept due to the policy. As shown in Table 2, the interaction "*low-edu*policy*" is statistically significant (coefficient=-0.019), which suggests that the policy effect is larger among low-educated people.

As before, using the values in Table 3 we calculated the policy effect on absolute health inequality as the change of the rate difference: 10.99–9.14=1.85. This means that the policy reduced the rate difference between low- and high-educated people by almost 2 percentage points. Similarly, we calculated the policy effect on relative health inequality as the change of the rate ratio, which resulted in the finding that the policy has reduced the relative excess risk among the low educated by almost 12%.

DISCUSSION

Summary of findings

This study demonstrated that all seven quantitative analytical methods identified can be used to quantify the equity impact of policies on absolute and relative inequalities in health by conducting an analysis stratified by socioeconomic position. Further, all but one (propensity score matching) can be used to quantify equity impacts by inclusion of an interaction term between socioeconomic position and policy exposure.

Methodological considerations

In our example, we assessed the effects of the policy in stratified analysis, and modeled it by including an interaction term between policy exposure and education. Apart from our finding that an interaction term could not be included in propensity score matching and appeared to be slightly more complicated in instrumental variable analysis, some differences between both approaches have to be considered before deciding which approach to use. Stratification by education is intuitively attractive; the method, however, requires additional analyses to statistically test whether the policy effects for higher and lower socioeconomic groups differ. Comparing the overlap in confidence intervals provides some further insight, but is still not a formal test [33]. Further, in our simple example, we only had two levels for our indicator of socioeconomic position, which made stratification easy. Including more levels of socioeconomic position results in smaller strata, with loss of statistical power as a consequence. Moreover, some indicators of socioeconomic position can be measured on a continuous scale, such as number of years of education, or household income. Categorizing continuous values requires making (arbitrary) decisions, and results in a loss of information. Analyses using an interaction term allow indicators of socioeconomic position to be continuous variables. The results, however, can sometimes be more complex to interpret. For example, one issue to consider is that whether the effect of the policy on health inequalities changes in a linear way with an increase of one unit of the socioeconomic indicator.

Caution is needed when interpreting the results from the instrumental variable approach. Under certain conditions, the instrumental variable reveals a local average treatment effect [34], namely the intervention effect among individuals affected by the observed changes due to the instrument ("compliers"). It is a local parameter since it is specific to the population defined by the instrument [28]. Different instrumental variables, although all valid, will be associated with different local average treatment effect estimators and the population of corresponding compliers cannot be identified in the data [35]. Thus, when we apply it to health inequalities, for example in stratified analysis, the estimated policy effects are the effects among the corresponding compliers within each socioeconomic group given a set of instruments. The generalization of the conclusion to the whole population or to other populations is normally uncertain. However, when the change of policy is used as the instrument for the exposure, the local average treatment effect might be extremely useful, since it focuses on an important subpopulation whose exposure status is changed by the policy and may provide an informative measure of the impact of the policy [28].

The above mentioned problem of a low external validity also applies to regression discontinuity. Analysis are only performed based on the observations around a cutoff point (e.g. 1800 – 2200 euro per month in our example), and as a result, the method does not produce a policy effect on health inequalities among the whole population.

Persons were either exposed or unexposed in our fictitious example; we did not include the possibility of graded exposure to the policy. Whereas regression adjustment would allow a graded exposure relatively easy, for other techniques this may be more complex (although not impossible), such as for propensity score matching [36].

Which method to use depends to a large extent on data availability (e.g. whether cross-sectional or longitudinal data are available) and the nature of the confounders in the analysis (whether observable or not, and whether time-variant or not). The appropriateness of the preferred methods further depends on the degree to which underlying assumptions are met. For example, instrumental variable analysis requires strong assumptions, and violations can lead to biased estimated [21,26]. Similarly, difference-in-differences analysis is based on the assumption of common trends between higher and lower educated in the absence of the policy.

Interpretation

Although the methods described seem quite different, they actually try to achieve the same aim, which is constructing counterfactual outcomes for exposed units had they not been exposed to the policy [37]. Doing this in a convincing way is a key ingredient of any serious evaluation method [28]. For example, in well-designed randomized controlled trials, the control group is a perfect counterfactual for the exposed group, since the pre-intervention differences have been eliminated by the random assignment of intervention. In the same way, if the key assumption holds that selection bias disappears after controlling for observed characteristics [26], both regression adjustment and propensity score matching restore randomization to some extent. Similarly, both instrumental variable and regression discontinuity approaches aim at finding exogenous factors which can fully or partly determine the assignment to a policy; as such, this mimics randomization to some extent. If in a difference-in-differences analysis, the trend over time is the same for unexposed and exposed units of analysis, the change in the unexposed unit can be potentially used as the counterfactual. When longitudinal data is available, the fixed effects model uses the exposed unit's own history prior to treatment to construct the counterfactual [20]. Likewise, the time trend of the exposed unit before the policy implementation is utilized as the counterfactual part when time-series data are available.

We constructed our fictitious data in the way that people with a low socioeconomic position were more exposed to the policy, but the policy effect among the exposed was equal between socioeconomic groups. In reality, health inequalities might also be reduced in cases where people with different socioeconomic positions are equally exposed to the policy, but where the policy effect among those exposed is larger among people with a low socioeconomic position. It can also happen that people with a low socioeconomic position are more exposed to the policy, and where the policy effect among those exposed is larger among people with a low socioeconomic position. The process of analysis and the interpretation of the results however, are similar for these cases. Moreover, although we mainly constructed the examples using individual level data (except for interrupted time-series), some methods can also be used with aggregated level data. For example, a fixed effects model can also be applied with country-level longitudinal data.

In this paper, we performed the analysis based on a standard setting of each method. The analysis however, can be easily extended to more complicated examples. We only used few covariates in our analysis, but more can be incorporated. It is also possible to use the methods with longitudinal or repeated cross-sectional data with multiple periods, continuous health outcomes, and more than one instrument. Moreover, extensions of the models described here with relaxed assumptions have been applied, such as quasi difference-in-differences model [38], changes-in-changes model [39] and dynamic fixed effects model [40,41]. The extended models are not covered by this paper, but the general way of applying them for assessing the impact of policy on health inequalities is similar to the standard models. Combining methods in one study is also possible. For example, some papers recommend to combine regression adjustment and matching by using weighted regression based on propensity score to reduce bias [31,42]. In this way, matching can be combined with many techniques such as difference-in-differences analysis and fixed effects model [43,44]. Another example is incorporating instrumental variables into fixed effect models to tackle the potential measurement error [26].

This study demonstrated quantitative tools to assess if and to what extent natural policy experiments impact upon socioeconomic inequalities in health. While our approach offers further insight in whether effects resulted from a policy effects and/or and the size of the populations exposed, it does not offer in-depth insight into how effects were achieved. Quantitatively, (causal) mediation analyses could be used to assess explanations for potential effects, whereby the effect of the policy experiment on potential determinants could be assessed, as well as the effects of the potential explanatory factors on the outcome [45]. Future research should explore to what extent mediation analysis can be used to assess explanations of the impact of NPE's in inequalities in health. Simultaneously, qualitative approaches can be used to further examine the processes leading to an impact [46,47].

The demonstrated possibility to use the techniques described in this paper for studying the impact of NPE's on socioeconomic inequalities raises the question as to whether all policy evaluations should include an evaluation of the equity impact. Researchers evaluating an equity impact of interventions are often criticized by statisticians for conducting unreliable (underpowered) analyses; those who don't are at the same time however, criticized by policymakers in need of evidence what works to close the gap in health between socioeconomic groups [48]. Following guidelines in which a logic model includes theoretically plausible mechanisms for a reduction on inequalities in health, and in which statistical power is not a real issue, we recommend that an equity impact analyses should be an integral part of any policy experiment.

In conclusion, application of methods commonly applied in economics and econometrics can be applied to assess the equity impact of natural policy experiments. The low external validity of results from instrumental variable and regression discontinuity makes these methods less desirable for assessing policy effects on population-level health inequalities. Increased used in social epidemiological research will help to build an evidence base to support policy making in this area.

REFERENCES

1. Mackenbach JP. Health inequalities: Europe in profile. 2006. http://ec.europa.eu/health/ph_determinants/socio_economics/documents/ev_060302_rd06_en.pdf. Accessed 6 January 2016.
2. Mackenbach JP. Socioeconomic inequalities in health in high-income countries: the facts and the options. In: Detels R, Gulliford M, Karim QA, Tan CC, editors. Oxford Textbook of Global Public Health. Oxford, Oxford University Press, 2015, p.106–126.
3. Mackenbach JP, Stirbu I, Roskam AJ, Schaap MM, Menvielle G, Leinsalu M, et al. Socioeconomic inequalities in health in 22 European countries. N Engl J Med. 2008;358:2468–81.
4. Commission on Social Determinants of Health, Closing the gap in a generation. Health equity through the social determinants of health. Geneva: World Health Organization, 2008.
5. Mackenbach JP, Bakker MJ. Tackling socioeconomic inequalities in health: analysis of European experiences. Lancet. 2003;362:1409–1414.
6. Judge K, Platt S, Costongs C, Jurczak K. Health inequalities: a challenge for Europe. 2006. http://ec.europa.eu/health/ph_determinants/socio_economics/documents/ev_060302_rd05_en.pdf. Accessed 2 Feb 2015.
7. Mackenbach JP, Stronks K. The development of a strategy for tackling health inequalities in the Netherlands. Int J Equity Health. 2004;3:11.
8. Mackenbach JP. Has the English strategy to reduce health inequalities failed? Soc Sci Med. 2010;71:1249–53.
9. Mackenbach JP. Socioeconomic inequalities in health in western Europe: from description to explanation to intervention. In: Siegrist J and Marmot M, editors. Social inequalities in health: New evidence and policy implications. New York: Oxford University Press, 2006, p223–250.
10. Kulik MC, Hoffmann R, Judge K, Looman C, Menvielle G, Kulhánová I, et al. Smoking and the potential for reduction of inequalities in mortality in Europe. Eur J Epidemiol. 2013;28:959–71.
11. Eikemo TA, Hoffmann R, Kulik MC, Kulhánová I, Toch-Marquardt M, Menvielle G, et al. How can inequalities in mortality be reduced? A quantitative analysis of 6 risk factors in 21 European populations. PLoS One. 2014, 9:e110952.
12. Schrijvers CTM, van de Mheen HD, Stronks K, Mackenbach JP. Socioeconomic inequalities in health in the working population: the contribution of working conditions. Int J Epidemiol. 1998; 27:1011–8.
13. Bambra C, Gibson M, Sowden AJ, Wright K, Whitehead M, Petticrew M. Working for health? Evidence from systematic reviews on the effects on health and health inequalities of organisational changes to the psychosocial work environment. Prev Med. 2009;48:454–61.
14. Kaikkonen R, Rahkonen O, Lallukka T, Lahelma E. Physical and psychosocial working conditions as explanations for occupational class inequalities in self-rated health. Eur J Public Health. 2009;19:458–63.
15. Nutbeam D. How does evidence influence public health policy? Tackling health inequalities in England. Health Promot J Aust. 2003;14:154–8.
16. Bonell CP, Hargreaves J, Cousens S, Ross D, Hayes R, Petticrew M, et al. Alternatives to randomisation in the evaluation of public health interventions: design challenges and solutions. J Epidemiol Comm Health. 2011;65:582–7.
17. Black N. Why we need observational studies to evaluate the effectiveness of health care. BMJ. 1996;312:1215–8.
18. Petticrew M, Cummins S, Ferrell C, Findlay A, Higgins C, Hoy C, et al. Natural experiments: an underused tool for public health? Public Health. 2005;119:751–7.
19. Roe BE, Just D. Internal and external validity in economics research: Tradeoffs between experiments, field experiments, natural experiments, and field data. Am J Agricult Economics. 2009; 91:1266–1271.

20. Jones AM, Rice N. Econometric evaluation of health policies. In: Glied S, Smith PC editors. The Oxford Handbook of Health Economics. Oxford, Oxford University Press, 2011, p890–923.
21. Oakes JM, Kaufman JS. Methods in social epidemiology. 2006, San Francisco: John Wiley & Sons.
22. Harper S, Strumpf EC, Burris S, Davey Smith G, Lynch J. The effect of mandatory seat belt laws on seat belt use by socioeconomic position. J P Anal Manage 2014;33:141–161.
23. Ludwig J, Miller DL. Does head start improve children's life chances? Evidence from a regression discontinuity design. Quart J Economics. 2007;122:159–208.
24. Mackenbach JP, Kunst AE. Measuring the magnitude of socio-economic inequalities in health: an overview of available measures illustrated with two examples from Europe. Soc Sci Med. 1997;44: 757–71.
25. Wooldridge JM. Econometric Analysis of Cross Section and Panel Data. Cambridge: MIT Press, 2001.
26. Angrist JD, Pischke JS. Mostly harmless econometrics: an empiricist's companion. Princeton: Princeton University Press, 2009.
27. Shadish WR, Cook TD, Campbell DT. Experimental and quasi-experimental designs for generalized causal Inference. Houghton Mifflin Company, 2002.
28. Blundell R, Dias MC. Alternative approaches to evaluation in empirical microeconomics. J Hum Res. 2009:565–640.
29. Dinardo, J. Lee DS. Program evaluation and research designs. Cambridge, MA: NBER Working Paper 16016, 2010.
30. Cameron AC, Trivedi PK. Microeconometrics: Methods and Applications. New York: Cambridge University Press, 2005.
31. Imbens GW, Wooldridge JM. Recent developments in the econometrics of program evaluation. J Economic Lit. 2009;47:5–86.
32. Cousens S, Hargreaves J, Bonell C, Armstrong B, Thomas J, Kirkwood BR. et al. Alternatives to randomisation in the evaluation of public-health interventions: statistical analysis and causal inference. J Epidemiol Comm Health. 2011; 65:576–81.
33. Cumming G. Inference by eye: reading the overlap of independent confidence intervals. Stat Med. 2009;28:205–20.
34. Imbens GW, Angrist JD, Identification and Estimation of Local Average Treatment Effects. Econometrica. 1994;62: 467–475.
35. Gangl M. Causal Inference in Sociological Research. Ann Rev Sociol. 2010; 36:21–47.
36. Hirano K, Imbens GW, Ridder G. Efficient estimation of average treatment effects using the estimated propensity score. Econometrica. 2003;71:1161–1189.
37. Rubin DB. Estimating causal effects of treatments in randomized and nonrandomized studies. J Educat Psychol. 1974;66:688–701.
38. Lee, M., Quasi-differences in differences with panel data: complete pairing and linear-model LSE. 2012: Working paper, http://www.econ.sinica.edu.tw/mobile/conferences/contents/2013090215153345654/?MSID=2013102818183246335. Accessed 3 July 2015.
39. Athey S, Imbens GW. Identification and inference in nonlinear difference-in-differences models. Econometrica. 2006;74:431–497.
40. Arellano M, Bond S. Some tests of specification for panel data - Monte-Carlo evidence and an application to employment equations. Rev Econom Studies. 1991;58:277–297.
41. Bun MJG, Carree MA. Bias-corrected estimation in dynamic panel data models with heteroscedasticity. Economics Letters. 2006, 92: 220–227.
42. Freedman DA, Berk RA. Weighting regressions by propensity scores. Evaluation Review. 2008. 32:392–409.

43. Nichols A. Causal inference with observational data. Stata Journal. 2007;7: 507–541.
44. García-Gómez P, van Kippersluis H, O'Donnell O, van Doorslaer E. Long-term and spillover effects of health shocks on employment and income. J Human Res. 2012; 48:873–909.
45. Vanderweele T. Explanation in causal inference: Methods for mediation and interaction. New York: Oxford University Press, 2015.
46. Whitehead MA. A typology of actions to tackle social inequalities in health. J Epidemiol Comm Health. 2007;61:473–8.
47. Morestin F, Gauvin FP, Hogwe M, Benoit F. Method for synthesizing knowledge about public policies 2010: National Collaborating Centre for Healthy Public Policy. http://www.ncchpp.ca/docs/MethodPP_EN.pdf. Accessed 1 March 2016.
48. Petticrew M, Tugwell P, Kristjansson E, Oliver S, Ueffing E, Welch V. Damned if you do, damned if you don't: subgroup analysis and equity. J Epidemiol Comm Health. 2012; 66:95–8.
49. Chaloupka FJ, Pacula RL. Sex and race differences in young people's responsiveness to price and tobacco control policies. Tob Control. 1999;8:373–7.
50. Melhuish E. Effects of fully-established Sure Start Local Programmes on 3-year-old children and their families living in England: a quasi-experimental observational study. Lancet. 2008; 372:1641–1647.
51. Gruber J. Youth smoking in the U.S.: prices and polices. 2000: NBER Working paper No. 7506. http://www.nber.org/papers/w7506. Accessed 19 August 2014.
52. Barr B, Bambra C, Whitehead M. The impact of NHS resource allocation policy on health inequalities in England 2001–11: longitudinal ecological study. BMJ. 2014; 348; g3231.
53. Evans WN, Lien DS. The benefits of prenatal care: evidence from the PAT bus strike. J Econometrics. 2005;125:207–239.
54. Federico B, Mackenbach JP, Eikemo TA, Kunst AE. Impact of the 2005 smoke-free policy in Italy on prevalence, cessation and intensity of smoking in the overall population and by educational group. Addiction. 2012;107:1677–86.
55. Kramer D, Droomers M, Jongeneel-Grimen B, Wingen M, Stronks K, Kunst AE. The impact of area-based initiatives on physical activity trends in deprived areas; a quasi-experimental evaluation of the Dutch District Approach. Int J Behav Nutr Phys Act. 2014;11:36.
56. Kuipers MA, Nagelhout GE, Willemsen MC, Kunst AE. Widening educational inequalities in adolescent smoking following national tobacco control policies in the Netherlands in 2003: a time-series analysis. Addiction. 2014;109:1750–1759.

Chapter 6

Did the English strategy reduce inequalities in health?

Yannan Hu, Frank J. van Lenthe, Ken Judge, Eero Lahelma, Giuseppe Costa, Rianne de Gelder, Johan P. Mackenbach. Did the English strategy reduce inequalities in health? A difference-in-differences analysis comparing England with three other European countries. (submitted)

ABSTRACT

Background

Between 1997 and 2010, the English government pursued an ambitious programme to reduce health inequalities, the explicit and sustained commitment of which was historically and internationally unique. Previous evaluations have produced mixed results. None of these evaluations have, however, compared the trends in health inequalities within England with those in other European countries. We carried out an innovative analysis to assess whether changes in trends in health inequalities observed in England after the implementation of its programme, have been more favourable than those in other countries without such a programme.

Methods

Data were obtained from nationally representative surveys carried out in England, Finland, the Netherlands and Italy for years around 1990, 2000 and 2010. A modified difference-in-difference approach was used to assess whether trends in health inequalities in 2000–2010 were more favourable as compared to the period 1990–2000 in England, and the changes in trends in inequalities after 2000 in England were then compared to those in the three comparison countries. Health outcomes were self-assessed health, long-standing health problems, smoking status and obesity. Education was used as indicator of socioeconomic position.

Results

After the implementation of the English strategy, more favourable trends in some health indicators were observed among low-educated people, but trends in health inequalities in 2000–2010 in England were not more favourable than those observed in the period 1990–2000. For most health indicators, changes in trends of health inequalities after 2000 in England were also not significantly different from those seen in the other countries.

Conclusions

In this rigorous analysis comparing trends in health inequalities in England both over time and between countries, we could not detect a favourable effect of the English strategy. Our analysis illustrates the usefulness of a modified difference-in-difference approach for assessing the impact of policies on population-level health inequalities.

BACKGROUND

Between 1997 and 2010, the English government made reducing health inequalities part of its core political programme [1]. It developed and implemented a strategy that – in the government's own words – was "the most comprehensive programme of work to tackle health inequalities ever undertaken in this country" [2]. This contained a number of comprehensive and coordinated policies, which were clearly documented and monitored in a series of reports [2–9].

The English strategy to reduce health inequalities was shaped in two steps [10, 11], of which the first was taken in 1999, when the Department of Health issued "Reducing Health Inequalities: an Action Report" [2]. This set out national actions across a broad front including raising living standards and tackling low income, family support policies, tax-reduction and long-term care for the elderly, anti-smoking policies, improving early education ("Sure Start") and promoting healthy communities, as well as some broader policies in the areas of education, employment and housing. It largely followed the recommendations of the Acheson committee which were based on the best available evidence in the late 1990s [5].

The second step followed in 2003 when a more focused strategy was laid down in "Tackling Health Inequalities: a Program for Action" [8]. Following an interdepartmental review of progress [9], it announced a revised strategy which contained 12 "headline indicators" (i.e., specific targets for intermediate outcomes) and 82 "departmental commitments", that together were expected to ensure the timely delivery of two new overall targets: "to narrow the gap in life expectancy between areas and the difference in infant mortality across social classes by 10% in 2010". The revised strategy also had a stronger emphasis on "downstream" policies than the 1999 Action Report, such as reducing smoking in manual social groups, managing other risks for coronary heart disease and cancer (e.g., poor diet and obesity, physical inactivity, hypertension), improving housing quality by tackling cold and dampness, and reducing accidents at home and on the road. The total budget exceeded £20 billion [8].

A remarkable series of reports systematically assessing and reviewing progress in achieving "headline indicators" and fulfilling "departmental commitments" followed. The high level of government commitment to reducing health inequalities was matched by an equally remarkable commitment to critically review, revise and then re-review its policies [10]. It has been noted that, quite clearly, the English strategy to reduce health inequalities was both historically and internationally unique [12, 13].

When the strategy came to an end, however, after the election in 2010 of a new government, the results turned out to be less encouraging than most people had expected. On the one hand, all the departmental commitments were fulfilled, indicating that all elements of the strategy as originally planned had been implemented, from "Sure Start" to the creation of sports facilities, from neighbourhood renewal programmes to smoking cessation support, and from improving access to health care services to reducing fuel poverty [6, 7, 11]. This by itself was a great achievement, but only some of the headline indicators showed reduced inequalities, in terms of smaller relative or absolute inequalities in intermediate outcomes like educational outcomes, child poverty or cardiovascular risks. Others, including those that matter for inequalities in life expectancy and infant mortality, such as on primary care, diet and smoking, suggested stable or even increased inequalities between socioeconomic groups [6, 7, 11, 14]. There was no evidence at all for a reduction of inequalities in infant mortality or life expectancy, as stipulated in the overall targets [3, 11].

However, one potential problem with most existing evaluations of the English strategy is that these mainly investigated the trends in health inequalities within England after the implementation of the strategy, sometimes as compared to the trends before the implementation of the strategy, but never as compared to the trends in other countries. Given that a widening of health inequalities over the past decades has been observed in many European countries [15, 16], a relevant question is whether health inequalities in England have perhaps widened less than elsewhere thanks to the English strategy, as compared to the trends in other countries which have taken less action to tackle health inequalities.

This paper therefore extends existing evaluations by first assessing the change in trend in health inequalities in England between 1990–2000 (during which the English strategy had not yet been generally implemented) and 2000–2010 (during which the main effects of the strategy could be expected), and then comparing this change in trend, if any, with the trend change occurring in 3 comparison countries.

For comparison we selected countries that were in a similar stage of awareness of health inequalities, but that had not implemented a national strategy to tackle health inequalities. Our selection of countries was guided by several studies that have characterized national policy developments in this area in European countries [12, 17, 18]. Based on a strong tradition of measuring and investigating health inequalities, Finland launched a national public health programme with explicit priorities for reducing health inequalities, which were elaborated in a specific programme on reducing health inequalities conducted between 2008 and 2011. No resources however, were allocated for the latter programme except for doing more research, and it was not implemented in practice [19]). In terms of Whitehead's action spectrum, the Netherlands was already in a phase of "structured development" in the late 1990s (Whitehead, 1998). After two five-year research programmes, a national programme to tackle health inequalities was proposed in 2001, but was never implemented, mainly because of a sudden change in government [20, 21]. In Italy, a country in a stage of "concern" with regard to health inequalities according to Whitehead's action spectrum, a serious level of awareness was evidenced by national research programmes, but again, no coordinated action to tackle health inequalities was taken by the national government [12, 18]). We did not choose other parts of the United Kingdom (Scotland, Wales and Northern Ireland) as the comparison countries because, although they may be more comparable to England, they also underwent significant policy changes to reduce health inequalities after 2000 [18].

For our evaluation of the English strategy, we used individual-level data from national health interview surveys, which allowed us to study trends in inequalities in health (self-assessed health, long-standing health problems) and in determinants of health (smoking and obesity). Like mortality and life expectancy, self-assessed health and long-standing health problems are generic health outcomes for which socioeconomic inequalities have been extensively documented [22, 23]. Many elements of the English strategy may have contributed to a favourable change in the trend in inequalities for these health outcomes, including improvements in material living conditions, health-related behaviours like smoking, diet and exercise, and access and quality of care [2, 8]. Smoking and obesity were directly targeted by the English strategy, which explicitly aimed to reduce inequalities in smoking (e.g., by increasing access to smoking cessation services) and to improve diet and physical activity [8].

We aimed to investigate the effect of the English strategy by assessing whether trends in inequalities in these health outcomes were more favourable in 2000–2010 as compared to those in 1990–2000 in England, and whether the changes in trends in inequalities after 2000 in England were more favourable than those in the three comparison countries.

METHODS

Data

We obtained nationally representative health surveys or multipurpose surveys with a health component from England, Finland, the Netherlands and Italy (Table 1). The available years of surveys differed slightly between countries, but all of them were around 1990, 2000 or 2010. The selected surveys were identical over time for England and Finland, but not for the Neth-

Table 1 Countries included in the analysis and sources of data

Country	Survey year	Survey names
England	1991–1992; 2000; 2010	Health Survey for England
Finland	1989; 1999; 2009	Health Behaviour and Health
The Netherlands	1990	Ongoing Survey of Living Conditions (DLO)
	2000; 2009	Permanent Survey of Living Conditions (POLS)
Italy	1990	Multipurpose Family Survey
	2000	Health and Health Care Utilization
	2010	Multipurpose Family Survey - Aspects of daily living

Note: The Finnish data used in this study are the data combined from the two Finish studies: "Health behaviour and health among Finnish adult population (AVTK)", which includes respondents who are 15–64 years old, and "Health behaviour and health among the Finnish elderly (EVTK)", which includes respondents who are older than 64 years.

erlands and Italy. Given that our main aim was to investigate the changes in trends in health inequalities between 1990–2000 and 2000–2010 in England and whether these changes in trends were more favourable in England than in the three comparison countries, we focused on the comparability over time within each country and considered the risk of bias due to between-country variations in data collection to be limited. All the selected surveys had a high degree of comparability within-country over time in the aspects of sampling strategy, survey questions and answers, thus could be used to analyse the trends over time [24–30]. Details on data collection in each country are reported in the appendix (table A1). The age range used in the analyses was 16–79 years. Older respondents were excluded to avoid the potential bias caused by the exclusion of institutionalized population in most surveys. Survey weights which were designed to make the sample representative of the whole population were available in some countries and years. Specifically, weights were available for the data from England in 2010, the Netherlands in 1990/2000/2010 and Italy in 1990/2000. Most of the weights are normal sampling weights, which make the samples nationally representative. Weights in the Dutch survey in 2000 and 2010 are "expansion" weights, which are used so that the weighted data reflect the size of the total Dutch population. In order to be comparable to those in the other years and other countries, weights in the Dutch survey in 2000 and 2010 were scaled in our analysis (i.e. divide each year by the mean weight).

Based on data availability, four health outcomes were chosen: self-assessed health, long-standing health problems, smoking status and obesity. Self-assessed health was generated based on a question which was framed in a way similar to "how is your health in general?", and was recoded into a binary variable indicating whether the respondent had less-than-good self-assessed health. Long-standing health problems was a binary variable measuring whether or not the respondent reported any long-standing health problems. Smoking status was measured as whether the respondent was a current smoker. Obesity was based on the body mass index (BMI) of 30 or higher, calculated from the measured or self-reported height and weight (kg/m^2). The precise survey questions and answer categories varied slightly between countries, but consistency over time was retained in all countries (appendix table A1).

Socioeconomic position was measured by the highest level of education completed or currently being attended by a person. It was harmonized on the basis of the International Standard Classification of Education (ISCED) and reclassified into 3 categories: levels 0–2 (no, primary or lower secondary education, considered "low-educated"), levels 3–4 (upper secondary and post-secondary non-tertiary education, considered "middle-educated"), levels 5–6 (tertiary education, considered "high-educated"). Details in the classification of education in England are reported in the appendix (table A2). Comparable indicators for other measures of socioeconomic position, such as occupational class or income level, were not available in all surveys.

Statistical methods

Our analysis started with a comparison of changes in health occurring between 1990–2000 (control condition) and those occurring between 2000–2010 (treatment condition) among low-educated people in England. We assessed whether there was a larger improvement of health among low-educated people in England after the introduction of the strategy than before the introduction of the strategy.

The model for this analysis can be written as:

$$outcomeist = \beta_0 + \beta_1 endyear_t + \beta_2 policyperiod_s + \beta_3 \mathrm{endyear}_t * policyperiod_s + X_{ist}$$

where $outcome_{ist}$ is one of the chosen health measures of individual i in period s and year t, β_0 is a constant, $endyear_t$ is a dummy indicating whether it is the end year of each period, $policyperiod_s$ is a dummy indicating whether it is the period 2000–2010 (treatment period), $endyear_t$*$policyperiod_s$ is the interaction between $endyear_t$ and $policyperiod_s$, X_{ist} represents the control variables which are age and sex. For the period 1990–2000, $policyperiod_s$ was 0, and $endyear_t$ was 0 for data from 1990 and was 1 for data from 2000. For the period 2000–2010, $policyperiod_s$ was 1, and $endyear_t$ was 0 for data from 2000 and was 1 for data from 2010. The $\beta_1$coefficient measures the trend in health in the control condition (i.e. trend in 1990–2000). β_2 measures the difference in the level of health between the control and treatment condition at the beginning (i.e. difference in health between the year 1990 (the beginning year of the control condition) and the year 2000 (the beginning year of the treatment condition)). β_3 is the key parameter (further referred to as "two-way interaction" parameter) that quantifies the difference in the trend between the two conditions.In order to make a causal interpretation, the assumption we need is that in the absence of the strategy, the trend in health among low-educated people in 2000–2010 (treatment condition) would have been the same as the trend in health in 1990–2000 (control condition).

In a second step, in order to assess whether there is a more favourable trend in health inequalities after the introduction of the strategy, we made a comparison between the changes in improvement of health between low- and high-educated people. Therefore, we introduced an additional difference, i.e. the difference between low- and high-educated people, into the regression, by adding the binary variable for education, and all possible interactions with education in the equation.

The model can be written as:

$$\begin{aligned} outcomeist = {} & \beta_0 + \beta_1\, endyear_{\mathrm{t}} + \beta_2\, policyperiod_s + \beta_3\, endyear_t * policyperiod_s \\ & + \beta_4\, ledu_{ist} + \beta_5\, ledu_{ist} * endyear_t + \beta_6\, ledu_{ist} * policyperiod_s + \beta_7 \\ & ledu_{ist} * endyear_t * policyperiod_s + X_{ist} \end{aligned}$$

Chapter 6

where a new variable $ledu_{ist}$ indicating whether the respondent is low-educated, and the interactions between $ledu_{ist}$ and other variables were added. Now β_7 is the key parameter (further referred to as "three-way interaction" parameter), which quantifies the difference between low and high educated in the difference in the trend between the two periods. In other words, this assessed whether the trend in health inequalities in England was different in the period 2000–2010 as compared to the period 1990–2000. In order to interpret this as the effect of the strategy, the assumption we need to make is that, in the absence of the strategy, the trend in health inequalities in both periods would have been the same. This model was also applied, independently, to the three comparison countries.

In the last and our main step, we added each of the comparison countries separately to the analysis of the English data, following the idea of "difference-in-differences analysis" [31, 32]. Our aim was to investigate whether the changes in trends in health inequalities between 1990–2000 and 2000–2010 were more favourable in England than those in the three comparison countries. The rationale, as mentioned in the introduction, is that even if there is no more reduction in health inequalities after the implementation of the strategy than before, the changes in trends in England could still be more favourable than those in other European countries that have done less to reduce health inequalities. Therefore, we pooled data from England and each of the comparison countries, and added an additional difference, i.e. the difference between England and the comparison country, into the regression, by adding a dummy for England and the corresponding interactions. Here the difference in trend in health inequalities between the period 1990–2000 and 2000–2010 in the comparison country was regarded as the control condition.

The model can be written as:

$$\begin{aligned} outcome_{istj} = {} & (\beta_0 + \beta_1\ endyear_{tj} + \beta_2\ policyperiod_{sj} + \beta_3\ endyear_{tj} * \\ & policyperiod_{sj} + \beta_4\ ledu_{istj} + \beta_5\ ledu_{istj} *\ endyear_{tj} + \beta_6\ ledu_{istj} * \\ & policyperiod_{sj} + \beta_7\ ledu_{istj} *\ endyear_{tj} * policyperiod_{sj} + X_{istj}) \\ & + (\beta_0' + \beta_1'\ endyear_{tj} + \beta_2'\ policyperiod_{sj} + \beta_3'\ endyear_{tj} * \\ & policyperiod_{sj} + \beta_4'\ ledu_{istj} + \beta_5'\ ledu_{istj} *\ endyear_{tj} + \beta_6'\ ledu_{istj} \\ & *\ policyperiod_{sj} + \beta_7'\ ledu_{istj} *\ endyear_{tj} *\ policyperiod_{sj} + X_{istj}) \\ & *\ england_j \end{aligned}$$

where $outcome_{istj}$ is one of the chosen health measures of individual i in period s, year t and country j. Now β_7', the coefficient of the quadruple interaction term "$ledu_{istj}*endyear_{tj}*policyperiod_{sj}*england_j$", is the key parameter of this model (further referred to as "four-way interaction" parameter), which quantifies the difference between the "three-way interaction" parameter of England and that of each comparison country. In other words, it assesses whether changes in trends in health inequalities observed in England after the implementation of its programme, were more favourable than those in other countries without such a programme. In order to make a causal interpretation, the assumption is that in the absence of the strategy, the changes in trends in health inequalities between the two periods in England would have been the same as those in the comparison countries.

Logistic regression was used in all the analyses. When the outcomes are non-linear, as in the case of binary outcomes, difference-in-difference models for non-linear models (such as logistic regression) are to be preferred [33]. The interpretation of the interaction terms in difference-in-differences logistic models is essentially similar to that in the more common linear models, except that they indicate the relative change of the odds of the health outcome in the treatment group relative to that in the control group, instead of the absolute change of the rate of the health outcome in the treatment group minus that in the control group. Robust standard errors were used to account for potential heteroskedasticity. Unweighted results are reported in the results section. Analysis with weighting factors when available are reported in the appendix as a sensitivity analysis (table A3).

All regression analyses were performed in Stata 13.1. Results with a p-value lower than 0.1 were regarded as significant. The specific significance level was indicated for each significant result. The coding of the variables and more explanations are reported in the appendix.

RESULTS

Summary statistics of key variables are presented in Table 2. Compared to the three comparison countries, the sample in England appeared to have a more equal distribution of the three education categories, a relatively lower proportion of less-than-good self-assessed health, a higher proportion of long-standing health problems, an average level of smoking prevalence, but a much higher rate of obesity.

The main results are reported in Table 3. The full model results are reported in the appendix. The “two-way interaction” parameter estimates for low-educated people in England show that more favourable trends after 2000 were found in all health measures, although not statistically significant for obesity. A favourable change in trend is shown by odds ratios (OR) that are smaller than 1. For example, although the odds of less-than-good self-assessed health increased during both periods (Figure 1.a), the increase of the odds in 2000–2010 was 24% less than that in 1990–2000 (OR=0.76, $p<0.01$). Similarly, although the odds of being a current smokers decreased during both periods (Figure 1.c), the decrease of the odds in 2000–2010 was 18% more than that in 1990–2000 (OR=0.82, $p<0.05$).

Table 3 also presents the “three-way interaction” parameter estimates for each country, which tested the differences in trends in health inequalities between 1990–2000 and 2000–2010. An odds ratio below 1.00 indicates that the trend in health inequalities was more

Table 2 Summary statistics of key variables, pooled for all years in each country

	England		Finland		The Netherlands		Italy	
	N	%	N	%	N	%	N	%
Number of respondents	22,442		14,296		18,353		204,963	
Gender								
male	10,255	46%	6,654	47%	8,712	47%	99,888	49%
female	12,187	54%	7,642	53%	9,641	53%	105,075	51%
Age								
16–25	3,057	14%	1,874	13%	2,816	15%	32,779	16%
26–35	4,102	18%	1,959	14%	3,484	19%	37,242	18%
36–45	4,366	19%	2,208	15%	3,627	20%	37,626	18%
46–55	3,787	17%	2,181	15%	3,179	17%	34,499	17%
56–65	3,403	15%	2,096	15%	2,648	15%	30,930	15%
66–79	3,727	17%	3,978	28%	2,599	14%	31,887	16%
Education								
ISCED 0–2	7,796	36%	4,277	34%	7,880	43%	125,976	61%
ISCED 3–4	8,127	37%	6,037	48%	6,538	36%	64,068	31%
ISCED 5–6	5,864	27%	2,179	18%	3,877	21%	14,919	8%
missing	655	3%	1,803	13%	58	0%	0	0%
Self-assessed health								
less-than-good	5,311	24%	5,668	40%	4,155	23%	-	-
good or above	17,115	76%	8,522	60%	14,197	77%	-	-
missing	16	0%	106	0%	1	0%	-	-
Long-standing health problems								
yes	9,338	42%	-	-	6,298	34%	-	-
no	13,094	58%	-	-	12,050	66%	-	-
missing	10	0%	-	-	5	0%	-	-
Smoking status								
current smoker	5,812	26%	3,409	25%	5,154	33%	52,622	26%
ex or never smoker	16,527	74%	10,380	75%	10,571	67%	151,675	74%
missing	103	0%	507	3%	2,628	14%	666	0%
Obesity								
yes	4,087	21%	1,778	13%	-	-	17,266	9%
no	15,568	79%	12,238	87%	-	-	181,925	91%
missing	2,787	12%	280	2%	-	-	5,772	3%

Note: The population distribution for each variable is given as % of subjects, excluding those with missing information. The % missing for each variable is given as a % of total subjects.

favourable in the 2000s (i.e., less increase or more decrease). All three-way interactions were statistically non-significant in England, implying that trends in health inequalities after 2000 did not significantly differ from those observed in the 1990s. As shown in Figure 1, this is because high-educated people also experienced favourable changes in trends after 2000. For the Netherlands, all three-way interactions were also statistically non-significant. Finland showed a significantly less favourable trend in inequalities in obesity after 2000 (OR=1.90, p<0.1). A more favourable change in trends of inequalities in obesity after 2000 was found in Italy (OR=0.76, p<0.1).

Table 3 "Two-way interaction" parameter estimates comparing the trends in health between 1990s and 2000s, "three-way interaction" parameter estimates comparing the trends in health inequalities between 1990s and 2000s, and "four-way interaction" parameter estimates comparing the "three-way interaction" parameter estimates between countries

	Odds ratios (logistic)			
	less-than-good self-assessed health	long-standing health problems	smoker	obesity
1. two-way interaction parameter estimates[a]				
England (low-edu)	0.76***	0.78***	0.82**	0.97
	(0.064)	(0.065)	(0.073)	(0.097)
2. three-way interaction parameter estimates[b]				
England	1.22	0.95	1.19	1.25
	(0.197)	(0.125)	(0.182)	(0.213)
Finland	0.78	-	1.28	1.90*
	(0.173)	-	(0.308)	(0.652)
the Netherlands	1.18	1.16	1.00	-
	(0.221)	(0.181)	(0.165)	-
Italy	-	-	0.97	0.76*
	-	-	(0.072)	(0.121)
3. four-way interaction parameter estimates[c]				
England vs Finland	1.57	-	0.93	0.66
	(0.433)	-	(0.267)	(0.253)
England vs the Netherlands	1.04	0.82	1.20	-
	(0.257)	(0.167)	(0.270)	-
England vs Italy	-	-	1.23	1.64**
	-	-	(0.209)	(0.383)

a. Based on the "two-way interaction" analysis for low-educated people in England. An odds ratio below 1.00 indicates a larger health improvement in the period 2000–2010 than in the period 1990–2000.
b. Based on the "three-way interaction" analysis within each country. An odds ratio below 1.00 indicates a more favourable trend in health inequalities in the period 2000–2010 than in the period 1990–2000.
c. Based on the "four-way interaction" analysis for England and each of the comparison countries. An odds ratio below 1.0 indicates a more favourable change (between 1990–2000 and 2000–2010) in the trend in health inequalities in England as compared to the other country.
Robust standard errors in parentheses. *** p<0.01, ** p<0.05, * p<0.1

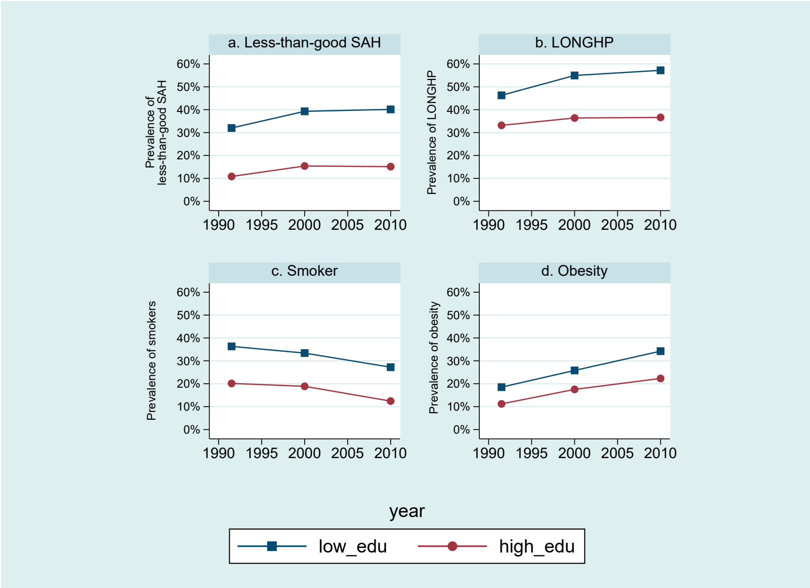


Figure 1 Trends in self-reported health outcomes in England by education

The results of the "four-way interaction" parameter estimates are reported in the last part of Table 3, which tested the differences in the "three-way interaction" parameter estimates between England and the comparison countries. An odds ratio below 1.00 indicates that the change in trends of inequalities was more favourable in England than in the comparison countries (i.e., a stronger change towards decreasing inequalities, or a weaker change towards increasing inequalities). Results showed that changes in trends of inequalities after 2000 were not statistically significantly different between England and any of the other countries, with the single exception of obesity for which the change was less favourable in England than in Italy (OR=1.64, p<0.05).

Using the amount of smoking per day among the current smokers as the health outcome did not change our conclusions (reported in the appendix, together with the full model results of the other outcomes). Essentially similar results were obtained in a sensitivity analysis (appendix, table A3), where weighting factors were incorporated when available. Trends in the health outcomes in the comparison countries during the study period are also reported in the appendix (Figure A1-A3). Additionally adding age square and the interaction between age and sex in the regressions to account for a potential nonlinear or sex-specific effect of age on health did not essentially change our results (available upon request).

DISCUSSION

Summary of findings

After the implementation of the English strategy, more favourable trends in some health indicators were observed among low-educated people, but trends in health inequalities in 2000–2010 in England were not more favourable than those observed in the period 1990–2000. For most health indicators, changes in trends of health inequalities after 2000 in England were also not statistically significantly different from those seen in the other countries.

Strengths and limitations

To the best of our knowledge, this is the first attempt to evaluate the population-level effects of the English strategy comparing trends in health inequalities in England before and after the implementation of the strategy, and between England and countries without a national programme to reduce health inequalities. We assessed whether there were larger improvements of health or health inequalities in England after the introduction of the strategy as compared to the pre-treatment trends, and simultaneously incorporated each of the comparison countries into the analysis to assess whether the changes in trends in health inequalities observed in England between the period 2000–2010 and the period 1990–2000 were more favourable than those in other countries without such a programme. For these analyses we developed a modified difference-in-difference analysis, based on a four-way interaction framework (education*time*policy*country) that may also be useful for the evaluation of other programs and interventions to tackle health inequalities. However, our study also has a number of potential weaknesses that need to be taken into account in interpreting our findings.

Our positive findings of changes in trends for the low educated in England will be biased if in the absence of the strategy, trends in health among the low educated would also have been more favourable in the second than in the first period. Potential reasons could include changes in major background causes of ill health not targeted by the strategy, resulting in, for example, less unemployment or higher incomes among lower educated. These changes should then be unique to England, since more favourable trends in health among the low educated in the period 2000–2010 were not generally observed in the comparison countries (appendix, table A4). We consider this unlikely, and therefore think that the more favourable changes in trends among the low educated in England can be interpreted as possible effects of the English strategy on health outcomes in this group.

Analogously, our finding of an absence of significant differences in trends in inequalities before and after the implementation of the strategy in England will be biased if in the absence of the strategy, trends in health inequalities would have been less favourable in the second than in the first period. One possible candidate for a background factor which may have increased health inequalities is the Great Recession that started in the late 2000s, which may have been especially harmful for the health of vulnerable populations and may have increased health inequalities [34–36]. However, there are two reasons why we believe the resulting bias can only be very limited. First, our last measure is for 2010, i.e. shortly after the recession started, and several of our health measures (e.g. long-standing health problems and obesity) are not likely to change within 1–2 years. Second, and more importantly, we compared the changes of trends in health inequalities in England to those in 3 comparison countries that also went through the recession and we found the changes in trends in England were not significantly more favourable. One remaining concern is that if the effect of the recession on health inequalities was different for each country, our results may still be biased. However, in order to mask the effect of the English strategy, England should be the country in which health inequalities were most affected by the recession. This is unlikely, given that the UK had similar or even smaller percentage decreases of GDP and employment as compared to Italy and Finland during the period of recession [37], and the UK did not show an increase in inequalities in self-reported health and some other health measures caused by the recession [36, 38]. We therefore think that the absence of more favourable changes in health inequalities in England can be interpreted as evidence for the absence of an effect of the English strategy on health inequalities. Changes in social mobility could also affect the degree of inequalities over time, but it is unclear whether a rise in social mobility would lead to wider or narrower health inequalities [39].

The validity of the comparison with the three other countries also hinges on whether trends in inequalities in health before the implementation of the English strategy were the same in England and the other countries, and whether other countries have indeed done much less than England to reduce health inequalities. Kunst et al., 2005 investigated trends in socioeconomic inequalities in self-assessed health in 10 European countries between the 1980's and 1990's. Their analyses showed a high degree of stability of socioeconomic inequalities in self-assessed health across the 10 countries, which included England, Finland, Italy and the Netherlands. Trends in educational inequalities in self-assessed health between 1980 and 1990 appeared to be similar in all four comparison countries, with the exception of males in Italy. In our own analysis, the changes in the odds ratios of educational inequalities in self-assessed health between 1980 and 1990 also appeared to be similar in all four comparison countries, with the exception of males in Italy. A possible improvement on our approach is the creation of a weighted "synthetic" control group, including several comparison countries [40]. In this innovative approach, weights are calculated such that the resulting synthetic control group best reproduces the values of a set of predictors of health inequalities in England before the implementation of the English strategy. We recommend to explore the usefulness of this approach for cross-national policy evaluations in future studies.

As stated in the introduction, there can be little doubt that England has done more to reduce health inequalities, but efforts to reduce health inequalities were not completely absent in the other countries. For example, by building a systematic evidence-base for interventions and policies, the Dutch government pursued a research-based approach to tackle socioeconomic inequalities in health [20]. Some of the recommendations were adopted by health policy-makers and health care practice, although more so at the local than at the national level [21]. Nevertheless, the absence of a reduction in health inequalities in the comparison countries suggests that it is unlikely that we have missed major policy effects.

Policies not specifically implemented with the purpose to reduce inequalities potentially could have influenced our findings. For example, the Dutch government increased the health care expenditure after 2001 [41]. Italy introduced a much more comprehensive smoke-free legislation in 2005, which made public transport completely smoke-free and extended coverage to bars and restaurants [42]. These policies could affect our results if their impacts were larger among lower as compared to higher socioeconomic groups. However, the extent to which these policies effectively reduce health inequalities is still rather unknown [43, 44]. Moreover, such policy changes are often specific to one country or one health outcome. Given that we have used three comparison countries and different health outcomes, and observe no significant decline of health inequalities in the comparison countries, it seems unlikely that such policies have influenced our findings substantially.

The validity of our analysis would also be compromised if the composition of the population would have changed differently in England as compared to the comparison countries. This "common composition" assumption would be violated if the UK had larger inward migration of persons with poorer or better health than the comparison countries. However, there is no evidence that changes in population composition are substantially different between these countries, as shown by statistics on the distribution of foreigners in European countries in 1980 and in 2000 (Wanner, 2002). Moreover, participation of migrants in surveys is usually low, so the potential for bias in our findings by immigration is also low.

Our analysis is limited by the fact that we have only used data on self-reported health measures obtained from the survey data (self-assessed health, long-standing health problems, smoking and obesity), and not on life expectancy and infant mortality, the overall targets of the English strategy. The main reason for this is the lack of comparable data for a sufficient number of countries and correct time-periods. There is evidence suggesting that inequalities in infant mortality between manual and non-manual occupational groups started to decrease in England after 2007 [45] – a possible effect of the English strategy which we have missed in our study. The only mortality outcome for which we could repeat our analyses is all-cause mortality in England and Finland (see web appendix Table A5). The analysis shows that the trend in mortality among the low educated in England was more favourable after the year 2000 than before, and that there was a smaller increase of inequalities in mortality in England after the year 2000 than before. However, the trend in inequalities in England was not statistically significantly different from that observed in Finland, which is consistent with our findings based on the survey data.

The two general health measures used in our analysis have been widely used in other comparative studies [46, 47] and have been shown to be reliable and valid indicators of general health and well-being [48]. They are also more likely to be changed in a short time span than mortality. The latter is also true for smoking, but less so for obesity. As explained in the introduction, favourable trends in inequalities in the chosen measures could reasonably be expected as a result of the English strategy, either because it directly targeted these outcomes (as in the case of smoking and obesity) or because it had more generalized effects beyond mortality (as in the case of self-assessed health and long-standing health problems). However, there is a potential concern that the implementation of the English strategy may change the willingness to report health problems of the respondents. It is therefore important to repeat our analysis with mortality and other more objective outcomes directly relevant for the strategy, if adequate data can be found.

We have used education as an indicator of socioeconomic position, which is one of the common socioeconomic indicators used in measuring health inequalities in European countries [22, 23, 47, 49]. Furthermore, education is strongly (albeit not perfectly) associated with both occupational class and income, and trends in health inequalities by education are often similar to those by occupational class or income (Hu et al, 2016). However, the objectives of the English strategy were phrased in terms of occupational class or area-based deprivation. Comparable measures of occupational class or area-based deprivation were not available in our data. We believe that it is reasonable to assume that policies that have effectively reduced inequalities in health by occupational class or area-based deprivation, will also have reduced inequalities in health by education. To the extent however, that the effect on the first was larger than that on the second, our use of education as socioeconomic indicator may have led to underestimation of the effect of the English strategy. We therefore recommend replication of our findings in a cross-national framework with other socioeconomic indicators, if these can be found.

We adjusted for gender in all analyses, in order to increase statistical power. Because trends in health outcomes may differ between men and women, especially for smoking [50], we repeated the analyses stratified by gender (see web appendix Tables A6 and A7). Among both men and women, for most of the health indicators changes in trends of health inequalities after 2000 in England were not statistically significantly different from those seen in the other countries. Out of 22 four-way interactions, only two showed more favourable trends in inequalities in England as compared to one of the other countries: long-standing health problems among men as compared to the Netherlands, and amount of smoking among women as compared to Italy.

In all countries, the response rate of the survey went down overtime, except in the Netherlands where it went up (Appendix Table A1). If the non-response population mainly consisted of people with low socioeconomic status or bad health, a decreasing response rate may potentially lead to more favourable trends in health inequalities. This might have biased the comparison between England and the Netherlands (but not the other two comparison countries). In our analysis, data from 3 years with 10-year gaps were used for each country. Although the choice of year is meaningful (the period of 2000–2010 is the period when the main effects of the strategy could be expected and the period of 1990–2000 is a comparable period when the strategy had not yet been generally implemented), the limited number of years implies that our measure of change may be unreliable. Further study may consider to repeat our analysis by using time-series data if available, which can help to model the trends better and improve the robustness. We mainly focused on the odds ratios of the core parameters and their statistical significance in the models, since these parameters could directly answer our study questions. In order to interpret the results better, future research may consider to present the predicted probabilities and use them to calculate meaningful results, such as the potential health or health inequalities in England in 2010 if it followed the trends of health or health inequalities in Finland. This was not done in our analysis since most of the "four-way interaction" parameters were insignificant.

Interpretation

Taking into account the trends in inequalities before the implementation of the strategy in England and the trends in the other three European countries, we found that the effects of the English strategy on inequalities in self-assessed health, long-standing health problems, smoking and obesity were limited. Our study confirms previous evaluations which have also not found clear effects of the English strategy on the population level [3, 6, 7]. Although evaluation studies have sometimes found positive effects in specific sections of the population, e.g. small but significant reductions in the absolute and relative rate gaps in smoking prevalence between Spearhead areas and others [51], and some beneficial effects of the Sure Start Local Programmes on children and their families living in deprived communities [52], the general consensus is that population-level effects have been largely absent. A possible exception might be tackling inequalities in infant and maternal health outcomes, where a national support team was established and some positive results were reported [53].

The potential reasons for why the English strategy was not more successful have been discussed in some reviews [3, 4, 6, 10, 11, 54–56]. One widely acknowledged reason is that the design of the English strategy was not based on policies or interventions with proven effectiveness in reducing health inequalities [1, 10, 11, 57, 58]. This is partly because of the reality that the current evidence for the effectiveness of policies is limited. Another reason is that the English strategy might have chosen the wrong entry-points. Mackenbach [11] pointed out that the strategy spent resources on entry-points which were irrelevant for life expectancy or infant mortality, at least within the chosen time frame. Marmot [3] noticed that the strategy had not systematically addressed the background causes of ill health and had relied more on tackling proximal causes (such as smoking). The inadequate delivery of the English strategy was also criticized [3, 10, 11].

The detected trends in inequalities in the three comparison countries are also not generally consistent with the efforts that have been made in each country in reducing health inequalities. Finland, which has a long tradition in eliminating inequalities [18] showed less favourable trends (although many not statistically significant) in inequalities in several outcomes in recent years as compared to the 1990s. Italy, which has made less efforts to tackle health inequalities, displayed decreasing trends in inequalities in all available measures, although significant decreases could only be shown in inequalities in obesity. Similar findings are reported in the literature on trends in health inequalities in Europe, and has been attributed to the fact that Italy is relatively late in many modern epidemic transitions [15, 16]. Apparently, more effective policies together with a deeper exploration of the causes for changes in health inequalities are needed.

Conclusions

In this rigorous analysis comparing trends in health inequalities in England both over time and between countries, we could not detect a favourable effect of the English strategy on national trends in educational inequalities in self-assessed health, long-standing health problems, smoking or obesity. However, our analysis illustrates the usefulness of a modified difference-in-difference approach for assessing the impact of policies on population-level health inequalities.

REFERENCES

1. Horton, R., *What the UK government is (not) doing about health inequalities*. Lancet, 2002. **360**(9328): p. 186.
2. Department of Health, *Reducing health inequalities: An action report*. 1999: London. http://www.swslim.org.uk/downloads/SL1264.pdf (accessed 02 Feb 2015).
3. Department of Health, *Fair society, healthy lives (the Marmot Review)*. 2010: London. http://www.instituteofhealthequity.org/projects/fair-society-healthy-lives-the-marmot-review (accessed 02 Feb 2015).
4. House of Commons, *Health inequalities*. 2009: London: Health Committee of the House of Commons. http://www.publications.parliament.uk/pa/cm200809/cmselect/cmhealth/286/28602.htm (accessed 02 Feb 2015).
5. Department of Health, *Independent inquiry into inequalities in health (the Acheson report)*. 1998: London. https://www.gov.uk/government/publications/independent-inquiry-into-inequalities-in-health-report (accessed 16 Oct 2014).
6. Department of Health, *Tackling health inequalities: 10 years on*. 2009: London. http://webarchive.nationalarchives.gov.uk/20130107105354/http:/www.dh.gov.uk/prod_consum_dh/groups/dh_digitalassets/documents/digitalasset/dh_098934.pdf (accessed 02 Feb 2015).
7. Department of Health, *Tackling health inequalities: 2007 status report of the program for action*. 2007: London. http://webarchive.nationalarchives.gov.uk/+/www.dh.gov.uk/en/Publicationsandstatistics/Publications/DH_083471 (accessed 02 Feb 2015).
8. Department of Health, *Tackling health inequalities: A program for action*. 2003: London. http://www.dh.gov.uk/prod_consum_dh/groups/dh_digitalassets/@dh/@en/documents/digitalasset/dh_4019362.pdf (accessed 02 Feb 2015).
9. Department of Health, *Tackling health inequalities: Cross-cutting review*. 2002: London. http://www.dh.gov.uk/prod_consum_dh/groups/dh_digitalassets/@dh/@en/documents/digitalasset/dh_4068003.pdf (accessed 02 Feb 2015).
10. Mackenbach, J.P., *Has the English strategy to reduce health inequalities failed?* Soc Sci Med, 2010. **71**(7): p. 1249–53; discussion 1254–8.
11. Mackenbach, J.P., *Can we reduce health inequalities? An analysis of the English strategy (1997–2010).* Journal of Epidemiology and Community Health, 2011. **65**(7): p. 568–575.
12. Mackenbach, J.P., et al., *Tackling socioeconomic inequalities in health: analysis of European experiences.* Lancet, 2003. **362**(9393): p. 1409–14.
13. Mackenbach, J.P., *Socio-economic inequalities in health in western Europe: from description to explanation to intervention. In Siegrist, J. & Marmot, M. (Eds.), Social inequalities in health: New evidence and policy implications.* 2006, New York: Oxford University Press.
14. Mackenbach, J.P., *The English strategy to reduce health inequalities.* Lancet, 2011. **377**(9782): p. 1986–1988.
15. Mackenbach, J.P., et al., *Widening socioeconomic inequalities in mortality in six Western European countries.* Int J Epidemiol, 2003. **32**(5): p. 830–7.
16. Mackenbach, J.P., et al., *Trends in inequalities in premature mortality: a study of 3.2 million deaths in 13 European countries.* J Epidemiol Community Health, 2014.
17. Whitehead, M., *Diffusion of ideas on social inequalities in health: a European perspective.* Milbank Q, 1998. **76**(3): p. 469–92, 306.
18. Judge, K., et al., *Health inequalities: a challenge for Europe*. 2006: An independent, expert report commissioned by, and published under the auspices of, UK Presidency of the EU. http://ec.europa.eu/

health/ph_determinants/socio_economics/documents/ev_060302_rd05_en.pdf (accessed 02 Feb 2015).

19. Ministry of Social Affairs and Health, *National Action Plan to Reduce Health Inequalities 2008–2011*. Publications of the Ministry of Social Affairs and Health Vol. 25. 2008, Helsinki.
20. Mackenbach, J.P. and K. Stronks, *A strategy for tackling health inequalities in the Netherlands.* BMJ, 2002. **325**(7371): p. 1029–32.
21. Mackenbach, J.P. and K. Stronks, *The development of a strategy for tackling health inequalities in the Netherlands.* Int J Equity Health, 2004. **3**(1): p. 11.
22. Kunst, A.E., et al., *Trends in socioeconomic inequalities in self-assessed health in 10 European countries.* Int J Epidemiol, 2005. **34**(2): p. 295–305.
23. Eikemo, T.A., et al., *Health inequalities according to educational level in different welfare regimes: a comparison of 23 European countries.* Sociol Health Illn, 2008. **30**(4): p. 565–82.
24. European Foundation for the Improvement of Living and Working Conditions, *Trends in quality of work in the Netherlands*. 2007: http://eurofound.europa.eu/sites/default/files/ef_files/ewco/surveys/NL0601SR01/NL0601SR01.pdf (accessed 11 December 2014).
25. Italian National Institute of Statistics (ISTAT), *Multipurpose Family survey: aspects of daily life*. Italian National Institute of Statistics, http://www.istat.it/it/archivio/4630 (accessed 16 July 2015).
26. Italian National Institute of Statistics (ISTAT), *Condizioni di salute e ricorso ai servizi sanitari [Health and Health Care Utilization]*. Italian National Institute of Statistics, http://www.istat.it/it/archivio/5471 (accessed 12 December 2014).
27. Katikireddi, S.V., C.L. Niedzwiedz, and F. Popham, *Trends in population mental health before and after the 2008 recession: a repeat cross-sectional analysis of the 1991–2010 Health Surveys of England.* BMJ Open, 2012. **2**(5).
28. National Institute for Health and Welfare. *Population studies: Survey on Health Behaviour and Health among the Finiish Adult Population (AVTK), Health Behaviour and Health among the Finnish Retirement-Age Populatin (EVTK)*.
29. Italian National Institute of Statistics (ISTAT), *Multipurpose Family Survey*. Italian National Institute of Statistics, http://www.istat.it/it/archivio/4394 (accessed 16 July 2015).
30. Carreras, G. and G. Gorini, *Time Trends of Italian Former Smokers 1980–2009 and 2010–2030 Projections Using a Bayesian Age Period Cohort Model.* International Journal of Environmental Research and Public Health, 2014. **11**(1): p. 1–12.
31. Angrist, J.D. and J.S. Pischke, *Mostly harmless econometrics*. 2009: Princeton University Press.
32. Wooldridge, J.M., *Introductory Econometrics: A Modern Approach*. 2012: Cengage Learning; 5 edition.
33. Lechner, M., *The Estimation of Causal Effects by Difference-in-Difference Methods.* Foundations and Trends in Econometrics, 2011. **4**(3): p. 165–224.
34. De Vogli, R., *The financial crisis, health and health inequities in Europe: the need for regulations, redistribution and social protection.* Int J Equity Health, 2014. **13**: p. 58.
35. Suhrcke, M. and D. Stuckler, *Will the recession be bad for our health? It depends.* Soc Sci Med, 2012. **74**(5): p. 647–53.
36. Bacigalupe, A. and A. Escolar-Pujolar, *The impact of economic crises on social inequalities in health: what do we know so far?* Int J Equity Health, 2014. **13**.
37. Bell, D.N.F. and D.G. Blanchflower, *Recession and Unemployment in the OECD*. 2010: CESifo Forum 1/2010, https://www.cesifo-group.de/portal/page/portal/968433588CCE0D9FE04400144FAFBA7C (accessed 22 July 2015).
38. Astell-Burt, T. and X. Feng, *Health and the 2008 economic recession: evidence from the United Kingdom.* PLoS One, 2013. **8**(2): p. e56674.

39. Bartley, M. and I. Plewis, *Does health-selective mobility account for socioeconomic differences in health? Evidence from England and Wales, 1971 to 1991*. Journal of Health and Social Behavior, 1997. **38**(4): p. 376–386.
40. Abadie, A., A. Diamond, and J. Hainmueller, *Synthetic control methods for comparative case studies: estimating the effect of California's Tobacco Control Program*. Journal of the American Statistical Association, 2010. **105**(490): p. 493–505.
41. Peters, F., et al., *Quantifying the contribution of changes in healthcare expenditures and smoking to the reversal of the trend in life expectancy in the Netherlands*. BMC Public Health, 2015. **15**: p. 1024.
42. Currie, L., *Appendix A. Tobacco Control Policy Index. In: Nguyen L., Rosenqvist G. & Pekurinen M. Demand for Tobacco in Europe: An Econometric Analysis of 11 Countries for the PPACTE Project*. . 2012, Tampere, Finland: Juvenes Print.
43. Brown, T., S. Platt, and A. Amos, *Equity impact of population-level interventions and policies to reduce smoking in adults: a systematic review*. Drug Alcohol Depend, 2014. **138**: p. 7–16.
44. Mackenbach, J.P., *Socioeconomic inequalities in health in high-income countries: the facts and the options*. In: Detels, R. Gulliford, M., Karim, Q.A., Tan, C.C., eds. Oxford Textbook of Global Public Health. 2015, Oxford: Oxford University Press.
45. Bambra, C., *Reducing health inequalities: new data suggest that the English strategy was partially successful*. Journal of Epidemiology and Community Health, 2012. **66**(7): p. 662–662.
46. Mackenbach, J.P., et al., *Socioeconomic inequalities in morbidity and mortality in western Europe. The EU Working Group on Socioeconomic Inequalities in Health*. Lancet, 1997. **349**(9066): p. 1655–9.
47. Cavelaars, A.E., et al., *Differences in self reported morbidity by educational level: a comparison of 11 western European countries*. J Epidemiol Community Health, 1998. **52**(4): p. 219–27.
48. Manderbacka, K., *Examining what self-rated health question is understood to mean by respondents*. Scandinavian Journal of Social Medicine, 1998. **26**(2): p. 145–153.
49. Mackenbach, J.P., et al., *Socioeconomic inequalities in health in 22 European countries*. N Engl J Med, 2008. **358**(23): p. 2468–81.
50. Giskes, K., et al., *Trends in smoking behaviour between 1985 and 2000 in nine European countries by education*. J Epidemiol Community Health, 2005. **59**(5): p. 395–401.
51. Bauld, L., K. Judge, and S. Platt, *Assessing the impact of smoking cessation services on reducing health inequalities in England: observational study*. Tob Control, 2007. **16**(6): p. 400–4.
52. Melhuish, E., et al., *Effects of fully-established Sure Start Local Programmes on 3-year-old children and their families living in England: a quasi-experimental observational study*. Lancet, 2008. **372**(9650): p. 1641–7.
53. Department of Health, *Tackling health inequalities in infant and maternal health outcomes (report of the infant mortality national support team)*. 2010: London. https://www.gov.uk/government/uploads/system/uploads/attachment_data/file/215869/dh_122844.pdf (accessed 16 Apr 2015).
54. Judge, K. and L. Bauld, *Learning from policy failure? Health action zones in England*. European Journal of Public Health, 2006. **16**(4): p. 341–343.
55. Exworthy, M., L. Berney, and M. Powell, *'How great expectations in Westminster may be dashed locally': the local implementation of national policy on health inequalities*. Policy and Politics, 2002. **30**(1): p. 79–96.
56. Hunter, D.J., et al., *Getting to grips with health inequalities at last?* British Medical Journal, 2010. **340**.
57. Macintyre, S., et al., *Using evidence to inform health policy: case studs*. British Medical Journal, 2001. **322**(7280): p. 222–225.
58. Bambra, C., et al., *A labour of Sisyphus? Public policy and health inequalities research from the Black and Acheson Reports to the Marmot Review*. J Epidemiol Community Health, 2011. **65**(5): p. 399–406.

WEB APPENDIX

Details on the models, coding of the variables and the full results

1.The two-way interaction model:

$$outcome_{ist} = \beta_0 + \beta_1\ endyear_t + \beta_2 policyperiod_s + \beta_3\ endyear_t * policyperiod_s + age_{ist} + sex_{ist}$$

Independent Variables	Definitions/Coding
endyear	For the period 1990–2000: *endyear*=1 if year=2000 *endyear*=0 if year=1990; For the period 2000–2010: *endyear*=1 if year=2010 *endyear*=0 if year=2000[a]
policyperiod	*policyperiod*=1 if the period is 2000–2010; *policyperiod*=0 if the period is 1990–2000[a]
*endyear*policyperiod*	the interaction between *endyear* and *policyperiod*
age	continuous variable
sex	*sex*=1 for males; *sex*=2 for females

a. Please note that since the period 1990–2000 was used as an independent control condition for the treatment condition 2000–2010, data from year 2000 were used twice (both as the starting year of the treatment period and as the ending year of the control period). In this way, we could give the data from year 2000 different values for both variables "*endyear*" and "*policyperiod*", depending on whether the year 2000 was regarded as the ending year of the control period or as the starting year of the treatment period.
As for the specific year, year 2000 was always considered as the year when the English strategy was not generally implemented (no matter whether it was considered as the ending year of 1990–2000 or the starting year of 2000–2010). Being the ending year of 2000–2010 (the treatment period), year 2010 was the only year that was considered as the year when the English strategy potentially had effects.

The full results of the two-way interaction models conducted among low-educated people in England are:

	Odds ratios (logistic)				Coefficients (linear)
Variables	Less-than-good SAH	LONGHP	smoker	obesity	amount of smoking
policyperiod	1.307***	1.325***	0.949	1.501***	-0.453
	(0.0736)	(0.0738)	(0.0551)	(0.103)	(0.411)
endyear	1.307***	1.325***	0.949	1.501***	-0.453
	(0.0736)	(0.0738)	(0.0551)	(0.103)	(0.411)
policyperiod*endyear	0.762***	0.780***	0.822**	0.970	-0.643
	(0.0639)	(0.0654)	(0.0732)	(0.0967)	(0.631)
age	1.024***	1.037***	0.968***	1.013***	-0.0387***
	(0.00131)	(0.00131)	(0.00123)	(0.00143)	(0.00906)
sex	0.975	0.929*	0.884***	1.363***	-2.712***
	(0.0408)	(0.0387)	(0.0387)	(0.0690)	(0.318)
Constant	0.141***	0.146***	3.546***	0.0720***	22.26***
	(0.0143)	(0.0145)	(0.361)	(0.00865)	(0.694)
Observations	10,336	10,341	10,316	8,875	3,377

Robust standard errors in parentheses, *** p<0.01, ** p<0.05, * p<0.

We did not use the clustered standard errors in our analysis since the clustered standard errors can be biased if the number of clusters is small (Angrist & Pischke, 2009), and the cure can be worse than the disease (Nichols & Schaffer, 2007). There is no agreement on the required number of clusters for reliable inference. Angrist and Pischke semi-jokingly gave the number of 42 as the minimum number of clusters. In our study, there are only 2 countries at most in each regression, which is far from enough for the clustered standard errors to be reliable. Therefore, we only used the robust standard errors to account for potential heteroscedasticity.

However, given that clustered standard errors are generally larger than normal or robust standard errors, we believe that using clustered standard errors will not change our general conclusions since most of the four-way interactions were already insignificant.

References

Angrist, J. D., & Pischke, J. S. (2009). Mostly harmless econometrics: Princeton University Press.

Nichols, A., & Schaffer, M. (2007). Clustered errors in Stata. United Kingdom Stata Users' Group Meetings, http://repec.org/usug2007/crse.pdf (accessed 7th July 2015).

2. The three-way interaction model:

$$outcome\_ist = \beta_0 + \beta_1\, endyear_t + \beta_2\, policyperiod_s + \beta_3\, endyear_t * policyperiod_s + \beta_4\, ledu_{ist} + \beta_5\, ledu_{ist} * endyear_t + \beta_6\, ledu_{ist} * policyperiod_s + \beta_7\, ledu_{ist} * endyear_t * policyperiod_s + X_{ist}$$

The coding for *endyear, policyperiod, age* and *sex* is exactly the same as that in the two-way interaction model. The dummy variable *ledu* and its interactions were additionally added. Specifically, *ledu*=1 for low-educated people, and *ledu*=0 for high-educated people. Middle-educated group was excluded from this analysis.

The full results of the three-way interaction models conducted for England are:

	Odds ratios (logistic)				Coefficients (linear)
Variables	Less-than-good SAH	LONGHP	smoker	obesity	amount of smoking
policyperiod	1.464***	1.112	0.955	1.657***	0.206
	(0.162)	(0.0871)	(0.0890)	(0.187)	(0.765)
endyear	1.464***	1.112	0.955	1.657***	0.206
	(0.162)	(0.0871)	(0.0890)	(0.187)	(0.765)
ledu	3.061***	1.209***	3.239***	1.535***	4.821***
	(0.308)	(0.0887)	(0.273)	(0.164)	(0.654)
policyperiod*endyear	0.624***	0.818**	0.694***	0.780*	-1.953*
	(0.0862)	(0.0822)	(0.0865)	(0.108)	(1.009)
policyperiod*ledu	0.891	1.191*	0.988	0.899	-0.692
	(0.111)	(0.115)	(0.108)	(0.119)	(0.869)
endyear*ledu	0.891	1.191*	0.988	0.899	-0.692
	(0.111)	(0.115)	(0.108)	(0.119)	(0.869)
policyperiod*endyear*ledu	1.220	0.952	1.189	1.250	1.299
	(0.197)	(0.125)	(0.182)	(0.213)	(1.191)
age	1.025***	1.039***	0.969***	1.014***	-0.00674
	(0.00115)	(0.00108)	(0.00109)	(0.00119)	(0.00822)
sex	0.994	0.994	0.808***	1.175***	-2.739***
	(0.0350)	(0.0315)	(0.0287)	(0.0459)	(0.267)
Constant	0.0412***	0.0990***	1.185*	0.0556***	15.98***
	(0.00471)	(0.00884)	(0.117)	(0.00674)	(0.766)
Observations	18,245	18,250	18,218	15,878	4,711

Robust standard errors in parentheses, *** p<0.01, ** p<0.05, * p<0.1

The full results of the three-way interaction models conducted for Finland are:

	Odds ratios (logistic)			Coefficients (linear)
Variables	Less-than-good SAH	smoker	obesity	amount of smoking
policyperiod	0.978	0.935	1.882**	0.177
	(0.151)	(0.141)	(0.503)	(1.332)
endyear	0.978	0.935	1.882**	0.177
	(0.151)	(0.141)	(0.503)	(1.332)
ledu	2.588***	2.062***	2.554***	4.793***
	(0.352)	(0.273)	(0.635)	(1.188)
policyperiod*endyear	1.133	0.919	0.586*	-2.217
	(0.220)	(0.182)	(0.183)	(1.621)
policyperiod*ledu	0.868	0.793	0.588*	0.0518
	(0.149)	(0.139)	(0.168)	(1.519)
endyear*ledu	0.868	0.793	0.588*	0.0518
	(0.149)	(0.139)	(0.168)	(1.519)
policyperiod*endyear*ledu	0.777	1.275	1.896*	1.228
	(0.173)	(0.308)	(0.652)	(2.003)
age	1.038***	0.969***	1.021***	0.0640***
	(0.00169)	(0.00168)	(0.00208)	(0.0180)
sex	0.997	0.555***	1.306***	-3.934***
	(0.0460)	(0.0305)	(0.0807)	(0.434)
Constant	0.0595***	2.420***	0.0143***	14.03***
	(0.00931)	(0.378)	(0.00379)	(1.441)
Observations	8,864	8,548	8,762	1,525

Robust standard errors in parentheses, *** p<0.01, ** p<0.05, * p<0.1

The full results of the three-way interaction models conducted for the Netherlands are:

	Odds ratios (logistic)			Coefficients (linear)
Variables	Less-than-good SAH	LONGHP	smoker	amount of smoking
policyperiod	0.841	1.410***	0.640***	-3.583***
	(0.106)	(0.149)	(0.0672)	(1.012)
endyear	0.841	1.410***	0.640***	-3.583***
	(0.106)	(0.149)	(0.0672)	(1.012)
ledu	1.938***	1.483***	1.663***	1.465*
	(0.215)	(0.149)	(0.156)	(0.791)
policyperiod*endyear	0.836	0.716**	1.195	2.836**
	(0.139)	(0.0947)	(0.167)	(1.367)
policyperiod*ledu	1.174	0.973	1.164	0.567
	(0.164)	(0.118)	(0.144)	(1.157)
endyear*ledu	1.174	0.973	1.164	0.567
	(0.164)	(0.118)	(0.144)	(1.157)
policyperiod*endyear*ledu	1.177	1.163	0.995	0.816
	(0.221)	(0.181)	(0.165)	(1.589)
age	1.033***	1.031***	0.983***	0.00593
	(0.00115)	(0.00104)	(0.00111)	(0.0109)
sex	1.211***	1.235***	0.672***	-1.059***
	(0.0467)	(0.0423)	(0.0257)	(0.369)
Constant	0.0335***	0.0556***	1.968***	16.62***
	(0.00416)	(0.00617)	(0.212)	(0.981)
Observations	16,496	16,492	13,438	3,092

Robust standard errors in parentheses, *** p<0.01, ** p<0.05, * p<0.1

The full results of the three-way interaction models conducted for Italy are:

	Odds ratios (logistic)		Coefficients (linear)
Variables	smoker	obesity	amount of smoking
policyperiod	0.783***	1.405***	-1.025***
	(0.0409)	(0.184)	(0.369)
endyear	0.783***	1.405***	-1.025***
	(0.0409)	(0.184)	(0.369)
ledu	0.962	2.641***	1.873***
	(0.0438)	(0.315)	(0.328)
policyperiod*endyear	1.103	1.063	-1.025**
	(0.0779)	(0.165)	(0.484)
policyperiod*ledu	1.190***	0.955	-0.408
	(0.0648)	(0.127)	(0.394)
endyear*ledu	1.190***	0.955	-0.408
	(0.0648)	(0.127)	(0.394)
policyperiod*endyear*ledu	0.967	0.761*	0.772
	(0.0724)	(0.121)	(0.522)
age	0.980***	1.026***	0.00574**
	(0.000271)	(0.000388)	(0.00285)
sex	0.397***	0.888***	-4.437***
	(0.00416)	(0.0126)	(0.0830)
Constant	3.758***	0.0115***	20.73***
	(0.177)	(0.00139)	(0.346)
Observations	218,223	213,977	53,436

Robust standard errors in parentheses, *** p<0.01, ** p<0.05, * p<0.1

3. The four-way interaction model:

$$outcome\_{istj} = \textit{the three-way interaction model} * (1 + \textit{england dummy})$$
$$= (\beta_0 + \beta_1\, endyear_{tj} + \beta_2\, policyperiod_{sj} + \beta_3\, endyear_{tj} * policyperiod_{sj} + \beta_4\, ledu_{istj} + \beta_5\, ledu_{istj} * endyear_{tj} + \beta_6\, ledu_{istj} * policyperiod_{sj} + \beta_7\, ledu_{istj} * endyear_{tj} * policyperiod_{sj} + X_{istj})$$
$$+ (\beta_0' + \beta_1'\, endyear_{tj} + \beta_2'\, policyperiod_{sj} + \beta_3'\, endyear_{tj} * policyperiod_{sj} + \beta_4'\, ledu_{istj} + \beta_5'\, ledu_{istj} * endyear_{tj} + \beta_6'\, ledu_{istj} * policyperiod_{sj} + \beta_7'\, ledu_{istj} * endyear_{tj} * policyperiod_{sj} + X_{istj}) * england_j$$

The coding for *endyear*, *policyperiod*, *ledu*, *age* and *sex* is exactly the same as that in the two-way and three-way interaction models. But since we included data from both England and one of the comparison countries in the regression, we additionally added a dummy for England and its interaction terms into the regression. Specifically, *england*=1 for data from England, and *england*=0 for data from the comparison countries.

The full results of the four-way interaction models comparing the changes in trends in health inequalities between 2000–2010 and 1990–2000 in England and those in Finland are:

	Odds ratios (logistic)			Coefficients (linear)
Variables	Less-than-good SAH	smoker	obesity	amount of smoking
policyperiod	0.978	0.935	1.882**	0.177
	(0.151)	(0.141)	(0.503)	(1.330)
endyear	0.978	0.935	1.882**	0.177
	(0.151)	(0.141)	(0.503)	(1.330)
ledu	2.588***	2.062***	2.554***	4.793***
	(0.352)	(0.273)	(0.635)	(1.186)
policyperiod*endyear	1.133	0.919	0.586*	-2.217
	(0.220)	(0.182)	(0.183)	(1.618)
policyperiod*ledu	0.868	0.793	0.588*	0.0518
	(0.149)	(0.139)	(0.168)	(1.517)
endyear*ledu	0.868	0.793	0.588*	0.0518
	(0.149)	(0.139)	(0.168)	(1.517)
policyperiod*endyear*ledu	0.777	1.275	1.896*	1.228
	(0.173)	(0.308)	(0.652)	(1.999)
england	0.693*	0.490***	3.890***	1.948
	(0.134)	(0.0906)	(1.133)	(1.630)
england*policyperiod	1.498**	1.021	0.881	0.0294

(continued)

	Odds ratios (logistic)			Coefficients (linear)
Variables	Less-than-good SAH	smoker	obesity	amount of smoking
	(0.285)	(0.181)	(0.255)	(1.534)
england*endyear	1.498**	1.021	0.881	0.0294
	(0.285)	(0.181)	(0.255)	(1.534)
england*ledu	1.182	1.571***	0.601*	0.0282
	(0.200)	(0.246)	(0.163)	(1.355)
england*policyperiod*endyear	0.551**	0.755	1.330	0.263
	(0.131)	(0.177)	(0.455)	(1.908)
england*policyperiod*ledu	1.026	1.247	1.530	-0.744
	(0.217)	(0.258)	(0.482)	(1.748)
england*endyear*ledu	1.026	1.247	1.530	-0.744
	(0.217)	(0.258)	(0.482)	(1.748)
england*policyperiod*endyear*ledu	1.571	0.933	0.659	0.0712
	(0.433)	(0.267)	(0.253)	(2.328)
age	1.038***	0.969***	1.021***	0.0640***
	(0.00169)	(0.00168)	(0.00208)	(0.0180)
sex	0.997	0.555***	1.306***	-3.934***
	(0.0460)	(0.0305)	(0.0807)	(0.434)
england*age	0.988***	1.000	0.993***	-0.0707***
	(0.00195)	(0.00206)	(0.00233)	(0.0198)
england*sex	0.997	1.456***	0.899	1.195**
	(0.0579)	(0.0952)	(0.0657)	(0.509)
Constant	0.0595***	2.420***	0.0143***	14.03***
	(0.00931)	(0.378)	(0.00379)	(1.438)
Observations	27,109	26,766	24,640	6,236

Robust standard errors in parentheses, *** p<0.01, ** p<0.05, * p<0.1

The full results of the four-way interaction models comparing the changes in trends in health inequalities between 2000–2010 and 1990–2000 in England and those in the Netherlands are:

	Odds ratios (logistic)			Coefficients (linear)
Variables	Less-than-good SAH	LONGHP	smoker	amount of smoking
policyperiod	0.841	1.410***	0.640***	-3.583***
	(0.106)	(0.149)	(0.0672)	(1.012)
endyear	0.841	1.410***	0.640***	-3.583***
	(0.106)	(0.149)	(0.0672)	(1.012)
ledu	1.938***	1.483***	1.663***	1.465*
	(0.215)	(0.149)	(0.156)	(0.790)
policyperiod*endyear	0.836	0.716**	1.195	2.836**
	(0.139)	(0.0947)	(0.167)	(1.366)
policyperiod*ledu	1.174	0.973	1.164	0.567
	(0.164)	(0.118)	(0.144)	(1.156)
endyear*ledu	1.174	0.973	1.164	0.567
	(0.164)	(0.118)	(0.144)	(1.156)
policyperiod*endyear*ledu	1.177	1.163	0.995	0.816
	(0.221)	(0.181)	(0.165)	(1.589)
england	1.232	1.779***	0.602***	-0.639
	(0.208)	(0.253)	(0.0879)	(1.244)
england*policyperiod	1.740***	0.789*	1.492***	3.789***
	(0.291)	(0.104)	(0.209)	(1.268)
england*endyear	1.740***	0.789*	1.492***	3.789***
	(0.291)	(0.104)	(0.209)	(1.268)
england*ledu	1.580***	0.816	1.948***	3.356***
	(0.237)	(0.102)	(0.246)	(1.026)
england*policyperiod*endyear	0.746	1.142	0.581***	-4.789***
	(0.161)	(0.190)	(0.109)	(1.699)
england*policyperiod*ledu	0.758	1.224	0.849	-1.260
	(0.142)	(0.190)	(0.140)	(1.447)
england*endyear*ledu	0.758	1.224	0.849	-1.260
	(0.142)	(0.190)	(0.140)	(1.447)
england*policyperiod*endyear*ledu	1.037	0.819	1.195	0.483
	(0.257)	(0.167)	(0.270)	(1.986)
age	1.033***	1.031***	0.983***	0.00593
	(0.00115)	(0.00104)	(0.00111)	(0.0109)
sex	1.211***	1.235***	0.672***	-1.059***
	(0.0467)	(0.0423)	(0.0257)	(0.369)
england*age	0.992***	1.007***	0.986***	-0.0127
	(0.00157)	(0.00146)	(0.00157)	(0.0136)
england*sex	0.821***	0.805***	1.202***	-1.680***
	(0.0429)	(0.0375)	(0.0627)	(0.455)
Constant	0.0335***	0.0556***	1.968***	16.62***
	(0.00415)	(0.00617)	(0.212)	(0.981)
Observations	34,741	34,742	31,656	7,803

Robust standard errors in parentheses, *** p<0.01, ** p<0.05, * p<0.1

The full results of the four-way interaction models comparing the changes in trends in health inequalities between 2000–2010 and 1990–2000 in England and those in Italy are:

	Odds ratios (logistic)		Coefficients (linear)
Variables	smoker	obesity	amount of smoking
policyperiod	0.783***	1.405***	-1.025***
	(0.0409)	(0.184)	(0.369)
endyear	0.783***	1.405***	-1.025***
	(0.0409)	(0.184)	(0.369)
ledu	0.962	2.641***	1.873***
	(0.0438)	(0.315)	(0.328)
policyperiod*endyear	1.103	1.063	-1.025**
	(0.0779)	(0.165)	(0.484)
policyperiod*ledu	1.190***	0.955	-0.408
	(0.0648)	(0.127)	(0.394)
endyear*ledu	1.190***	0.955	-0.408
	(0.0648)	(0.127)	(0.394)
policyperiod*endyear*ledu	0.967	0.761*	0.772
	(0.0724)	(0.121)	(0.522)
england	0.315***	4.816***	-4.755***
	(0.0345)	(0.823)	(0.840)
england*policyperiod	1.220*	1.180	1.231
	(0.130)	(0.204)	(0.849)
england*endyear	1.220*	1.180	1.231
	(0.130)	(0.204)	(0.849)
england*ledu	3.367***	0.581***	2.948***
	(0.323)	(0.0930)	(0.731)
england*policyperiod*endyear	0.629***	0.734	-0.929
	(0.0902)	(0.153)	(1.118)
england*policyperiod*ledu	0.830	0.942	-0.285
	(0.102)	(0.177)	(0.953)
england*endyear*ledu	0.830	0.942	-0.285
	(0.102)	(0.177)	(0.953)
england*policyperiod*endyear*ledu	1.230	1.642**	0.527
	(0.209)	(0.383)	(1.299)
age	0.980***	1.026***	0.00574**
	(0.000271)	(0.000388)	(0.00285)
sex	0.397***	0.888***	-4.437***
	(0.00416)	(0.0126)	(0.0830)
england*age	0.990***	0.988***	-0.0125
	(0.00114)	(0.00121)	(0.00869)
england*sex	2.036***	1.323***	1.699***
	(0.0754)	(0.0550)	(0.280)
Constant	3.758***	0.0115***	20.73***
	(0.177)	(0.00139)	(0.346)
Observations	236,441	229,855	58,147

Robust standard errors in parentheses, *** p<0.01, ** p<0.05, * p<0.1

Chapter 6

Table A1 Information on data collection and health outcome measures

Country	Survey year	Sampling strategy	Response rate	Question and answer for SAH	Question and answer for LONGHP	Measurement of smoking status	Measurement of amount of smoking	The way of collecting height / weight
England	1991–1992	multi-stage stratified random sample	84%	How is your health in general? 1= fair/bad/very bad 0= very good/good	Do you have any longstanding illness, disability or infirmity? 1=Yes 0=No	1= current smoker 0=ex- or non-smoker	Number of cigarettes per day	Measured
	2000	multi-stage stratified random sample	75%	How is your health in general? 1= fair/bad/very bad 0= very good/good	Do you have any longstanding illness, disability or infirmity? 1=Yes 0=No	1= current smoker 0=ex- or non-smoker	Number of cigarettes per day	Measured
	2010	multi-stage stratified random sample	66%	How is your health in general? 1= fair/bad/very bad 0= very good/good	Do you have any longstanding illness, disability or infirmity? 1=Yes 0=No	1= current smoker 0=ex- or non-smoker	Number of cigarettes per day	Measured
Finland	1989	(stratified) random sample[a]	77%/87%[b]	Assessment of own health status: 1= average/rather poor/poor 0= good/reasonably good	-	1= current smoker 0=ex- or non-smoker	Number of cigarettes per day	Self-reported

Table A1 Information on data collection and health outcome measures (continued)

Country	Survey year	Sampling strategy	Response rate	Question and answer for SAH	Question and answer for LONGHP	Measurement of smoking status	Measurement of amount of smoking	The way of collecting height / weight
	1999	(stratified) random sample[a]	69%/75%[b]	Assessment of own health status: 1= average/rather poor/ poor 0= good/reasonably good	-	1= current smoker 0=ex- or non-smoker	Number of cigarettes per day	Self-reported
	2009	(stratified) random sample[a]	59%/73%[b]	Assessment of own health status: 1= average/rather poor/ poor 0= good/reasonably good	-	1= current smoker 0=ex- or non-smoker	Number of cigarettes per day	Self-reported
The Netherlands	1990	three-stage sample: municipalities, addresses, people	47%	In general, how do you experience your health condition? 1= fair/poor/very poor 0= very good/good	Do you suffer from any long-standing illness, infirmity, or disability? 1=Yes 0=No	1= current smoker 0=ex- or non-smoker	Number of cigarettes per day	-
	2000	two-stage sample: municipalities, people	59%	How in general is your health condition? 1= fair/sometimes good sometimes poor/very poor 0= very good/good	Do you suffer from any long-standing illness, infirmity, or disability? 1=Yes 0=No	1= current smoker 0=ex- or non-smoker	Number of cigarettes per day	-

Table A1 Information on data collection and health outcome measures (continued)

Country	Survey year	Sampling strategy	Response rate	Question and answer for SAH	Question and answer for LONGHP	Measurement of smoking status	Measurement of amount of smoking	The way of collecting height / weight
	2009	two-stage sample: municipalities, people	63%	How in general is your health condition? 1= fair/sometimes good sometimes poor/very poor 0= very good/good	Do you suffer from any long-standing illness, infirmity, or disability? 1=Yes 0=No	1= current smoker 0=ex- or non-smoker	Number of cigarettes per day	-
Italy	1990	two stage stratified sample	96%	-	-	1= current smoker 0=ex- or non-smoker	Number of cigarettes per day	Self-reported
	2000	two stage stratified sample	87%	-	-	1= current smoker 0=ex- or non-smoker	Number of cigarettes per day	Self-reported
	2010	two stage stratified sample	82%	-	-	1= current smoker 0=ex- or non-smoker	Number of cigarettes per day	Self-reported

Notes:
a. The Finnish data used in this study are the data combined from the two Finish studies: "Health behaviour and health among Finnish adult population (AVTK)", which includes respondents who are 15–64 years old, and "Health behaviour and health among the Finnish elderly (EVTK)", which includes respondents who are older than 64 years. AVTK runs on a random sample of Finnish people aged 15–64 derived from the Population Register each year. EVTK runs on a random sample stratified by age and sex of Finnish people aged 65 or older every two years.
b. We report the response rates for AVTK/EVTK separately.

Table A2 Classification of education in England

ISCED category	Education classification in the survey data	Percentage
Year 1991–1992		
ISCED 0–2 (low-educated)	GCSE d-g or equiv; ungraded CSE; no qualification	48%
ISCED 3–4 (middle-educated)	a levels or equiv; GCSE a-c or equiv	34%
ISCED 5–6 (high-educated)	degree; prof not degree	18%
Year 2000		
ISCED 0–2 (low-educated)	NVQ1/CSE other grade equiv; No qualification	35%
ISCED 3–4 (middle-educated)	NVQ3/GCE A Level equiv; NVQ2/GCE O Level equiv	37%
ISCED 5–6 (high-educated)	NVQ4/NVQ5/Degree or equiv; Higher ed below degree	28%
Year 2010		
ISCED 0–2 (low-educated)	NVQ1/CSE other grade equiv; No qualification	26%
ISCED 3–4 (middle-educated)	NVQ3/GCE A Level equiv; NVQ2/GCE O Level equiv	40%
ISCED 5–6 (high-educated)	NVQ4/NVQ5/Degree or equiv; Higher ed below degree	34%

Table A3 Two-way interaction estimators comparing the trends in health between 1990s and 2000s, three-way interaction estimators comparing the trends in health inequalities between 1990s and 2000s, and four-way interaction estimators, comparing the three-way interaction estimators between countries, weighted results[a]

	Odds ratios (logistic)				Coefficients (linear)
	Less-than-good SAH	LONGHP	smoker	obesity	amount of smoking
1. two-way interaction estimators[a]					
England (low-edu)	0.75***	0.75***	0.80**	0.95	-0.91
	(0.064)	(0.064)	(0.073)	(0.096)	(0.644)
2. three-way interaction estimators[b]					
England	1.18	0.93	0.99	1.22	1.38
	(0.195)	(0.124)	(0.108)	(0.210)	(1.199)
Finland	0.78	-	1.27	1.90*	1.23
	(0.173)	-	(0.308)	(0.652)	(2.003)
the Netherlands	1.17	1.02	1.00	-	1.45
	(0.238)	(0.177)	(0.179)	-	(1.785)
Italy	-	-	0.93	0.81	0.39
	-	-	(0.083)	(0.152)	(0.612)
3. four-way interaction estimator[c]					
England vs Finland	1.52	-	0.89	0.64	0.15
	(0.421)	-	(0.256)	(0.247)	(2.332)
England vs the Netherlands	1.01	0.90	1.14	-	-0.08
	(0.263)	(0.198)	(0.271)	-	(2.150)
England vs Italy	-	-	1.21	1.50	0.99
	-	-	(0.219)	(0.381)	(1.345)

a. Weights are available for England 2010, the Netherlands 1990/2000/2010 and Italy 1990/2000. Weights included in the Dutch survey 2000 and 2010 are expansion weights, which are scaled (i.e. divide each year by the mean weight) in order to be comparable to the other years and other countries.
b. Based on the two-way interaction analysis for low-educated people in England. An odds ratio below 1.00 or a negative linear coefficient indicates a larger health improvement in the period 2000–2010 than in the period 1990–2000.
c. Based on the three-way interaction analysis within each country. An odds ratio below 1.00 or a negative linear coefficient indicates a more favourable trend in health inequalities in the period 2000–2010 than in the period 1990–2000.
d. Based on the four-way interaction analysis for England and each of the comparison countries. An odds ratio below 1.00 or a negative linear coefficient indicates a more favourable change (between 1990–2000 and 2000–2010) in the trend in health inequalities in England as compared to the other country.
Robust standard errors in parentheses. *** p<0.01, ** p<0.05, * p<0.1

Table A4 Two-way interaction estimators comparing trends in health among low-educated people between 1990s and 2000s, within the three comparison countries

	Odds ratios (logistic)				Coefficients (linear)
	Less-than-good SAH	LONGHP	smoker	obesity	amount of smoking
two-way interaction estimators					
Finland	0.91	-	1.15	1.07	-1.09
	(0.103)	-	(0.162)	(0.155)	(1.188)
the Netherlands	0.99	0.84**	1.20**	-	3.82***
	(0.085)	(0.069)	(0.109)	-	(0.809)
Italy	-	-	1.07**	0.81***	-0.25
	-	-	(0.027)	(0.028)	(0.194)

Robust standard errors in parentheses
*** p<0.01, ** p<0.05, * p<0.1

Table A5 Two-way interaction estimators comparing the trends in mortality between the period late 1990s – late 2000s and the period late 1980s – late 1990s, three-way interaction estimators comparing the trends in mortality inequalities between the two periods, and four-way interaction estimators, comparing the three-way interaction estimators between England and Finland[a]

	Incidence-rate ratios (Poisson regression)
	all-cause mortality
1. two-way interaction estimators[a]	
England (low- and mid- edu)	0.91***
	(0.022)
2. three-way interaction estimators[b]	
England	0.86**
	(0.060)
Finland	0.94
	(0.060)
3. four-way interaction estimators[c]	
England vs Finland	0.91
	(0.086)

a. We collected and harmonized mortality data by education among people aged 35–79 years in England and Finland between 1980s and 2000s. The period late 1980s – late 1990s was regarded as the period during which the English strategy had not yet been generally implemented and the period late 1990s – late 2000s was regarded as the period during which the main effects of the strategy could be expected. Due to data limitation, we could only classify two levels of education for England ("low+middle" vs "high"), thus the low and middle educated groups were also grouped together for Finland.
b. Based on a two-way interaction analysis for "low+middle" educated people in England. An incidence rate ratio below 1.00 indicates a larger reduction in all-cause mortality in the period late 1990s – late 2000s than in the period late 1980s – late 1990s.
c. Based on a three-way interaction analysis within each country. An incidence rate ratio below 1.00 indicates a more favourable trend in inequalities in mortality in the period late 1990s – late 2000s than in the period late 1980s – late 1990s.
d. Based on a four-way interaction analysis for England and Finland. An incidence rate ratio below 1.00 indicates a more favourable change (between late 1990s – late 2000s and late 1980s – late 1990s) in the trend in inequalities in mortality in England as compared to Finland.
Robust standard errors in parentheses. *** p<0.01, ** p<0.05, * p<0.1

Table A6 Two-way interaction estimators comparing the trends in health between 1990s and 2000s, three-way interaction estimators comparing the trends in health inequalities between 1990s and 2000s, and four-way interaction estimators, comparing the three-way interaction estimators between countries, among men

	Odds ratios (logistic)				Coefficients (linear)
	SAH	LONGHP	smoker	obesity	amount of smoking
1. two-way interaction estimators[a]					
England (low-edu)	0.68***	0.70***	0.90	0.91	-1.14
	(0.086)	(0.088)	(0.119)	(0.142)	(1.052)
2. three-way interaction estimators[b]					
England	1.32	0.79	1.25	1.34	2.36
	(0.304)	(0.149)	(0.261)	(0.325)	(1.741)
Finland	0.86	-	1.55	1.85	4.13
	(0.288)	-	(0.513)	(0.968)	(3.060)
the Netherlands	1.43	1.36	1.47*	-	0.74
	(0.392)	(0.310)	(0.335)	-	(2.506)
Italy	-	-	0.99	0.83	-0.03
	-	-	(0.097)	(0.161)	(0.754)
3. four-way interaction estimator[c]					
England vs Finland	1.54	-	0.80	0.72	-1.77
	(0.627)	-	(0.314)	(0.415)	(3.514)
England vs the Netherlands	0.92	0.59*	0.85	-	1.62
	(0.329)	(0.173)	(0.263)	-	(3.050)
England vs Italy	-	-	1.26	1.61	2.39
	-	-	(0.291)	(0.499)	(1.894)

a. Based on the two-way interaction analysis for low-educated people in England. An odds ratio below 1.00 or a negative linear coefficient indicates a larger health improvement in the period 2000–2010 than in the period 1990–2000.
b. Based on the three-way interaction analysis within each country. An odds ratio below 1.00 or a negative linear coefficient indicates a more favourable trend in health inequalities in the period 2000–2010 than in the period 1990–2000.
c. Based on the four-way interaction analysis for England and each of the comparison countries. An odds ratio below 1.00 or a negative linear coefficient indicates a more favourable change (between 1990–2000 and 2000–2010) in the trend in health inequalities in England as compared to the other country.
Robust standard errors in parentheses. *** p<0.01, ** p<0.05, * p<0.1

Table A7 Two-way interaction estimators comparing the trends in health between 1990s and 2000s, three-way interaction estimators comparing the trends in health inequalities between 1990s and 2000s, and four-way interaction estimators, comparing the three-way interaction estimators between countries, among women

	Odds ratios (logistic)				Coefficients (linear)
	SAH	LONGHP	smoker	obesity	amount of smoking
1. two-way interaction estimators[a]					
England (low-edu)	0.83	0.86	0.76**	1.01	-0.36
	(0.093)	(0.096)	(0.092)	(0.131)	(0.743)
2. three-way interaction estimators[b]					
England	1.08	1.13	1.13	1.07	-0.73
	(0.246)	(0.209)	(0.259)	(0.260)	(1.594)
Finland	0.70	-	1.18	1.87	-2.38
	(0.209)	-	(0.431)	(0.860)	(2.532)
the Netherlands	1.03	1.12	0.72	-	2.00
	(0.270)	(0.241)	(0.180)	-	(2.021)
Italy	-	-	0.92	0.72	2.38***
	-	-	(0.101)	(0.210)	(0.708)
3. four-way interaction estimator[c]					
England vs Finland	1.54	-	0.96	0.57	1.66
	(0.582)	-	(0.414)	(0.296)	(2.985)
England vs the Netherlands	1.04	1.01	1.58	-	-2.73
	(0.362)	(0.288)	(0.534)	-	(2.574)
England vs Italy	-	-	1.23	1.49	-3.10*
	-	-	(0.313)	(0.566)	(1.742)

a. Based on the two-way interaction analysis for low-educated people in England. An odds ratio below 1.00 or a negative linear coefficient indicates a larger health improvement in the period 2000–2010 than in the period 1990–2000.
b. Based on the three-way interaction analysis within each country. An odds ratio below 1.00 or a negative linear coefficient indicates a more favourable trend in health inequalities in the period 2000–2010 than in the period 1990–2000.
c. Based on the four-way interaction analysis for England and each of the comparison countries. An odds ratio below 1.00 or a negative linear coefficient indicates a more favourable change (between 1990–2000 and 2000–2010) in the trend in health inequalities in England as compared to the other country.
Robust standard errors in parentheses. *** p<0.01, ** p<0.05, * p<0.1

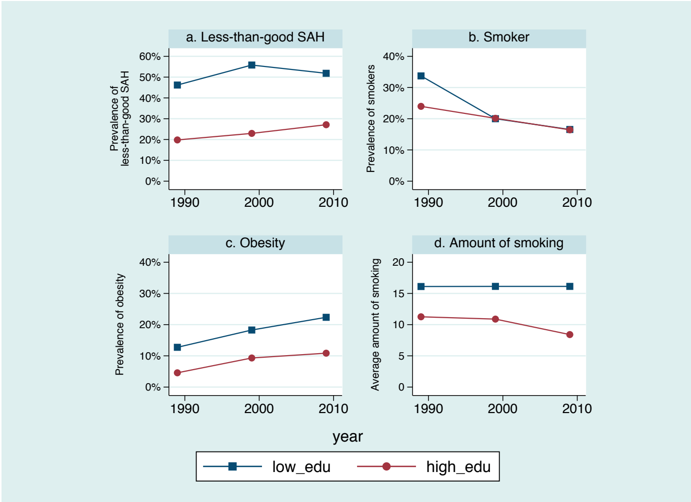


Figure A1 Trends in self-reported health outcomes in Finland by education

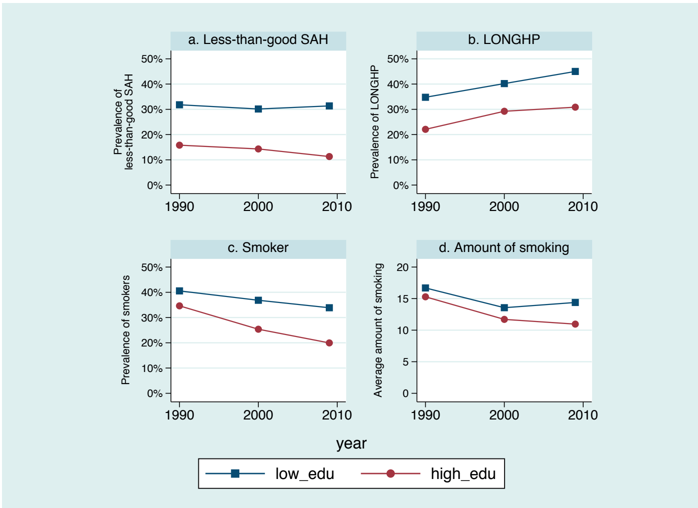


Figure A2 Trends in self-reported health outcomes in the Netherlands by education

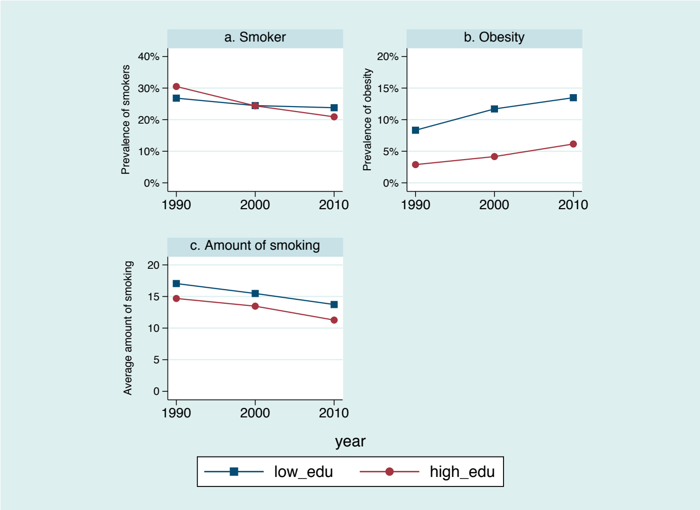


Figure A3 Trends in self-reported health outcomes in Italy by education

Chapter 7

The impact of population-based tobacco control policies on smoking among socioeconomic groups

Yannan Hu, Frank J. van Lenthe, Stephen Platt, Jizzo R. Bosdriesz, Eero Lahelma, Gwenn Menvielle, Enrique Regidor, Paula Santana, Rianne de Gelder, Johan P. Mackenbach. The impact of tobacco control policies on smoking among socioeconomic groups in nine European countries, 1990–2007: a fixed effects study. (submitted)

ABSTRACT

Background

It is uncertain whether tobacco control policies have contributed to a narrowing or widening of socioeconomic inequalities in smoking in European countries during the past two decades. This paper aims to investigate the impact of price and non-price related population-wide tobacco control policies on smoking by socioeconomic group in nine European countries between 1990 and 2007.

Methods

Individual-level education, occupation and smoking status were obtained from nationally representative surveys. Country-level price-related tobacco control policies were measured by the relative price of cheapest cigarettes and of cigarettes in the most popular price category. Country-level non-price policies were measured by a summary score covering four policy domains: smoking bans or restrictions in public places and workplaces, bans on advertising and promotion, health warning labels and cessation services. The associations between policies and smoking were explored using logistic regressions, stratified by education and occupation, and adjusted for age, GDP, period and country fixed effects.

Results

The price of popular cigarettes and non-price policies were negatively associated with smoking among men. The price of the cheapest cigarettes was negatively associated with smoking among women. While these favourable effects were generally in the same direction for all socioeconomic groups, they were larger and statistically significant in lower socioeconomic groups only.

Conclusions

Tobacco control policies as implemented in nine European countries, have probably helped to reduce the prevalence of smoking in the total population, particularly in lower socioeconomic groups. Widening inequalities in smoking may be explained by other factors. Policies with larger effects on lower socioeconomic groups are needed to reverse this trend.

INTRODUCTION

Socioeconomic inequalities in smoking widened in many European countries in the 1980s and 1990s, mainly because the decline in the prevalence of smoking was larger among those in higher as compared to lower socioeconomic groups [1]. Evidence relating to the first decade of the 21st century suggests that the higher educated more often quit smoking, although trends in smoking cessation have varied between countries [2].

During the period of widening inequalities in smoking, there were intensified tobacco control policies in many European countries [3]. For example, comprehensive tobacco control policies were implemented in Great Britain, including health warnings on cigarette packages since 1991, and the introduction of comprehensive NHS smoking cessation services since the late 1990s.[3] Similar developments in tobacco control policies were found in many other countries [4].

It is important to ascertain whether tobacco control policies adopted in European countries affected socioeconomic groups differentially. The most consistent evidence from systematic reviews is that higher prices for cigarettes have had a disproportionately greater impact on the most disadvantaged smokers and as such contributed to a reduction in smoking inequalities [5, 6]. The equity impacts of many other tobacco control policies are less consistent; in fact, voluntary, regional and partial smoke-free policies might have even increased inequalities in smoking [6]. Further, previous studies evaluating the impact of population-level tobacco control policies on smoking inequalities are mainly restricted to a single country (e.g. the United States or the United Kingdom) [7–10] or a small number of countries at best [11–14]. Differences in the societal context in which these policies were implemented, however, may undermine the external validity of the effects. One notable exception is a cross-sectional study of 18 European countries in which a comprehensive package of tobacco control policies, as measured by the tobacco control scale (TCS) [15], was related to higher quit ratios, without significant differences between educational groups [16]. As the original TCS score is only available for a limited number of recent years [15, 17–20], it can only be used for evaluations over a relatively short time period.

For important price and non-price related tobacco control policies, data are available for a longer period of time. Blecher *et al.* developed an indicator of cigarette affordability, defined as the percentage of per capita GDP required to purchase the 100 cheapest packs of cigarettes [21, 22]. Bosdriesz constructed a cigarette price index which measured the relative price of a pack of cigarettes in the most popular price category [23, 24]. Currie developed a non-price-related policy indicator (the Tobacco Control Policy Index (TCPI)), which included four domains of the TCS: smoking bans or restrictions in public places and workplaces, comprehensive bans on advertising and promotion, health warning labels and cessation services [4]. This paper aims to investigate the impact of both price and non-price related tobacco control policies on smoking by socioeconomic group in nine European countries between 1990 and 2007.

DATA AND METHODS

Data

We used country-level data made available by Blecher, Bosdriesz and Currie to assess key aspects of tobacco control policies [4, 21–24]. To measure the affordability of cigarettes, Blecher estimated the relative income price of cigarettes (the percentage of per capita GDP required to purchase 100 packs of the cheapest cigarettes) for a large number of countries for the period 1990–2008, using the cigarette price data from the Economist Intelligence Unit [21, 22]. While the cheapest packs may be most relevant for those in lower socioeconomic groups, they may not be the packs most often bought. We therefore also included a cigarette price index which measured the price of a pack of cigarettes in the most popular price category divided by GDP per capita expressed in purchasing power standards [23, 24]. Currie developed an index for 11 European countries from 1950–2010 [4], which estimated scores for four non-price domains of the TCS [15], which then was rebased to a maximum score of 100 (smoking bans or restrictions (maximum score=40), comprehensive bans on advertising and promotion (maximum score=23.64), health warning labels (maximum score=18.18) and cessation services (maximum score=18.18)). More information about this index is provided in the online supplementary file. Both the data on cigarette price and the non-price index were measured annually.

Data on individual-level smoking status, age, sex and socioeconomic position were obtained from nationally representative health surveys for nine European countries (Finland, Ireland, Great Britain, Austria, the Netherlands, France, Italy, Portugal and Spain) for a number of years between 1990 and 2007 (Table 1). The selected surveys were either identical (for most of the countries) or had a high degree of comparability within country over time (for Austria, the Netherlands and Italy) [25–28]. We linked the individual-level survey data to the country-level data on Blecher's price measure (further referred to as "cheapest cigarette price") or Bosdriesz' price measure (further referred to as "popular cigarette price") and Currie's non-price index (further referred to as "non-price"). For some countries and years, price and non-price policy measures were excluded from the analysis, if the national health survey data of the corresponding countries or years were not available. The final dataset included observations of 563,987 individuals from nine countries with 33 country-year observations between 1990 and 2007. The age range used for most countries in the analysis was 30–79 years. Younger respondents were excluded because many of them were still receiving full-time education. Older respondents were excluded to avoid the potential bias caused by the exclusion of the institutionalized population in most surveys. In some countries, upper age limits were lower than 79 years (France (30–74), Ireland (30–74) and Great Britain (30–69)).

Table 1 Countries included in the analysis, sources of data, age ranges and sample sizes

Country	Survey year	Survey name	Included age range	Number of included respondents[1]
Finland	1993/1995/1997/1999/2001/ 2003/2005/2007	Health Behaviour and Health[2]	30–79	3792 ~ 4069
Ireland	1998/2002/2007	Survey of Lifestyle and Nutrition	30–74	4235 ~ 7638
Great Britain	1990/1996/2000/2005	General Household Survey	30–69	9967 ~ 15722
Austria	1991/1999	Micro Census	30–79	27817 ~ 28817
	2006	Health Interview Survey	30–79	11421
Netherlands	1990	Ongoing Survey of Living Condition (DLO)	30–79	3472
	1997/2000/2005	Permanent Survey on Living Conditions (POLS)	30–79	5665 ~ 6169
France	2000/2005	Baromètre santé	30–74	9641 ~ 20105
Italy	1990	Multipurpose family survey	30–79	38591
	2000/2005	Health and Health Care Utilization	30–79	82040 ~ 87673
Portugal	1995–1996/1998– 1999/2005–2006	National Health Survey	30–79	26091 ~ 30199
Spain	1993/2001/2006	National Health Survey	30–79	14187 ~ 23396

[1] The range of the number of respondents per year for each country.
[2] The Finnish data used are combined from two Finnish studies: "Health behaviour and health among Finnish adult population (AVTK)", which includes respondents who are 15–64 years old, and "Health behaviour and health among the Finnish elderly (EVTK)", which includes respondents who are older than 64 years.

Smoking status was defined as current smoking (daily or occasional smoking in all countries, except in Austria where only daily smoking was included). Socioeconomic position was measured by educational level and occupational class. Educational levels were recorded as the highest level of education completed or currently attended by a person. Data were harmonized on the basis of the International Standard Classification of Education (ISCED) and reclassified into three categories: levels 0–2 (no, primary or lower secondary education, considered "low education"); levels 3–4 (upper secondary and post-secondary non-tertiary education, considered "middle education"); and levels 5–6 (tertiary education, considered "high education"). Occupational classes were classified as "manual" versus "non-manual". Farmers and self-employed respondents were excluded from the occupation-related analyses. Economically inactive respondents who could not be classified on the basis of their last or main occupation, were classified as missing.

Gross Domestic Product (GDP) per capita was used as a confounding variable, as it may be related to both smoking behaviour and the implementation of tobacco control policies [16]. Moreover, it may also be used as a proxy measure of the stages in the spread of the smoking epidemic [29]. GDP per capita (constant prices, constant purchasing power parity rates, expressed in dollars) was extracted from OECD Stat (http://stats.oecd.org/).

Statistical methods

Logistic regressions were used to analyse the associations between the country-level tobacco control policy measures and the individual-level smoking status. We applied fixed-effects analysis by adding country dummy variables into the models. Fixed effects analysis compares multiple observations over time, and reveals the average change in the outcome due to the policies. Because each country is compared with itself over time, differences between countries that remain constant over time – even if unmeasured – are eliminated and cannot confound the results. It allowed us to adjust for unobserved time-invariant country heterogeneities related to both the implementation of tobacco control efforts and smoking in countries. Additionally, we adjusted the analysis for age, age squared, logarithmic form of GDP and period dummy variables. Clustered sandwich estimators were used to allow for within-country correlation between error terms [30].

The basic model can be written as:

$$smoking_{ijt} = \beta_0+\beta_1\, ln(price)_{jt} + \beta_2\, (non\text{-}price)_{jt} + \beta_3\, age_{ijt} +\beta_4\, age^2_{ijt} + \beta_5\, lngdp_{jt} + T_t + X_j$$

where $smoking_{ijt}$ is a dummy variable indicating whether the respondent i in country j was a current smoker in year t; β_0 is a constant; $ln(price)_{jt}$ represents the logarithmic form of the cigarette price (either cheapest or popular cigarette price) for country j in year t, accounting for the potential non-linear relationship between price and smoking; $(non\text{-}price)_{jt}$ is the non-price policy measure for country j in year t; age_{ijt} and age^2_{ijt} are the age and age squared of the respondent; $lngdp_{jt}$ is the logarithmic form of GDP per capita; T_t is a vector of the 5-year period dummies controlling the shared time trend in smoking; X_j is a vector of country dummies, controlling the time-invariant country heterogeneities in propensity towards smoking or implementing tobacco control policies, which are affected by the long-term country-specific cultural or political factors, e.g. the religious composition of population [31] or the power of the tobacco industry [32, 33]. Additionally, by adding country dummies we controlled for the potential national differences in reporting or recording smoking behaviours.

The model was run separately for men and women. Given that our primary interest was to assess whether tobacco control policies affected socioeconomic groups equally, we further stratified the analysis by education and occupation. To empirically test whether the associations between the policy measures and smoking differed significantly between educational or occupational groups, interactions between the policy measures and education or occupation were added, together with the corresponding interactions between the other variables in the basic model and education or occupation.

In the online supplementary file, supplementary analyses include an analysis a) incorporating available weighting factors in surveys of some countries or years (Table S1); b) using the amount of smoking per day among current smokers as the outcome (Table S2); c) allowing potential lagged effects of non-price policies (Table S3); d) combining data for men and women (Table S4); e) using a constructed measure combining the price and non-price measures (Table S5); f) linking each of the four domains of the non-price score to smoking (Table S6).

All regression analyses were performed in Stata 13.1.

RESULTS

Summary statistics of key variables are presented in Table 2. The distribution of socioeconomic position differed between countries, with a higher percentage of high-educated persons in Ireland, France and Great Britain, and a higher percentage of non-manual workers in France, the Netherlands and Finland (especially among women). The smoking prevalence was higher among men (ranging from 26% (Ireland) to 41% (Spain)) than among women (ranging from 8% (Portugal) to 29% (the Netherlands)) in all countries. The mean GDP was higher in Ireland and the Netherlands, and was generally lower in the Mediterranean countries. Averaged over time, the cheapest cigarette price index was highest in Great Britain and lowest in Spain; the popular cigarette price index was highest in Ireland and lowest in Spain. The non-price policy measure was highest in Ireland and lowest in Austria.

Figure 1 shows the trends in the price and non-price measures by country over the study period. A relatively stable trend in the cheapest cigarette price index was observed in most of the countries, except for an increasing trend in the Netherlands, France and Great Britain, and a decreasing trend in Ireland. An increasing trend in the popular cigarette price index was observed in most of the countries, except for a relatively stable trend in Finland. The non-price indicator constantly increased over time in all available countries.

Online supplementary figures S1.a and S1.b show the trends in age-standardised smoking prevalence by education in each country for men and women separately over the study period [34]. Among men, smoking was more prevalent in the low-educated group in almost all countries and years. The smoking prevalence generally decreased to the same extent within each educational group. Among women, smoking was more prevalent among the lower-educated group in many countries. While the smoking prevalence among the low-educated was generally stable, it often declined among the high-educated. Deviant trends were found in Italy, Portugal and Spain, where the smoking prevalence was higher among high-educated in the 1990s, but almost equal to the prevalence among the low-educated in 2005. Reasonably similar trends were observed in age-standardised smoking prevalence by occupation (online supplementary figure S2.a and S2.b).

Table 3 shows the results from logistic regressions linking the tobacco control policy measures to smoking. In the analysis including the cheapest cigarette price and the non-price

Table 2 Summary statistics of key variables, pooled for all years in each country

	Finland	Ireland	Great Britain	Austria	Netherlands	France	Italy	Portugal	Spain
Gender									
Men	47%	43%	48%	46%	48%	43%	48%	47%	44%
Women	53%	57%	52%	54%	52%	57%	52%	53%	56%
Men									
Age									
30–39	16%	28%	29%	24%	26%	28%	25%	21%	26%
40–49	18%	28%	27%	25%	25%	25%	24%	23%	24%
50–59	19%	20%	24%	22%	21%	23%	21%	21%	19%
60–69	21%	17%	20%	18%	17%	18%	18%	20%	18%
70–79	26%	7%	-	11%	11%	6%	12%	15%	13%
Education									
ISCED 0–2	45%	23%	37%	24%	38%	23%	63%	87%	64%
ISCED 3–4	39%	46%	34%	68%	36%	46%	28%	7%	19%
ISCED 5–6	16%	31%	29%	8%	26%	31%	9%	6%	17%
Occupation									
Manual	48%	55%	45%	46%	39%	31%	52%	59%	73%
non-manual	52%	45%	55%	54%	61%	69%	48%	41%	27%
Smoking status									
current smoker	27%	26%	28%	30%	36%	34%	32%	34%	41%
ex/ never smoker	73%	74%	72%	70%	64%	66%	68%	66%	59%
Women									
Age									
30–39	17%	33%	29%	23%	27%	26%	24%	20%	24%
40–49	21%	28%	27%	23%	24%	23%	22%	22%	22%
50–59	19%	18%	23%	21%	21%	24%	21%	21%	19%
60–69	20%	14%	21%	19%	16%	19%	18%	21%	19%
70–79	23%	7%	-	14%	12%	8%	15%	16%	16%

Table 2 Summary statistics of key variables, pooled for all years in each country (continued)

	Finland	Ireland	Great Britain	Austria	Netherlands	France	Italy	Portugal	Spain
Education									
ISCED 0–2	39%	16%	51%	46%	52%	31%	66%	87%	72%
ISCED 3–4	43%	51%	25%	49%	30%	40%	26%	6%	14%
ISCED 5–6	17%	33%	24%	5%	18%	29%	8%	7%	14%
Occupation									
Manual	16%	50%	25%	30%	24%	11%	50%	40%	77%
non-manual	84%	50%	75%	70%	76%	89%	50%	60%	23%
Smoking status									
current smoker	17%	24%	27%	19%	29%	26%	18%	8%	21%
ex/ never smoker	83%	76%	73%	81%	71%	74%	82%	92%	79%
GDP[1]									
Mean	27679	37555	29629	30454	33716	29959	28322	20068	25720
Min	21057	29463	24886	26991	27820	29043	24566	18136	20407
Max	34813	42114	34519	35680	36974	30398	29544	22073	28569
Cheapest cigarette price[2]									
Mean	1.40	1.64	2.25	1.15	0.97	1.45	1.08	1.67	0.93
Min	1.24	1.55	1.86	1.10	0.82	1.12	0.98	1.65	0.80
Max	1.50	1.79	2.65	1.22	1.07	1.60	1.20	1.72	0.99
Popular cigarette price[3]									
Mean	14.44	17.75	17.06	8.50	8.48	17.27	8.89	9.36	7.20
Min	12.90	13.50	8.30	7.60	5.20	11.80	4.40	7.10	2.80
Max	15.60	21.40	24.50	12.00	12.30	19.90	12.50	13.20	9.60
Non-price policy[4]									
Mean	55.79	55.80	29.78	14.46	35.37	49.51	55.55	35.26	49.73
Min	43.64	24.55	9.09	1.82	15.45	40.91	30.00	30.00	25.45
Max	68.18	86.36	52.73	38.18	60.91	53.64	75.45	46.36	73.64

[1] GDP per capita (constant prices, constant purchasing power parity rates, expressed in dollars).
[2] Relative income price of cigarettes (the percentage of per capita GDP required to purchase 100 cheapest packs).
[3] The index measured the price of a pack of cigarettes in the most popular price category divided by GDP per capita expressed in purchasing power standards.
[4] Tobacco Control Policy Index estimating scores for four non-price domains of the TCS (ranging from 0 to 100).

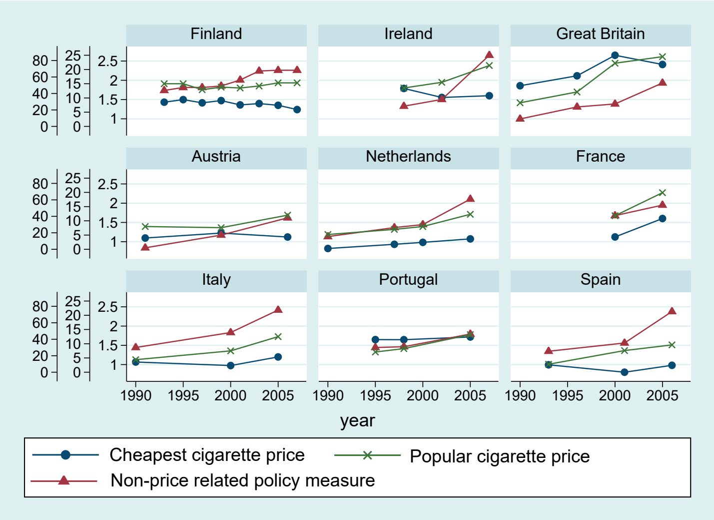


Figure 1 Trends in tobacco control policies in each country

Notes:

[1]The cheapest cigarette price is the percentage of per capita GDP required to purchase 100 cheapest packs.

[2]The popular cigarette price is the index measuring the price of a pack of cigarettes in the most popular price category divided by GDP per capita expressed in purchasing power standards.

[3]The non-price related policy measure is the Tobacco Control Policy Index estimating scores for four non-price domains of the TCS (ranging from 0 to 100).

[4]The three y-axes represent the cheapest cigarette price, the popular cigarette price and the non-price related policy measure (from right to left).

[5]The policy measures for the United Kingdom were used to represent the policy measures for Great Britain.

policy measure, significant associations were found between the non-price-related policy index and smoking among men (OR=0.95, 95% CI (0.91~0.99)), and between the cheapest cigarette price index and smoking among women (OR=0.50, 95% CI (0.34~0.75)). Thus, a 10 units increase in the non-price-related policy index was associated with a 5% decrease in the odds of smoking among men, and a 2.72-fold increase in the cheapest cigarette price index (the cigarette price was log-transformed) was associated with a 50% decrease in the odds of smoking among women. Stratified by education or occupation, most odds ratios were still below 1 and significant associations were found in low socioeconomic groups only. Specifically, the non-price related policy measure was negatively related to smoking among men with low education (OR=0.96, 95% CI (0.92~0.99)), and the cheapest cigarette price index was negatively related to smoking among women with low education (OR=0.55, 95% CI (0.37~0.82)) or manual jobs (OR=0.49, 95% CI (0.35~0.68)). While the odds ratios were always smaller in the low socioeconomic groups, they were not statistically different from those in the high socioeconomic groups, as indicated by the p-value of the interactions.

In analyses using the popular cigarette price index, men were found to be more responsive to the price change than women. With regard to the impact on smoking by socioeconomic

Table 3 Associations between price and non-price related tobacco control policy measures and smoking status among men and women, stratified by education and occupation

	Men				Women			
	Cheapest cigarette price[1]		Non-price[2]		Cheapest cigarette price[1]		Non-price[2]	
	OR[3,4]	95% CI	OR	95% CI	OR	95% CI	OR	95% CI
Total population	0.89	(0.63 ~ 1.26)	**0.95**	**(0.91 ~ 0.99)**	**0.50**	**(0.34 ~ 0.75)**	0.96	(0.88 ~ 1.05)
Low education	0.99	(0.60 ~ 1.63)	**0.96**	**(0.92 ~ 0.99)**	**0.55**	**(0.37 ~ 0.82)**	0.94	(0.85 ~ 1.05)
High education	1.22	(0.73 ~ 2.05)	0.97	(0.90 ~ 1.04)	0.89	(0.55 ~ 1.45)	1.00	(0.95 ~ 1.05)
P value[5]	0.260		0.623		0.128		0.444	
Manual[6]	0.85	(0.64 ~ 1.14)	0.98	(0.91 ~ 1.05)	**0.49**	**(0.35 ~ 0.68)**	0.96	(0.88 ~ 1.04)
Non-manual	1.01	(0.64 ~ 1.59)	0.98	(0.93 ~ 1.04)	0.66	(0.43 ~ 1.03)	0.94	(0.83 ~ 1.06)
P value[5]	0.309		0.600		0.196		0.691	
	Popular cigarette price[7]		Non-price[2]		Popular cigarette price[7]		Non-price[2]	
	OR[3,4]	95% CI	OR	95% CI	OR	95% CI	OR	95% CI
Total population	**0.86**	**(0.74 ~ 0.99)**	**0.96**	**(0.92 ~ 0.995)**	1.04	(0.67 ~ 1.63)	0.96	(0.88 ~ 1.04)
Low education	0.85	(0.70 ~ 1.03)	**0.96**	**(0.93 ~ 0.995)**	1.09	(0.76 ~ 1.56)	0.93	(0.85 ~ 1.03)
High education	1.15	(0.81 ~ 1.64)	0.97	(0.90 ~ 1.04)	0.77	(0.54 ~ 1.10)	1.00	(0.96 ~ 1.05)
P value[5]	0.145		0.915		0.258		0.272	
Manual[6]	**0.80**	**(0.67 ~ 0.94)**	0.98	(0.91 ~ 1.06)	0.87	(0.59 ~ 1.28)	0.97	(0.89 ~ 1.05)
Non-manual	0.84	(0.66 ~ 1.06)	0.99	(0.94 ~ 1.04)	0.90	(0.69 ~ 1.17)	0.96	(0.83 ~ 1.10)
P value[5]	0.651		0.524		0.890		0.817	

[1] The relative income price of cigarettes (the percentage of per capita GDP required to purchase the 100 cheapest packs of cigarettes).
[2] Summary scores for four domains of non-price related policies (smoking bans or restrictions, comprehensive bans on advertising and promotion, health warning labels and cessation services; ranging from 0 to 100). The odds ratios for the non-price related policies relate to an increase of 10 units in the scores.
[3] Odds ratios and 95% confidence intervals based on clustered standard errors are reported, derived from logistic regressions.
[4] All models were adjusted by age, age squared, logarithmic form of GDP, periods and country dummies. Significant associations and significant interactions ($p<0.05$) are highlighted in bold.
[5] P-value for the interactions between the tobacco control policy measures and education or occupation.
[6] In the stratified analysis based on occupation, some recent years for Finland were excluded because information on occupation was not available.
[7] The index measured the price of a pack of cigarettes in the most popular price category divided by GDP per capita expressed in purchasing power standards

group, however, the results were similar to those found for the cheapest cigarette price index. The potentially favourable associations were statistically significant in lower socioeconomic groups only.

Broadly similar results were obtained in a sensitivity analysis (table S1), in which weighting factors were incorporated if available. Using the daily amount of smoking as the outcome, significant associations between tobacco control policy measures and smoking were mainly found again in the lower socioeconomic groups (table S2). Analyses in which the 1-year and 2-year lagged terms of the non-price policy measure were additionally added revealed significant results mainly in respect of the contemporaneous non-price policy measure (table S3). Consistent with the main findings, none of the interactions between the contemporaneous or lagged non-price measures and socioeconomic position was significant. Essentially similar results were obtained when data from men and women were combined (table S4).

DISCUSSION

Summary of findings

The price of popular cigarettes and non-price policies were negatively associated with smoking among men. The price of the cheapest cigarettes was negatively associated with smoking among women. While these favourable effects were generally in the same direction for all socioeconomic groups, they were larger and statistically significant in lower socioeconomic groups only.

Strengths and limitations

By covering nine European countries, multiple elements of tobacco control efforts and a period of nearly 20 years, we comprehensively investigated the impact of tobacco control policies on smoking by socioeconomic group. Both education and occupation were used to capture the multidimensional concept of socioeconomic position. Fixed-effects models were used to reduce potential bias caused by unobserved country characteristics related to both smoking and the implementation of tobacco control policies. As such, these models improved the causal inference of the relationships reported.

Although we were able to control for unobserved time-invariant confounding, the results may be still biased due to omitted time-variant confounders [30]. For example, increasing knowledge about the health consequences of smoking may have been related to the implementation of tobacco control policies, as well as to changes in smoking prevalence. Further studies may consider the inclusion of potentially omitted time-variant confounders, if appropriate data are available.

By using the TCPI, we tried to cover the most important domains of non-price-related policies. The TCPI excluded the domain of public (mass media) information campaigns, which were present in some countries. For example, in the Netherlands a large campaign ("The Netherlands starts quitting/The Netherlands continues with quitting") was implemented in 2003–2004 [35], and in England a free telephone helpline advertised by TV and radio was introduced in 1994 [36]. There is no consistent evidence on the equity impact of mass media campaigns [6], although some existing European studies suggest an effect of widening of inequalities [36–38]. Thus, it cannot be excluded that our findings may become less favourable when mass media campaigns are included into the policy measures. If appropriate data become available, this should be subject of further investigation.

We included measures for price and non-price tobacco control efforts simultaneously in the analysis. Mutual adjustment enabled us to distinguish their potentially different impacts on smoking inequalities [6]. However, this approach does not fully capture the equity impact of a comprehensive tobacco control implemented as a whole. We therefore constructed a summary measure of the price and non-price indicators using the same weights as the TCS [15]. Analysis using this summary measure showed that more tobacco control efforts were significantly related to less smoking among men and among women in lower socioeconomic groups (Table S5). Again, significant favourable results were found in lower socioeconomic groups only.

Our analyses covered the period between 1990 and 2007. Data on smoking and tobacco control policies were available for nine European countries, for which the data were mostly available. As such, our findings cannot be generalized to the most recent years and to Europe as a whole. Future research may consider to cover more recent years, to include more countries, and to use annual data.

Interpretation

Our results suggest that both price and non-price tobacco control efforts as implemented in the nine European countries have helped to reduce smoking among the total population and among people in low socioeconomic groups. This is consistent with some previous findings [5, 6, 39]. Our findings about the negative association between tobacco control policies and the overall smoking prevalence are generally consistent with those using the TCS [40, 41]. The number of studies analysing the associations between the TCS and smoking inequalities is very limited. Apart from one study which focused on one specific country [23], the only international study we are aware of is one covering 18 European countries around the year 2000 [16]. It found that the TCS was related to high quit ratios among both educated groups, and there were no consistent differences in policy effects between education levels when the subgroups based on age and gender were studied. Negative cross-sectional associations for both educational levels were also found with our data (results available upon request). However, it is important to notice that we have controlled for the fixed effects of country and period, which can reduce potential bias caused by unobserved country characteristics and common global shocks over time. Besides the usage of the fixed effects models, subtle differences between studies may also play a role, such as differences in the study periods and study countries, and the different policy coverage between the TCS and our policy measures.

The majority of existing studies suggest that increasing the cigarette price may be more effective in reducing smoking among lower socioeconomic groups, although there is also some evidence suggesting neutral or negative equity impacts of the cigarette price [5, 6]. In our study, no matter which price measure we used, the significantly negative associations between price and smoking were only found among the lower socioeconomic groups, suggesting that people in lower socioeconomic groups are more responsive to price change. People in higher socioeconomic groups may not regard the cost of cigarettes as a heavy financial burden, which would make them less sensitive to the change of cigarette price [42]. There are also some other countervailing factors, however, which might weaken the potential equity impact of the cigarette price. For example, people in lower socioeconomic groups are more frequently heavier smokers and tobacco-dependent, and may turn to hand-rolled cigarettes, discount brands and black market when the cigarette price increases [43, 44]. Moreover, smokers in lower socioeconomic groups may satisfy more essential needs with smoking (e.g. the reduction of negative feelings, cheap leisure), and harsh living conditions may make quitting smoking harder [43]. In combination, these factors could make the differences in the reactions to the cigarette price between socioeconomic groups much smaller than expected. As our study suggests, while the impact was larger and statistically significant among lower socioeconomic groups only, differences in the associations between price and smoking were not significant between socioeconomic groups.

Of the four sub-domains of the non-price policy measure to smoking, none consistently showed a differential impact by socioeconomic position on smoking (Table S6). This is consistent with the uncertain equity impact of the non-price tobacco control efforts reported in previous reviews [5, 6]. To interpret the findings, we have to be aware that policy indicators, such as the TCPI, measure the extent to which tobacco control policies have been formulated, but contain little information on their implementation and enforcement [45]. It is possible that policies were not well implemented and faced considerable problems with compliance in some European countries [3, 46]. If we were able to measure the actual delivery of these non-price policies, the equity impact on smoking might be larger.

The effects of tobacco control policies were found to be larger and statistically significant among the lower socioeconomic groups only, which suggests that these policies may have contributed to a reduction of smoking inequalities. We tried to quantify the impact of these policies on smoking inequalities, using the average smoking prevalence in all countries by education in the first available year of our study, and the change of smoking by education over a ten-year period as predicted by the average yearly change in the policy indexes (results available upon request). We found that, keeping all other variables constant, educational inequalities would attenuate slightly among men (a reduction of relative inequalities by 12%, and a reduction of absolute inequalities by 16%), and more substantially among women (a reduction of relative inequalities by 42% and a reduction of absolute inequalities by 43%, in countries where smoking was more prevalent among low-educated women around 1990). This suggests that tobacco control policies, as implemented in the nine European countries, have not contributed to a widening of smoking inequalities, but may instead have contributed to a certain degree of narrowing. Other factors may have been responsible for this widening trend in smoking inequalities, such as differences across socioeconomic groups in knowledge of the risks of smoking [47, 48], increases in income inequalities [49], mass media quit smoking campaigns [36–38], and the economic recessions in some European countries (e.g. in the early 1990s and the late 2000s) [50].

Conclusions

Tobacco control policies as implemented in nine European countries, have probably helped to reduce the prevalence of smoking in the total population, particularly in lower socioeconomic groups. Widening inequalities in smoking may be explained by other factors. Policies with larger effects on lower socioeconomic groups are needed to reverse this trend.

REFERENCES

1. Giskes, K., et al., *Trends in smoking behaviour between 1985 and 2000 in nine European countries by education*. J Epidemiol Community Health, 2005. **59**(5): p. 395–401.
2. Bosdriesz, J.R., et al., *Socioeconomic inequalities in smoking cessation in 11 European countries from 1987 to 2012*. J Epidemiol Community Health, 2015. **69**(9): p. 886–92.
3. Currie, L. and A.B. Gilmore, *Tobacco*, in *Successes and Failures of Health Policy in Europe: four decades of divergent trends and converging challenges*, J.P. Mackenbach and M. Mckee, Editors. 2013, Open University Press: Buckingham.
4. Nguyen, L., G. Rosenqvist, and M. Pekurinen, *Demand for Tobacco in Europe: An Econometric Analysis of 11 Countries for the PPACTE Project*. 2012, Tampere, Finland: Juvenes Print.
5. Thomas, S., et al., *Population tobacco control interventions and their effects on social inequalities in smoking: systematic review*. Tob Control, 2008. **17**(4): p. 230–7.
6. Brown, T., S. Platt, and A. Amos, *Equity impact of population-level interventions and policies to reduce smoking in adults: a systematic review*. Drug Alcohol Depend, 2014. **138**: p. 7–16.
7. Colman, G.J. and D.K. Remler, *Vertical equity consequences of very high cigarette tax increases: If the poor are the ones smoking, how could cigarette tax increases be progressive?* Journal of Policy Analysis and Management, 2008. **27**(2): p. 376–400.
8. Dinno, A. and S. Glantz, *Tobacco control policies are egalitarian: A vulnerabilities perspective on clean indoor air laws, cigarette prices, and tobacco use disparities*. Social Science & Medicine, 2009. **68**(8): p. 1439–1447.
9. Hawkins, S.S., T.J. Cole, and C. Law, *Examining smoking behaviours among parents from the UK Millennium Cohort Study after the smoke-free legislation in Scotland*. Tobacco Control, 2011. **20**(2): p. 112–118.
10. Taggar, J.S., et al., *The impact of the Quality and Outcomes Framework (QOF) on the recording of smoking targets in primary care medical records: cross-sectional analyses from The Health Improvement Network (THIN) database*. BMC Public Health, 2012. **12**.
11. Nagelhout, G.E., et al., *Prevalence and predictors of smoking in "smoke-free" bars. Findings from the International Tobacco Control (ITC) Europe Surveys*. Social Science & Medicine, 2011. **72**(10): p. 1643–51.
12. Hitchman, S.C., et al., *Effectiveness of the European Union text-only cigarette health warnings: findings from four countries*. Eur J Public Health, 2012. **22**(5): p. 693–9.
13. Kasza, K.A., et al., *The effectiveness of tobacco marketing regulations on reducing smokers' exposure to advertising and promotion: findings from the International Tobacco Control (ITC) Four Country Survey*. Int J Environ Res Public Health, 2011. **8**(2): p. 321–40.
14. King, B.A., et al., *Socioeconomic variation in the prevalence, introduction, retention, and removal of smoke-free policies among smokers: findings from the International Tobacco Control (ITC) Four Country Survey*. Int J Environ Res Public Health, 2011. **8**(2): p. 411–34.
15. Joossens, L. and M. Raw, *The Tobacco Control Scale: a new scale to measure country activity*. Tob Control, 2006. **15**(3): p. 247–53.
16. Schaap, M.M., et al., *Effect of nationwide tobacco control policies on smoking cessation in high and low educated groups in 18 European countries*. Tob Control, 2008. **17**(4): p. 248–55.
17. Joossens, L. *Effective Tobacco Control Policies in 28 European Countries*. 2004.
18. Joossens, L. and M. Raw *Progress in Tobacco Control in 30 European Countries, 2005 to 2007*. 2007.
19. Joossens, L. and M. Raw *The Tobacco Control Scale 2010 in Europe*. 2011.
20. Joossens, L. and M. Raw *The Tobacco Control Scale 2013 in Europe*. 2014.

21. Blecher, E.H., *The Economics of Tobacco Control in Low- and Middle-income Countries*. 2011: Thesis presented for the degree of Doctor or Philosophy, School of Economics, University of Cape Town. Available online at: http://uctscholar.uct.ac.za/PDF/32601_Blecher,%20E.pdf (accessed 12 Nov 2015).
22. Blecher, E.H. and C.P. van Walbeek, *Cigarette affordability trends: an update and some methodological comments.* Tob Control, 2009. **18**(3): p. 167–75.
23. Bosdriesz, J.R., et al., *The Association Between Tobacco Control Policy and Educational Inequalities in Smoking Cessation in the Netherlands from 1988 Through 2011.* Nicotine Tob Res, 2015. **17**(11): p. 1369–76.
24. Bosdriesz, J.R., et al., *Patterns of Tobacco Control Policy Progress in 27 European Countries.* Tobacco Regulatory Science, 2015. **1**(3): p. 254–264.
25. Großschadl, F. and W.J. Stronegger, *Long-term trends in obesity among Austrian adults and its relation with the social gradient: 1973–2007*. European Journal of Public Health, 2013. **23**(2): p. 306–12.
26. European Foundation for the Improvement of Living and Working Conditions, *Trends in quality of work in the Netherlands*. 2007: http://eurofound.europa.eu/sites/default/files/ef_files/ewco/surveys/NL0601SR01/NL0601SR01.pdf (accessed 11 December 2014).
27. Italian National Institute of Statistics (ISTAT), *Condizioni di salute e ricorso ai servizi sanitari [Health and Health Care Utilization]*. Italian National Institute of Statistics, http://www.istat.it/it/archivio/5471 (accessed 12 December 2014).
28. Italian National Institute of Statistics (ISTAT), *Multipurpose Family Survey*. Italian National Institute of Statistics, http://www.istat.it/it/archivio/4394 (accessed 16 July 2015).
29. Schaap, M., *Socioeconomic inequalities in smoking in Europe*. 2010: PhD dissertation. Erasmus University Rotterdam.
30. Wooldridge, J.M., *Econometric analysis of cross section and panel data (second edition)*. 2010: MIT Press.
31. Garrusi, B. and N. Nakhaee, *Religion and Smoking: A Review of Recent Literature.* International Journal of Psychiatry in Medicine, 2012. **43**(3): p. 279–292.
32. McKee, M. and P. Diethelm, *How the growth of denialism undermines public health.* BMJ, 2010. **341**: p. c6950.
33. Gilmore, A.B., et al., *Understanding tobacco industry pricing strategy and whether it undermines tobacco tax policy: the example of the UK cigarette market.* Addiction, 2013. **108**(7): p. 1317–26.
34. Ahmad, O.B., et al., *Age Standardization of Rates: a New WHO Standard.* . GBP Discussion Paper Series No. 31. World Health Organization http://www.who.int/healthinfo/paper31.pdf (accessed 10 October 2014), 2001.
35. van den Putte, B., et al. *Nederlands Start Met Stoppen/Nederlands Gaat Door Met Stoppen. Evaluatie van de STIVORO campagnes rondom de jaarwisseling 2003–2004*. 2005.
36. Owen, L., *Impact of a telephone helpline for smokers who called during a mass media campaign*. Tob Control, 2000. **9**(2): p. 148–54.
37. van Osch, L., et al., *Optimizing the efficacy of smoking cessation contests: an exploration of determinants of successful quitting.* Health Educ Res, 2009. **24**(1): p. 54–63.
38. Wiebing, M.A., S.M. Bot, and M.C. Willemsen, *Rokers verdienen 'n beloning', de 24-uur-niet-rokenactie. Blootstelling, deelname en effect op stopbereidheid bij hoge en lage welstandsgroepen.* Tijdschrift voor Gezondheidswetenschappen, 2010. **88**: p. 435–441.
39. Gallus, S., et al., *Price and cigarette consumption in Europe.* Tob Control, 2006. **15**(2): p. 114–9.
40. Gallus, S., et al., *Pricing Policies And Control of Tobacco in Europe (PPACTE) project: cross-national comparison of smoking prevalence in 18 European countries.* European Journal of Cancer Prevention, 2013. **23**(3): p. 177–185.

41. Martinez-Sanchez, J.M., et al., *Smoking behaviour, involuntary smoking, attitudes towards smoke-free legislations, and tobacco control activities in the European Union*. PLoS One, 2010. **5**(11): p. e13881.
42. Peretti-Watel, P., O. L'Haridon, and V. Seror, *Responses to increasing cigarette prices in France: how did persistent smokers react?* Health Policy, 2012. **106**(2): p. 169–76.
43. Peretti-Watel, P. and J. Constance, *"It's all we got left". Why poor smokers are less sensitive to cigarette price increases.* Int J Environ Res Public Health, 2009. **6**(2): p. 608–21.
44. Licht, A.S., et al., *Socio-economic variation in price minimizing behaviors: findings from the International Tobacco Control (ITC) Four Country Survey.* Int J Environ Res Public Health, 2011. **8**(1): p. 234–52.
45. Bosdriesz, J.R., et al., *Tobacco control policy development in the European Union: do political factors matter?* Eur J Public Health, 2015. **25**(2): p. 190–4.
46. Nagelhout, G.E., et al., *Age and educational inequalities in smoking cessation due to three population-level tobacco control interventions: findings from the International Tobacco Control (ITC) Netherlands Survey.* Health Educ Res, 2013. **28**(1): p. 83–91.
47. Layte, R. and C.T. Whelan, *Explaining Social Class Inequalities in Smoking: The Role of Education, Self-Efficacy, and Deprivation.* European Sociological Review, 2009. **25**(4): p. 399–410.
48. Kickbusch, I.S., *Health literacy: addressing the health and education divide.* Health Promotion International, 2001. **16**(3): p. 289–297.
49. Jutz, R., *The role of income inequality and social policies on income-related health inequalities in Europe.* Int J Equity Health, 2015. **14**(1): p. 117.
50. Bacigalupe, A. and A. Escolar-Pujolar, *The impact of economic crises on social inequalities in health: what do we know so far?* Int J Equity Health, 2014. **13**: p. 52.

SUPPLEMENTARY FILE

The Tobacco Control Policy Index(TCPI) developed by Currie

The classical TCS score covers six tobacco control areas (100 score in total), including price (maximum score=30), smoking free legislation (maximum score=22), public campaigns (maximum score 15), bans on advertising (maximum score=13), health warning labels (maximum score=10) and smoking cessation service (maximum score=10). However, it was only available for limited countries and years. Using the same weight, Currie developed an index for 11 European countries from 1950–2010, which estimated scores for four non-price domains of the TCS (excluding "price" and "public campaigns"). Therefore, the original maximum score of Currie index was 55 ("smoking free legislation" + "bans on advertising" + "health warning labels" + "smoking cessation" = 22+13+10+10=55), which was rebased to a maximum score of 100. Then the final maximum score for "smoking free legislation" is 22/55*100=40. The final maximum score for "comprehensive bans on advertising and promotion" is 13/55*100=23.64. The final maximum score for "health warning labels" is 10/55*100=18.18. The final maximum score for "cessation services" is 10/55*100=18.18.

More information about the breakdown of the four non-price policy domains, annual data for each country and the country-specific description of TCPI over the study period can be found in:

Currie, L. Appendix A. Tobacco Control Policy Index. In the Nguyen L, Rosenqvist G, Pekurinen M. Demand for Tobacco in Europe: An Econometric Analysis of 11 Countries for the PPACTE Project. Tampere, Finland: Juvenes Print, 2012.

Table S1 Associations between price and non-price related tobacco control policy measures and smoking status among men and women, stratified by education and occupation, weighted results

	Men				Women			
	Cheapest cigarette price [1]		Non-price[2]		Cheapest cigarette price [1]		Non-price[2]	
	OR[3,4]	95% CI	OR	95% CI	OR	95% CI	OR	95% CI
Total population	0.98	(0.70 ~ 1.35)	**0.93**	**(0.90 ~ 0.96)**	**0.54**	**(0.34 ~ 0.85)**	0.94	(0.85 ~ 1.04)
Low education	1.00	(0.61 ~ 1.66)	**0.95**	**(0.92 ~ 0.98)**	**0.53**	**(0.37 ~ 0.75)**	0.94	(0.85 ~ 1.04)
High education	1.31	(0.85 ~ 2.04)	0.96	(0.90 ~ 1.02)	0.98	(0.60 ~ 1.60)	0.96	(0.92 ~ 1.01)
P value[5]	0.056		0.696		**0.035**		0.706	
Manual[6]	0.91	(0.73 ~ 1.13)	0.96	(0.91 ~ 1.01)	**0.58**	**(0.40 ~ 0.84)**	0.93	(0.85 ~ 1.01)
Non-manual	1.18	(0.80 ~ 1.74)	0.96	(0.90 ~ 1.03)	0.67	(0.43 ~ 1.05)	0.92	(0.81 ~ 1.05)
P value[5]	0.122		0.786		0.520		0.926	

[1] The relative income price of cigarettes (the percentage of per capita GDP required to purchase the 100 cheapest packs of cigarettes).

[2] Summary scores for four domains of non-price related policies (smoking bans or restrictions, comprehensive bans on advertising and promotion, health warning labels and cessation services; ranging from 0 to 100). The odds ratios for the non-price related policies relate to an increase of 10 units in the scores.

[3] Odds ratios and 95% confidence intervals based on clustered standard errors are reported, derived from logistic regressions.

[4] All models were adjusted by age, age squared, logarithmic form of GDP, periods and country dummies. Significant associations and significant interactions ($p<0.05$) are highlighted in bold.

[5] P-value for the interactions between the tobacco control policy measures and education or occupation.

[6] In the stratified analysis based on occupation, some recent years for Finland were excluded because information on occupation was not available.

Table S2 Associations between price and non-price related tobacco control policy measures and amount of smoking[1] among men and women, stratified by education and occupation

	Men				Women			
	Cheapest cigarette price [2]		Non-price[3]		Cheapest cigarette price [2]		Non-price[3]	
	Coeff[4,5]	95% CI	Coeff	95% CI	Coeff	95% CI	Coeff	95% CI
Total population	-0.34	(-0.73 ~ 0.05)	-0.09	(-0.23 ~ 0.06)	-0.30	(-0.64 ~ 0.05)	-0.10	(-0.27 ~ 0.07)
Low education	-0.26	(-0.66 ~ 0.14)	-0.14	(-0.29 ~ 0.01)	-0.27	(-0.63 ~ 0.08)	-0.15	(-0.35 ~ 0.06)
High education	-0.36	(-0.83 ~ 0.10)	-0.10	(-0.20 ~ 0.01)	-0.30	(-0.71 ~ 0.10)	**-0.14**	**(-0.26 ~ -0.02)**
P value[6]	0.355		0.163		0.688		0.865	
Manual[7]	-0.37	(-0.75 ~ 0.02)	**-0.16**	**(-0.28 ~ -0.04)**	**-0.35**	**(-0.64 ~ -0.05)**	**-0.25**	**(-0.37 ~ -0.13)**
Non-manual	-0.48	(-1.09 ~ 0.13)	-0.11	(-0.25 ~ 0.04)	-0.21	(-0.60 ~ 0.18)	-0.05	(-0.22 ~ 0.13)
P value[6]	0.320		0.084		0.202		**0.013**	

[1] The number of cigarettes were log-transformed.

[2] The relative income price of cigarettes (the percentage of per capita GDP required to purchase the 100 cheapest packs of cigarettes).

[3] Summary scores for four domains of non-price related policies (smoking bans or restrictions, comprehensive bans on advertising and promotion, health warning labels and cessation services; ranging from 0 to 100).

[4] Coefficients and 95% confidence intervals based on clustered standard errors are reported, derived from linear regressions.

[5] All models were adjusted by age, age squared, logarithmic form of GDP, periods and country dummies. Significant associations and significant interactions ($p<0.05$) are highlighted in bold.

[6] P-value for the interactions between the tobacco control policy measures and education or occupation.

[7] In the stratified analysis based on occupation, some recent years for Finland were excluded because information on occupation was not available.

Chapter 7

Table S3 Associations between the tobacco control policy measures and smoking status among men and women, stratified by education and occupation, adding the lagged terms of non-price policy indicator

	Men					
	Non-price [1]		Non-price_1d[1]		Non-price_2d[1]	
	OR[2,3]	95% CI	OR	95% CI	OR	95% CI
Total population	**0.93**	**(0.89 ~ 0.96)**	1.04	(0.99 ~ 1.10)	1.00	(0.93 ~ 1.08)
Low education	**0.95**	**(0.90 ~ 0.99)**	0.99	(0.92 ~ 1.06)	1.06	(0.95 ~ 1.18)
High education	0.95	(0.89 ~ 1.02)	1.02	(0.98 ~ 1.05)	1.03	(0.97 ~ 1.09)
P value[4]	0.874		0.474		0.692	
Manual[5]	0.94	(0.88 ~ 1.00)	1.04	(0.94 ~ 1.15)	1.03	(0.91 ~ 1.17)
Non-manual	0.96	(0.90 ~ 1.02)	1.05	(0.93 ~ 1.17)	1.00	(0.89 ~ 1.13)
P value[4]	0.212		0.843		0.573	
	Women					
	Non-price [1]		Non-price_1d[1]		Non-price_2d[1]	
	OR[2,3]	95% CI	OR	95% CI	OR	95% CI
Total population	1.01	(0.94 ~ 1.09)	0.94	(0.83 ~ 1.06)	0.97	(0.85 ~ 1.11)
Low education	0.97	(0.90 ~ 1.05)	0.97	(0.90 ~ 1.04)	0.91	(0.82 ~ 1.01)
High education	0.97	(0.92 ~ 1.02)	1.09	(0.99 ~ 1.20)	0.95	(0.85 ~ 1.06)
P value[4]	0.941		0.149		0.667	
Manual [5]	1.00	(0.91 ~ 1.09)	0.95	(0.85 ~ 1.06)	1.00	(0.88 ~ 1.12)
Non-manual	0.97	(0.89 ~ 1.07)	1.08	(0.96 ~ 1.21)	**0.85**	**(0.76 ~ 0.95)**
P value[4]	0.537		0.182		0.075	

[1] Non-price_1d represents the 1-year lagged term of the non-price related policy measure. Non-price_2d represents the 2-year lagged term of the non-price related policy measure.
[2] Odds ratios and 95% confidence intervals based on clustered standard errors are reported, derived from logistic regressions. Odds ratios relate to an increase of 10 units in the scores.
[3] All models were adjusted by relative income price of cigarettes, age, age squared, logarithmic form of GDP, periods and country dummies. Significant associations and significant interactions ($p<0.05$) are highlighted in bold.
[4] P-value for the interactions between the tobacco control policy measures and education or occupation.
[5] In the stratified analysis based on occupation, some recent years for Finland were excluded because information on occupation was not available.

Table S4 Associations between price and non-price related tobacco control policy measures and smoking status, men and women combined, stratified by education and occupation

	Cheapest cigarette price [1]		Non-price[2]	
	OR[3,4]	95% CI	OR	95% CI
Total population	**0.71**	**(0.58 ~ 0.87)**	**0.95**	**(0.92 ~ 0.98)**
Low education	0.87	(0.66 ~ 1.14)	**0.95**	**(0.91 ~ 0.99)**
High education	1.07	(0.67 ~ 1.70)	0.99	(0.94 ~ 1.04)
P value[5]	0.291		0.31	
Manual[6]	0.88	(0.67 ~ 1.15)	1.01	(0.94 ~ 1.08)
Non-manual	0.86	(0.57 ~ 1.29)	0.96	(0.89 ~ 1.03)
P value[5]	0.87		0.30	

[1] The relative income price of cigarettes (the percentage of per capita GDP required to purchase the 100 cheapest packs of cigarettes).
[2] Summary scores for four domains of non-price related policies (smoking bans or restrictions, comprehensive bans on advertising and promoting, health warning labels and cessation services; ranging from 0 to 100). The odds ratios for the non-price related policies relate to an increase of 10 units in the scores.
[3] Odds ratios and 95% confidence intervals based on clustered standard errors are reported, derived from logistic regressions.
[4] All models were adjusted by age, age squared, logarithmic form of GDP, periods and country dummies. Significant associations and significant interactions ($p<0.05$) are highlighted in bold.
[5] P-value for the interactions between the tobacco control policy measures and education or occupation.
[6] In the stratified analysis based on occupation, some recent years for Finland were excluded because information on occupation was not available.

Table S5 Associations between a constructed measure combining the cheapest cigarette price and non-price related policy measures and smoking status among men and women, stratified by education and occupation

	Men		Women	
	price and non-price combined[1]		price and non-price combined[1]	
	OR[2,3]	95% CI	OR	95% CI
Total population	**0.99**	**(0.98 ~ 0.9995)**	0.99	(0.97 ~ 1.00)
Low education	1.00	(0.99 ~ 1.00)	**0.98**	**(0.97 ~ 0.997)**
High education	1.00	(0.98 ~ 1.01)	1.00	(0.99 ~ 1.01)
P value[4]	0.404		0.132	
Manual[5]	1.00	(0.99 ~ 1.01)	**0.99**	**(0.98 ~ 0.995)**
Non-manual	1.00	(0.99 ~ 1.01)	0.99	(0.97 ~ 1.00)
P value[4]	0.350		0.845	

[1] Following the same weights used for TCS, we allocated 30 points to the highest price measure (Great Britain in 2000) and scaled the price measure in other years and countries correspondingly. We also rebased the non-price related policy measure to a maximum score of 55. The price and non-price combined was the sum of the rebased price measure and the rebased non-price measure.

[2] Odds ratios and 95% confidence intervals based on clustered standard errors are reported, derived from logistic regressions.

[3] All models were adjusted by age, age squared, logarithmic form of GDP, periods and country dummies. Significant associations and significant interactions ($p<0.05$) are highlighted in bold.

[4] P-value for the interactions between the tobacco control policy measures and education or occupation.

[5] In the stratified analysis based on occupation, some recent years for Finland were excluded because information on occupation was not available.

Table S6 Associations between each of the four domains of the non-price policy score and smoking status among men and women, stratified by education and occupation

	non-price_ smokefree[1]		non-price_ advertising[1]		non-price_ warning[1]		non-price_ cessation[1]	
Men								
	OR[2,3]	95% CI	OR	95% CI	OR	95% CI	OR	95% CI
Total population	**0.95**	**(0.91 ~ 0.995)**	0.99	(0.89 ~ 1.09)	0.97	(0.84 ~ 1.13)	1.05	(0.87 ~ 1.28)
Low education	0.94	(0.88 ~ 1.00)	1.05	(0.93 ~ 1.18)	0.97	(0.80 ~ 1.18)	1.05	(0.82 ~ 1.33)
High education	0.96	(0.89 ~ 1.04)	0.99	(0.90 ~ 1.08)	1.08	(0.83 ~ 1.40)	0.98	(0.80 ~ 1.19)
P value[4]	0.632		0.082		0.538		0.664	
Manual[5]	1.02	(0.95 ~ 1.09)	1.02	(0.91 ~ 1.15)	1.01	(0.80 ~ 1.26)	**1.28**	**(1.01 ~ 1.63)**
Non-manual	0.99	(0.92 ~ 1.06)	1.04	(0.92 ~ 1.18)	1.07	(0.82 ~ 1.40)	1.10	(0.85 ~ 1.42)
P value[4]	0.215		0.500		0.090		**0.024**	
Women								
	non-price_ smokefree[1]		non-price_ advertising[1]		non-price_ warning[1]		non-price_ cessation[1]	
	OR[2,3]	95% CI	OR	95% CI	OR	95% CI	OR	95% CI
Total population	0.95	(0.88 ~ 1.02)	0.95	(0.81 ~ 1.12)	0.80	(0.61 ~ 1.04)	0.79	(0.61 ~ 1.02)
Low education	0.92	(0.82 ~ 1.02)	0.92	(0.77 ~ 1.11)	0.78	(0.61 ~ 1.01)	0.78	(0.58 ~ 1.04)
High education	0.98	(0.93 ~ 1.04)	1.05	(0.87 ~ 1.26)	1.09	(0.86 ~ 1.39)	1.00	(0.77 ~ 1.32)
P value[4]	0.309		0.508		**0.041**		0.301	
Manual[5]	0.95	(0.86 ~ 1.04)	0.88	(0.78 ~ 1.00)	0.95	(0.77 ~ 1.19)	**0.76**	**(0.58 ~ 0.996)**
Non-manual	0.97	(0.89 ~ 1.07)	0.88	(0.70 ~ 1.12)	0.88	(0.68 ~ 1.12)	1.04	(0.83 ~ 1.30)
P value[4]	0.611		0.982		0.490		**0.021**	

[1] non-price_smokefree, non-price_advertising, non-price_warning and non-price_cessation represent the four sub-score of the non-price policy measure (smoking bans or restrictions (maximum score=40), comprehensive bans on advertising and promotion (maximum score=23.64), health warning labels (maximum score=18.18) and cessation services (maximum score=18.18)).

[2] Odds ratios and 95% confidence intervals based on clustered standard errors are reported, derived from logistic regressions. Odds ratios relate to an increase of 10 units in the scores.

[3] All models were adjusted by the cheapest cigarette price, age, age squared, logarithmic form of GDP, periods and country dummies. Significant associations and significant interactions (p<0.05) are highlighted in bold.

[4] P-value for the interactions between the tobacco control policy measures and education or occupation.

[5] In the stratified analysis based on occupation, some recent years for Finland were excluded because information on occupation was not available.

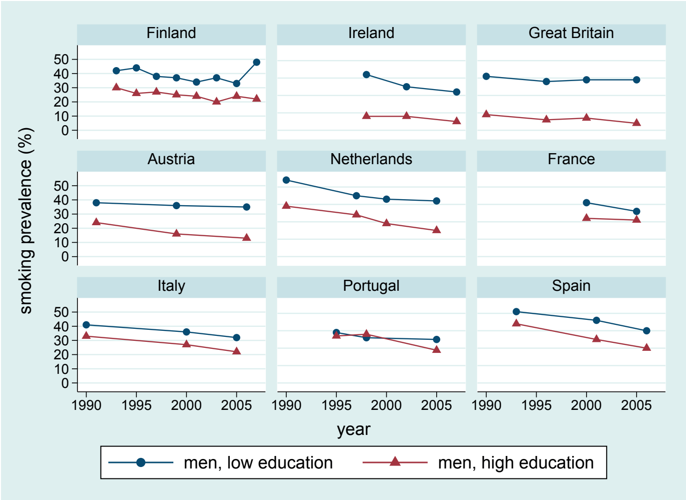


Figure S1.a Trends in age-standardised smoking prevalence by education among men in each country

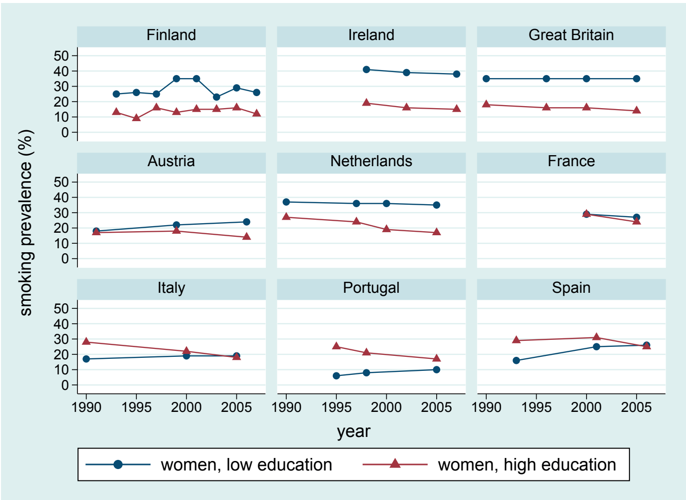


Figure S1.b Trends in age-standardised smoking prevalence by education among women in each country

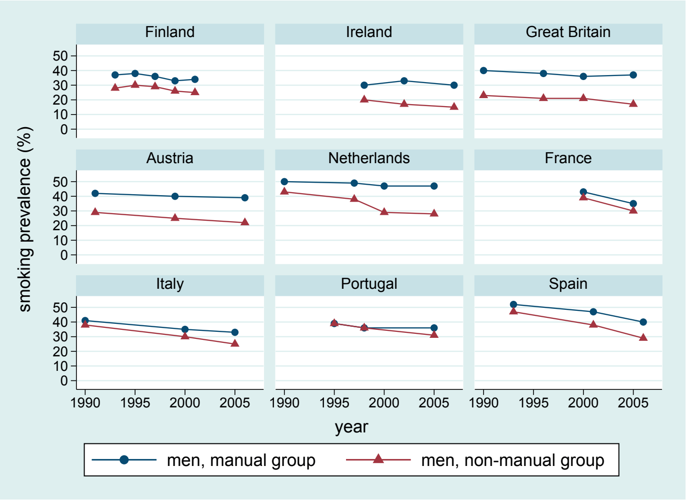


Figure S2.a Trends in age-standardised smoking prevalence by occupation among men in each country

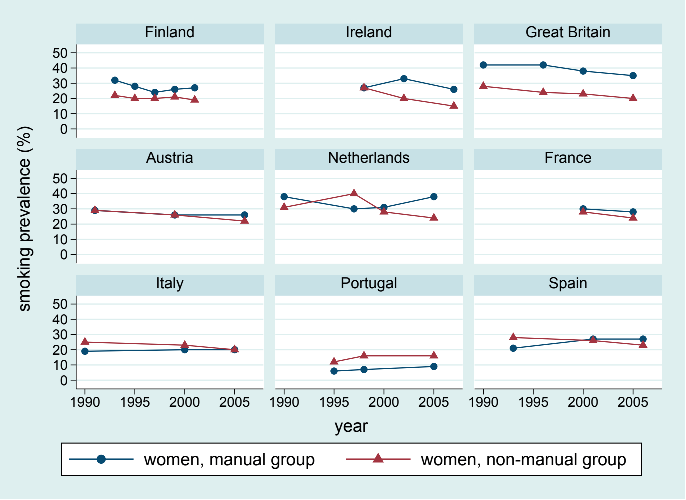


Figure S2.b Trends in age-standardised smoking prevalence by occupation among women in each country

Chapter 7

Chapter 8

General discussion

This thesis presents a collection of studies focusing on social and political determinants of population health and health inequalities in Europe. This final chapter summarizes and discusses the main findings. Moreover, the findings are presented in light of several methodological limitations and interpreted in reference to previous studies. Finally, we consider the policy implications and directions for future research.

8.1 MAIN FINDINGS

Answers to the study questions

1) Are international variations in population health related to social and political characteristics of European countries, particularly to levels of democracy and income inequality?

Chapter 2 and 3 explored two specific social and political country characteristics which are frequently studied but controversial determinants of population health, and their associations with variations in population health in Europe. Both studies were analysed using country fixed effects models, which eliminated time invariant country-level confounders.

In chapter 2, we assessed whether democratization was associated with improvements in population health as measured by life expectancy and cause-specific mortality rates, using data from 29 European countries in the period 1960–1990 and data from 43 European countries in the period 1987–2008. In the 1960–1990 period, current democracy (i.e. the democracy index measured in the same year as population health, which captured short-term effect of democracy) was more strongly associated with higher life expectancy than cumulative democracy (i.e. the weighted cumulative democracy index since an earlier year, which captured long-term and cumulative effect of democracy). The positive effects of current democracy on life expectancy were mainly the result from lower mortality from all heart diseases, pneumonia, liver cirrhosis, and suicide. In the 1987–2008 period, however, current democracy was associated with lower, and cumulative democracy with higher life expectancy, particularly among men. The positive effects of cumulative democracy on life expectancy were mainly the result from lower mortality from circulatory diseases, breast cancer, and external causes. Current democracy was associated with higher mortality from motor vehicle accidents in both periods, and also with higher mortality from cancer and all external causes in the second. Our results suggest that during these two periods democratization has overall had positive effects on life expectancy in Europe. That short-term changes in levels of democracy had positive effects in the first but not in the second period is probably due to the fact that democratization in Central and Eastern Europe in the second period was part of a complete system change which caused major societal disruptions.

In chapter 3, we assessed whether changes in income inequality were associated with improvements in population health as measured by life expectancy and cause-specific mortality rates in 43 European countries between 1987 and 2008. Most statistically significant associations between income inequality and many mortality indicators found in pooled cross-sectional regressions became statistically insignificant once the country fixed effects were added. Exceptions were deaths from external causes and homicide among men, and cancers among women. However, the statistically significant associations for homicide and cancers disappeared after further adjustment for indicators of democracy, education, transition to national independence, armed conflicts, and economic freedom. Our study thus suggests that cross-sectional associations between income inequality and mortality reflect the confounding effects of other country characteristics. In a European context, therefore, national levels of income inequality do not appear to have an independent effect on mortality.

2) Are international variations in health inequalities related to policies implemented in European countries, particularly to national policies to tackle socioeconomic inequalities in health and tobacco control policies?

Chapter 4 described trends in socioeconomic inequalities in self-assessed health in Europe. Chapter 5 reviewed the analytical methods for policy evaluation and explored how these methods can be used to evaluate policy effects on health inequalities. Chapters 6 and 7 empirically evaluated the effectiveness of the ambitious strategy to tackle health inequalities which was implemented in England and the impact of tobacco control policies on smoking inequalities, respectively.

In chapter 4, we provided a comprehensive overview of trends in socioeconomic inequalities in self-assessed health among men and women aged 30–79 years in seventeen European countries between 1990 and 2010. Declining trends in prevalence of less-than-good self-assessed health were observed in many countries, particularly in Southern and Eastern Europe and the Baltic States. In all countries, less-than-good self-assessed health was more prevalent in lower educational and occupational groups. Absolute inequalities in self-assessed health were mostly constant, whereas relative inequalities mostly increased. Almost no country consistently experienced a statistically significant decline in either absolute or relative inequalities. Our results do not support the idea that countries with national policies to tackle health inequalities have fared better, in terms of inequalities in self-assessed health, than countries without such policies.

In chapter 5, we identified seven econometric methods for the evaluation of the impacts of natural policy experiments: regression adjustment, propensity score matching, difference-in-differences, fixed effect analysis, instrumental variable, regression discontinuity and interrupted time-series. Using a fictitious numerical example, we explored how to apply these methods to assess policy effects on the magnitude of health inequalities. We found that all methods can be used to assess the impact of policies on health inequalities, either by doing a stratified analysis or by including interaction terms (the latter, however, was impossible with propensity score matching). The choice of the method depends on data availability and the nature of the confounders in the analysis (whether observable or not, and whether time-variant or not). However, the low external validity of results from instrumental variable and regression discontinuity makes these methods less desirable for assessing policy effects on population-level health inequalities.

In chapter 6, we empirically evaluated the effectiveness of an ambitious programme pursued by the English government to reduce health inequalities, by comparing trends in health inequalities in England both over time and between countries. After the implementation of the English strategy, more favourable trends in some health indicators were observed among low-educated people, but trends in health inequalities in 2000–2010 in England were not more favourable than those observed in the period 1990–2000. For most health indicators, changes in trends of health inequalities after 2000 in England were also not statistically significantly different from those seen in the other countries. Our results therefore do not suggest a favourable effect of the English strategy.

In chapter 7, we investigated the impact of price- and non-price-related population-wide tobacco control policies on smoking by socioeconomic group in nine European countries between 1990 and 2007. The price of popular cigarettes and non-price policies were negatively associated with smoking among men. The price of the cheapest cigarettes was negatively associated with smoking among women. While these favourable effects were generally in the same direction for all socioeconomic groups, they were larger and statistically significant in lower socioeconomic groups only. Our results suggest that tobacco control policies as implemented in these European countries have probably helped to reduce the prevalence of smoking in the total population, and particularly in lower socioeconomic groups, thereby contributing to a reduction of inequalities in smoking.

8.2 METHODOLOGICAL LIMITATIONS

Data considerations issues

The first part of this thesis (chapters 2 and 3) used period life expectancy and cause-specific mortality as health measures. Although both are common and easily interpretable measures of population health, some possible validity issues need to be noted. One potential limitation of using period life expectancy, which is typically calculated from age-specific death rates by life table methods, is that tempo effects can bias the measurement of life expectancy in times of rapidly declining or increasing mortality [1, 2], and that at lower rates of mortality larger declines are necessary for one unit increase in life expectancy [3]. The potential bias due to this issue in our analysis is probably limited, as the results from analyses using age-standardized mortality from all causes are broadly similar to those from analyses using life expectancy. International differences in processes of classifying and registering causes of death may cause problems in comparison between countries when cause-specific mortality rates are used [4, 5]. If these between-country differences can be regarded as time-invariant during the study period, they are unlikely to distort our findings as these country heterogeneities have been adjusted by using country fixed effects models. Yet, as more guidance for classifying and registering causes of death becomes available over time in many European countries, the validity of cause-specific mortality measures may improve over time which then would bias our results. For example, if the accuracy of the cause-of-death classification improves during periods of increased democratization, with the result that more deaths are correctly recorded under a certain classification during this process, the favourable effect of democratization on that specific mortality trend will be underestimated.

The second part of this thesis is based on health survey data compiled in a large European project, i.e. DEMETRIQ (Developing Methodologies to Reduce Inequalities in the Determinants of Health). In this project we have collected and harmonised data from nationally representative health interview surveys for 21 European countries over a period of between 2 and 4 decades. The data are used in chapters 4, 6 and 7 to give an overview of trends in inequalities in self-assessed health in Europe or to evaluate the effects of policies on inequalities in various self-reported health measures including smoking.

As data for different countries came from different surveys, the comparability between countries may be compromised due to differences in data collection across surveys, e.g. sampling strategy, population coverage and framing of survey questions. Despite great harmonization efforts, we were not able to remove all differences across surveys. However, as we either focused on general trends in health inequalities within countries (chapter 4), or constructed regression models in such a way that the models explored within-country variations in health over time by socioeconomic group and then compared the within-country variations between countries (chapter 6 and 7), we retained comparability in data over time within each country, and we therefore consider the risk of bias due to between-country variations in data collection to be limited.

Another possible concern of the health survey data is that we mainly used self-reported measures of health. The tendency to report health problems may differ between countries due to different cultural backgrounds [6, 7], e.g. persons in Central and Eastern European countries may tend to report their health as less good than persons in other European countries [8, 9], and Italians appear to generally assess their health more positively than Finns even after adjusting for social and health characteristics [10]. However, as the tendency of reporting health problems within each country does not likely change over the study period, and as we mainly explored the variations over time within countries (as indicated above), the comparability problem of using self-reported health measures across countries is considered to be limited.

Similarly, some studies suggest that the tendency to report health problems may also be different between socioeconomic groups [11, 12], although the differences are sometimes found to be very small and perhaps only limited to the extreme educational groups in men [12–15]. Although it is almost unavoidable to use self-reported data in trend studies, it would be preferable to estimate the magnitude of health inequalities from objectively measured data. However, as long as the potential differences in reporting health outcomes between socioeconomic groups are stable during the study period, the bias is likely to be limited as we mainly focused on changes in health inequalities within countries.

Due to lack of appropriate data across the entire life course, this thesis focused on inequalities in health among adults (aged 30–79 in chapters 4 and 7, aged 16–79 in chapter 6). Younger respondents were excluded because many of them were still receiving full-time education. Older respondents were excluded to avoid the potential bias caused by the exclusion of the institutionalized population in many surveys. It cannot be excluded that analyses focusing on adolescents or old people would lead to different conclusions. Specifically, in chapter 6, focusing on adults means that we may have missed the potential effects of policies implemented in the English strategy that mainly targeted children or old people, e.g. in the Sure Start Local Programmes, Healthy Schools Programme, the implementation of adjusted pension policies and long-term care for the elderly [16, 17]. For example, one study found some effects of the Sure Start Local Programmes in improving children's social development, social behaviours and independence, which may lead to better educational performance, lower unemployment and potentially a narrowing of health inequalities in the long run [18]. However, not all studies have discovered favourable effects of the Sure Start Local Programmes and their effectiveness is still under discussion [19, 20]. To the best of our knowledge, there are currently no studies specifically focusing on the effectiveness of the English strategy among older people. Because we focused on adults in chapter 7, we mainly evaluated the potential impact of tobacco control policies on quitting smoking instead of on smoking initiation. This is because most smokers already started smoking before the minimum age we looked at (i.e. 30 years old). Initiating and particularly re-starting smoking during the age period 30–79 years is possible and can be influenced by tobacco control policies, and this deserves further attention.

Causal inference

This thesis explored potential social and political determinants of population health and health inequalities using observational studies that compared performances across countries. The main methodological challenge of this approach is establishing causality. Cross-country comparative studies never provide evidence as strong as the evidence provided by randomized controlled trials. However, the limitations of randomized controlled trials in public health have also been clearly recognized [21, 22]. These limitations are particularly obvious with regard to the analysis of the effects of social and political factors on health and health inequalities. For example, randomization to democracy or low levels of income inequalities is simply impossible. Moreover, policies reducing health inequalities are usually implemented in large geographic areas, such as a whole country as in the case of the national strategy in England, and/or need a long time to show effects (e.g. tobacco control policies); in these circumstances randomized controlled trials are often neither appropriate or feasible. Cross-country comparative studies can be seen as a suitable alternative study design, especially if potential confounders can adequately be controlled for.

In most studies of this thesis, we controlled for several observable country characteristics potentially confounding the associations between the social or political factors that we were interested in and the health outcomes. Moreover, we additionally used country fixed effects models to control for unobservable time-invariant country heterogeneities due to long-term country-specific cultural, geographical or political factors (chapter 2, 3 and 7). In chapter 6, we included some comparison countries, which may have shared similar trends in health or health inequalities with the country studied. All these measures to control for confounding make it likely that the relation we report is closer to a causal relationship as compared to many previous studies.

Having said that, we were not able to exclude the potential bias caused by omitted time-variant confounders. For example, in chapter 2, one potentially omitted time-variant confounder was the change in value orientations preceding democratization in Southern Europe and Central and Eastern Europe [23]. These may have changed other collective and individual behaviours, independent from and parallel to the changes in the political system. Another example is the increase in knowledge about the health consequences of smoking (chapter 7), which may have changed smoking behaviour in the population and may also have increased the likelihood of implementing tobacco control policies. Unfortunately, many of these potential confounders are not available in current international datasets, which makes controlling for them difficult.

Generalizability

This thesis only contains studies of European countries. This reduced the potential for confounding because country-level confounding variables can be expected to be more similar within Europe than in a global context. It also produced results that are relevant for policy makers in European countries, who are likely to be more concerned with the effects of democracy or income inequality or tobacco control policies as observed within the range of variation prevailing in a European context, and with policies which have already been implemented in a neighbour country and can therefore be replicated more easily. However, this restriction to a European context limits generalizability of our results to a global context.

For example, democratization processes have also occurred in many Latin American and Asian countries during the study period [24]. Democratization in Latin America and Asia involved countries in a different stage of economic development, and the effects of democratization on cause-specific mortality may therefore have been different from those observed in Europe. However, the overall conclusion about the effect of democracy on population health may still be the same, as suggested by many studies that have found positive effects of levels of democracy in other parts of the world [25–28]. Similarly, extensive tobacco control policies have also been implemented in countries outside of Europe, especially in the United States, Canada and Australia [29]. Again, broadly similar findings about the equity impact of tobacco control policies are found in the studies focusing on countries outside of Europe [29].

Perhaps the most clear example of limits to generalizability is our study of income inequality. It has been noted that the most consistent evidence for an adverse effect of income inequality on population health comes from within-country differences in the United States or other countries with comparable or even larger income inequalities as in the US [30, 31]. Studies including non-European countries with a more equal income distribution than the US, such as Canada, Australia and Japan, also more often produced statistically insignificant findings [32, 33]. Therefore, one possible explanation for not finding an adverse effect of income inequality in this thesis is that we focused on a set of European countries with more equal income distributions (of which the mean GINI coefficient in 2008 is 0.3) than the United States (of which the observed GINI coefficient in 2008 is 0.38). It cannot be excluded that including countries from other regions which have higher level of income inequalities may make the adverse effect of income inequality more detectable.

8.3 INTERPRETATION OF FINDINGS

Social and political determinants of average population health

This thesis suggests a favourable effect of democracy on improving population health, either in the short or longer term. This is consistent with some previous studies which also found higher life expectancies in countries with more democratic governments [26–28, 34, 35]. However, many previous studies used less rigorous analytic methods [36], and more than half of the studies of democracy versus life expectancy or mortality were cross-sectional studies. As we wrote in chapter 2, there was also a strong cross-sectional relation between democracy and life expectancy in Europe among both men and women, even after controlling for GDP and education in our dataset. These results, however, were likely to be confounded by unobserved country characteristics. In our country-fixed effects models, we essentially related changes in democracy levels within each country to changes in health outcomes. After removing effects of unobserved country characteristics, we still found a positive association between indices of democracy and life expectancy.

Despite the fact that the relation between democracy and population health has been studied relatively frequently, the explanation of the study findings has generally remained speculative. The prevailing explanation is that, as democratic governments are more responsive to the needs of their populations including those that are health related, they generally have a higher social spending on public goods including or even especially those related to health promotion and protection, such as education and social security, than autocratic governments [27, 37–40]. As we discussed in chapter 2, among the specific countries and periods we were looking at, the association between democratization and life expectancy was likely to reflect the wider policy changes that were promoted by democratization: health care reform, economic restructuring, road traffic safety programs, health promotion campaigns, etc. [41]. In the case of Spain, Portugal and Greece in the 1970s and 1980s, changes in health care systems are likely to have been partial mediators [42]. The modernization of their health care systems during the study period led to a great improvement in access to new health care technologies, which enabled progress in medical outcomes after decades of stagnation [43] and which partly explained the observed declines in mortality from conditions amenable to medical intervention [44, 45].

Although we observed cross-sectional relationships between income inequality and mortality indicators, this thesis does not suggest an independent effect of income inequality on mortality, as once the country fixed effects were added, almost all associations between income inequality and mortality indicators became statistically insignificant. Differences between results from pooled cross-sectional analyses and fixed effects models indicate that the observed association between income inequality and mortality is likely to result from confounding by some country characteristics that are closely related to both the hierarchical nature of societies, as indicated by income inequality, and the health of their populations, and that income inequality as such is not a driving force of population health. As we wrote in chapter 3, one underlying factor that might be responsible for the disappearance of the effect could be social and health policies that vary across countries and are persistent over time, such as poverty reduction policies [46]. Other responsible factors could be some cultural and historical elements of a country, e.g. egalitarianism (importance of transcending self-interest and promoting the welfare of others), power distance (extent to which the less powerful accept that power is distributed unequally) and ethnic heterogeneity, which are potentially important determinants of population health [47, 48], and at the same time could be related to income inequality [49, 50].

After a careful examination of the existing studies, we found that the evidence-base to argue that tackling income inequality is an effective strategy to improve population health is not strong. As early as 2002, there was a BMJ editorial suggesting that evidence for a correlation between income inequality and population health was slowly dissipating due to the availability of better quality and interpretation of data, especially in the countries outside the United States, and within the United States it had still to be convincingly demonstrated that it was not due to curvilinear individual level relationships and confounding [51]. Going through the papers published in the recent 2 years, we still found large discrepancies between the results of different studies [52–56]. The latest review [57] tried to establish a causal interpretation between income inequality and health based on the existing findings. Available empirical studies were used to find evidence for Bradford Hill's criteria of causality, but only few studies using fixed effects analysis which were able to reduce confounding were included in this review. Moreover, the review attributed the statistically insignificant findings of previous studies to some methodological problems, including income inequality being measured at an inappropriate scale (e.g. regional instead of national measures of income inequalities), the inclusion of mediating variables as controls, the use of subjective rather than objective measures of health, or follow up periods which were too short. However, there are a number of studies like our study which have solved the methodological problems mentioned in the review, but still have produced statistically insignificant findings [58, 59]. Thus, there is no strong evidence-base to claim that reducing national income inequality can improve population health.

In this thesis, we did not study the potential effect of income inequality on health inequality. Although income inequality is not a causal determinant of average population health, could it be that income inequality determines the level of health inequality? The relation between income inequality and socioeconomic inequality in health has not been studied intensively. Some studies reported evidence for a positive relation between income inequality and health inequalities [52, 60–63], while others reported mixed or insignificant findings [64, 65]. Whether countries with higher income inequality have higher health inequalities and whether this relationship is causal still needs more formal tests in the future.

It is worthwhile to compare our findings about the health effect of democracy to those about the effects of income inequality. The claim that democracy promotes an egalitarian distribution of income can be traced back to the late 18th century [66]. Many scholars argued that democracy increases opportunities for participation, allows the poor to demand more equitable income redistribution, and helps the lower and middle classes by adopting redistributive policies such as welfare spending, progressive taxation, minimum wage laws and price subsidies [66–68]. With some exceptions [67, 69–71], several studies found evidence for a favourable effect of democracy in reducing income inequality or a curvilinear relationship between democracy and income inequality [66–69, 72–74]. However, since democracy may promote income equality, and income inequality does not have a causal impact on population health (as suggested in chapter 3), our findings suggest that income inequality is not a mediating factor between democracy and population health.

Social and political determinants of inequalities in health

Our findings about the impacts of health policies implemented in Europe on socioeconomic inequalities in health are generally not very positive. The observed trends in inequalities in self-assessed health were not consistent with the amount of efforts to reduce health inequalities made in some European countries. We could also not detect a favourable effect of the English strategy in tackling health inequalities, even though the implementation of the strategy was both historically and internationally unique. Some potential equity impacts were found with tobacco control policies as implemented in the available European countries, which appear to have helped to reduce the prevalence of smoking in the total population, particularly in the lower socioeconomic groups, and thereby may have contributed to a reduction of inequalities in smoking.

Specifically, countries with national policies to tackle health inequalities, such as England and Scotland [75], did not display more favourable trends in inequalities in self-assessed health than countries without such policies. One possible explanation is that self-assessed health must be understood as a summary statement about the way in which numerous aspects of health, both subjective and objective, are combined within the perceptual framework of the individual respondent [7]. It is a multidimensional measure, which is found to be mostly a measure of physical and mental health, with some limited contributions from the respondent's age, early life factors, family history, sociodemographic characteristics and health behaviours [76]. It is not a measure used for target setting in any national strategy tackling health inequalities, and it therefore might be better to include mode detailed outcomes in the evaluations of policy effects on reducing inequalities in health. However, the results from a previous study analysing trends in inequalities in mortality in the same period in Europe also suggest that the trends in inequalities in mortality did not correspond well with the amount of efforts to reduce health inequalities made in some European countries [77]. This may partly confirm the findings of this thesis.

After referring to the trends in inequalities before the implementation of the strategy in England and to the trends in three other European countries, we found that the effect of the English strategy in reducing inequalities in self-reported health, smoking and obesity is limited. Despite the fact that parts of the overall strategy seemed to have reduced inequalities in some outcomes in specific sections of the total population [18, 78, 79], the general consensus is that the population-level effects of the strategy as a whole have been largely absent. The new finding of our study is that the English strategy has not effectively slowed down the widening trends in inequalities in health in England, nor has it helped trends in health inequalities in England to widen less than those in countries without such a strategy. As we wrote in chapter 6, existing reviews have suggested many potential reasons for why the English strategy was not more successful [80–87]. One widely acknowledged reason is that the design of the English strategy was not based on policies with proven effectiveness in reducing health inequalities [83, 88–90]. This is partly due to the limited availability of evidence for the effectiveness of policies in this area at the time the strategy was developed. Another reason is that the English strategy might have chosen the wrong entry-points, which were irrelevant for life expectancy and infant mortality at least within the chosen time frame [84] or were only the proximal causes (such as smoking) but not the ultimate causes of inequalities in ill-health [80]. The inadequate delivery of the English strategy was also criticized [80, 83, 84]. The development of policies tackling population-level health inequalities in the future should draw lessons from the English strategy.

This thesis suggests that both price and non-price tobacco control efforts in the available European countries have helped to reduce smoking among the total population and among people in low socioeconomic groups. As the statistically significantly negative associations between policy measures and smoking were only found among the lower socioeconomic groups in the stratified analysis, the tobacco control policies might have helped to reduce smoking inequalities. Our results are consistent with many previous studies which also suggest that increasing the cigarette price may be more effective in reducing smoking among lower socioeconomic groups [29, 91]. A straightforward explanation is that people in higher socioeconomic groups may not regard the cost of cigarettes as a heavy financial burden, which would make them less sensitive to the change of cigarette price [92]. Non-price tobacco control policies may also have some effect in reducing inequalities in smoking, by reducing smoking among low-educated men. However, if we investigated the importance of the four sub-domains of the non-price policy measure separately, none of them consistently showed a differential impact by socioeconomic position on smoking (as shown in chapter 7). Therefore, the most effective domain of non-price tobacco policy is still uncertain, and perhaps a comprehensive package of non-price tobacco control policies as a whole which simultaneously covers several important domains such as smoking bans and restrictions and health warning labels, etc., will be most effective in reduce inequalities in smoking.

As we noted above, general findings of our studies into the effect of policies on health inequalities are not very positive. A potential question is whether there are other national strategies or policy areas which are not covered by this thesis but can be expected to be more effective in tackling health inequalities. As far as we know, the English strategy is the most comprehensive and promising national strategy that has ever been implemented in Europe, with a budget of more than £20 billion, enormous government efforts, and a remarkable commitment to critically review, revise and then re-review the strategy [17, 80–82, 93–96]. The national strategies for tackling health inequalities in Norway and Finland have had much less allocated resources and have never been directly evaluated so far [75, 97]. The national programme to tackle health inequalities proposed in 2001 in the Netherlands has never been implemented due to changes in the Dutch government [98, 99]. It seems that there is no national strategy which has been firmly demonstrated to be effective in reducing health inequalities for us to directly follow right now. Besides tobacco control policies, there may be some specific policies which do not focus on tackling health inequalities but may have a favourable equity impact on health. These include labour market policies, especially the policies helping to increase employment among disadvantaged people [100–103], poverty reduction policies such as minimum wage and unemployment insurance [104–107], and alcohol policies [108]. However, empirical studies directly testing the effectiveness of these policies in reducing health inequalities are still scarce. All in all, we are still a long way from having a rich arsenal of effective policies to reduce health inequalities.

8.4 IMPLICATIONS AND RECOMMENDATIONS FOR POLICY AND RESEACH

Policy implications and recommendations

We demonstrated that democratization has probably created favourable conditions for reducing mortality from diseases amenable to health intervention, if not in the short then in the longer term. The policy implications of these findings are significant. The way societies organise themselves through their political systems apparently has an important and independent impact on average population health. Policies that promote democracy, such as the establishment of government bodies which promote education, equality and democracy via international cooperation (e.g. the UK's Department for International Development), and the democratization policies of European Union towards its eastern partnership countries [109, 110], should be encouraged not only from the view of political justice, but also from the perspective of improving general health and well-being. Although complete system changes may cause major societal disruptions and deteriorations of population health in the short run, the cumulative level of democracy could still benefit population health in the longer run. We therefore need to allow democratization the time that is needed to realize important and persistent health benefits.

Income inequality may have some undesirable effects on average population outcomes in a range of areas. For example, a high level of income inequality may cause negative emotions to people living in the society such as shame, distrust and anxiety, and simultaneously these psychosocial problems may lead to antisocial behaviours, reduced civic participation, less social capital and cohesion within the community [111]. However, reducing income inequalities should not be regarded as an effective strategy to improve life expectancy and reduce mortality, at least in a European context.

We found persistent socioeconomic inequalities in self-assessed health over time in all European countries with available data. Almost no country consistently experienced a statistically significant decline in either absolute or relative inequalities. Similar findings were obtained in many previous studies using some other health measures [112–120]. There is a need for development of more effective policies, specifically for such policies which have a differentially greater impact on people with lower socioeconomic position, and which therefore have potential to reduce inequalities. A potential policy area to be considered is tobacco control, especially the price-related tobacco control policies, the effect of which on reducing inequalities in smoking has been consistently found in this thesis and some previous studies [29, 91], and which might help to reduce inequalities in some major chronic diseases and premature mortality in the long run [121–123].

Although the English strategy contained a number of comprehensive and coordinated policies, and had a remarkably high level of government commitment, we showed that its effect in reducing socioeconomic inequalities in health over the whole population was limited. Apparently, reducing socioeconomic inequalities in health is much more difficult than most scientists had foreseen [83]. The potential reasons for why the strategy was not more successful have been discussed in some reviews [80–87], which can provide some lessons for future policy design. In order to effectively reduce national health inequalities, policy actions should be directed towards the entry-points which are important determinants of health inequalities and give priority to the policies with proven effectiveness in reducing health inequalities [83, 88–90]. Moreover, policy actions should be carried out at a scale that is sufficient to have a population-wide impact, which requires a larger and quicker resource-allocation and effective local delivery systems [80, 83].

Research implications and recommendations

We analysed the validity of democracy and income inequality as two causal determinants of average on population health. There may be many other time-variant country characteristics for which the effects on population health can be measured with a similar fixed effects approach, such as investment in health care system, welfare state characteristics and social cohesion. Moreover, the effects of democracy and income inequality on socioeconomic inequalities are still unknown. Further research may need to explore whether democracy and income inequality causally determine the level of health inequality

In our analyses of the impacts of health policies implemented in Europe on inequalities in health, we mainly used data from adults (e.g. people who are 30 years or older). Future research may consider to explore the potential policy effects on inequalities in health among adolescents, if appropriate information such as the socioeconomic positions of the respondents' parents are available. Specifically, policies which aimed at preventing initiation of unhealthy behaviours (e.g. preventing smoking initiation) rather than changing unhealthy behaviours (e.g. smoking cessation) may be especially relevant for adolescents.

In this thesis, we did not find a favourable effect of the English strategy using general health measures (e.g. self-assessed health, long-standing illness and obesity), which may be relatively loosely related to the policies in the English strategy. In contrast, using smoking status which should be the direct outcome of tobacco control policies, we found some positive effects in reducing smoking inequalities. Therefore, we recommend future research to choose outcomes that are directly related to the policies in which we are interested in order to find a detectable impact on inequalities in health.

It is clear that the current evidence for the effectiveness of policies in reducing socioeconomic inequalities in health is limited. Discovering successful policies tackling health inequalities would be extremely helpful for the policy design in the future. We recommend that equity impact analyses should be an integral part of any policy experiment, and researchers need to do more to make the evidence base for reducing health inequalities available and accessible to policymakers.

We demonstrated that the identified methods which have been used to evaluate the effect of natural policy experiments can be useful tools to evaluate policy effects on health and health inequalities, and would help to build a good evidence base for policy makers and stakeholders. We recommend future research to consider to use these methods with the aim to formally evaluate the causal impact of social and political determinants of health and health inequalities. As shown in this thesis, fixed effects model and difference-in-differences analysis can be very useful and improve causal interpretations. Regression adjustment, propensity score matching and interrupted time-series also deserve a consideration, due to their relatively lower data requirements. If appropriate settings are available, instrumental variable and regression discontinuity analysis can be good options, as the results from both methods (if appropriate) have very high internal validity. However, their usage in evaluating the policies impacts on health inequalities over the whole population might be limited, as both instrumental variable and regression discontinuity analysis suffer from low external validity (i.e. the estimated policy effects are the effects on a proportion of people and the generalisation of the conclusion to the whole population is uncertain). Given the different data requirements and underlying assumptions these methods require, a sense of evaluation from the start of policy design and research is extremely helpful, which would facilitate good quality evaluations by providing systematically collected data during a necessarily long follow-up period, choosing well-defined exposed populations and predetermined control groups, and creating suitable settings for regression discontinuity and instrumental variable analyses. It helps to allow full play to the potential quantitative evaluation approaches and improve the availability of evidence on social and political determinants of health and health inequalities.

REFERENCES

1. Bongaarts, J. and G. Feeney, *How long do we live?* Population and Development Review, 2002. **28**(1): p. 13-+.
2. Luy, M., *Mortality tempo-adjustment: An empirical application*. Demographic Research, 2006. **15**: p. 561–590.
3. Preston, S., P. Heuveline, and M. Guillot, *Demography: measuring and modeling population processes*. 2001, Oxford: Blackwell.
4. Reid, D.D. and G.A. Rose, *Assessing the comparability of mortality statistics*. Br Med J, 1964. **2**(5422): p. 1437–9.
5. Inserm (French Institute of Health and Medical Research), *Comparability and quality improvement of European causes of death statistics*. 2001: http://ec.europa.eu/health/ph_projects/1998/monitoring/fp_monitoring_1998_frep_04_en.pdf (accessed 19 February 2016).
6. Jurges, H., *True health vs response styles: Exploring cross-country differences in self-reported health*. Health Economics, 2007. **16**(2): p. 163–178.
7. Jylha, M., *What is self-rated health and why does it predict mortality? Towards a unified conceptual model*. Soc Sci Med, 2009. **69**(3): p. 307–16.
8. Babones, S.J., *The consistency of self-rated health in comparative perspective*. Public Health, 2009. **123**(2): p. 199–201.
9. Reile, R., et al., *The recent economic recession and self-rated health in Estonia, Lithuania and Finland: a comparative cross-sectional study in 2004–2010*. J Epidemiol Community Health, 2014. **68**(11): p. 1072–9.
10. Jylha, M., et al., *Is self-rated health comparable across cultures and genders?* J Gerontol B Psychol Sci Soc Sci, 1998. **53**(3): p. S144–52.
11. Dowd, J.B. and A. Zajacova, *Does the predictive power of self-rated health for subsequent mortality risk vary by socioeconomic status in the US?* Int J Epidemiol, 2007. **36**(6): p. 1214–21.
12. Huisman, M., F. van Lenthe, and J. Mackenbach, *The predictive ability of self-assessed health for mortality in different educational groups*. Int J Epidemiol, 2007. **36**(6): p. 1207–13.
13. van Doorslaer, E. and U.G. Gerdtham, *Does inequality in self-assessed health predict inequality in survival by income? Evidence from Swedish data*. Soc Sci Med, 2003. **57**(9): p. 1621–9.
14. Wannamethee, G. and A.G. Shaper, *Self-assessment of health status and mortality in middle-aged British men*. Int J Epidemiol, 1991. **20**(1): p. 239–45.
15. Burstrom, B. and P. Fredlund, *Self rated health: Is it as good a predictor of subsequent mortality among adults in lower as well as in higher social classes?* J Epidemiol Community Health, 2001. **55**(11): p. 836–40.
16. Glass, N., *Sure Start: the development of an early intervention programme for young children in the United Kingdom*. Children & Society, 1999. **13**: p. 257–264.
17. Department of Health, *Reducing health inequalities: An action report*. 1999: London. http://www.swslim.org.uk/downloads/SL1264.pdf (accessed 02 February 2015).
18. Melhuish, E., et al., *Effects of fully-established Sure Start Local Programmes on 3-year-old children and their families living in England: a quasi-experimental observational study*. Lancet, 2008. **372**(9650): p. 1641–7.
19. Belsky, J., et al., *Effects of Sure Start local programmes on children and families: early findings from a quasi-experimental, cross sectional study*. BMJ, 2006. **332**(7556): p. 1476.
20. Rutter, M., *Is Sure Start an effective preventive intervention?* Child and Adolescent Mental Health, 2006. **11**(No.3): p. 135–141.

21. Black, N., *Why we need observational studies to evaluate the effectiveness of health care.* BMJ, 1996. **312**(7040): p. 1215–8.
22. Bonell, C.P., et al., *Alternatives to randomisation in the evaluation of public health interventions: design challenges and solutions.* J Epidemiol Community Health, 2011. **65**(7): p. 582–7.
23. Inglehart, R. and C. Welzel, *Modernization, cultural change, and democracy. The human development sequence*. 2005, Cambridge: Cambridge University Press.
24. Huntington, S.P., *The third wave: Democratization in the late 20th century*. 1991, Norman: University of Oklahoma Press.
25. J.W., M., *Wealth, Health, and Democracy in East Asia and Latin America*. 2010, Cambridge: Cambridge University Press.
26. Besley, T. and M. Kudamatsu, *Health and democracy.* American Economic Review, 2006. **96**(2): p. 313–318.
27. Baum, M.A. and D.A. Lake, *The political economy of growth: Democracy and human capital.* American Journal of Political Science, 2003. **47**(2): p. 333–347.
28. Franco, A., C. Alvarez-Dardet, and M.T. Ruiz, *Effect of democracy on health: ecological study.* British Medical Journal, 2004. **329**(7480): p. 1421–1423.
29. Brown, T., S. Platt, and A. Amos, *Equity impact of population-level interventions and policies to reduce smoking in adults: a systematic review.* Drug Alcohol Depend, 2014. **138**: p. 7–16.
30. Lynch, J., et al., *Is income inequality a determinant of population health? Part 1. A systematic review.* Milbank Quarterly, 2004. **82**(1): p. 5–99.
31. Kawachi, I. and S.V. Subramanian, *Income inequality*. In Berkman, L. F., Kawachi, I., Glymour, M. M. (Eds.), Social Epidemiology (pp. 126–151). 2014, United States of America: Oxford University Press.
32. Ross, N.A., et al., *Relation between income inequality and mortality in Canada and in the United States: cross sectional assessment using census data and vital statistics.* BMJ, 2000. **320**(7239): p. 898–902.
33. Ross, N.A., et al., *Metropolitan income inequality and working-age mortality: a cross-sectional analysis using comparable data from five countries.* J Urban Health, 2005. **82**(1): p. 101–10.
34. Alvarez-Dardet, C. and A. Franco-Giraldo, *Democratisation and health after the fall of the Wall.* J Epidemiol Community Health, 2006. **60**(8): p. 669–71.
35. Muntaner, C., et al., *Politics, welfare regimes, and population health: controversies and evidence.* Sociol Health Illn, 2011. **33**(6): p. 946–64.
36. Klomp, J. and J. de Haan, *Is the political system really related to health?* Soc Sci Med, 2009. **69**(1): p. 36–46.
37. Avelino, G., D.S. Brown, and W. Hunter, *The effects of capital mobility, trade openness, and democracy on social spending in Latin America, 1980–1999.* American Journal of Political Science, 2005. **49**(3): p. 625–641.
38. Navia, P. and T.D. Zweifel, *Democracy, dictatorship, and infant mortality revisited.* Journal of Democracy, 2003. **14**(3): p. 90–103.
39. Wigley, S. and A. Akkoyunlu-Wigley, *The impact of regime type on health: does redistribution explain everything?* World Politics, 2011. **63**(4): p. 647-+.
40. Iqbal, Z., *Health and human security: The public health impact of violent conflict.* International Studies Quarterly, 2006. **50**(3): p. 631–649.
41. McKee, M. and E. Nolte, *Lessons from health during the transition from communism.* British Medical Journal, 2004. **329**(7480): p. 1428–1429.
42. Figueras, J., et al., *Health systems in transition: learning from experience*. 2004: World Health Organization, European Observatory on Health Systems and Policies.

43. Stockholm Network Research Team, *Health care reform in Central and Eastern Europe: Setting the stage for discussion*. 2008: Stockholm.
44. Koupilova, I., M. McKee, and J. Holcik, *Neonatal mortality in the Czech Republic during the transition.* Health Policy, 1998. **46**(1): p. 43–52.
45. Nolte, E., R. Scholz, and M. McKee, *Progress in health care, progress in health? Patterns of amenable mortality in central and eastern Europe before and after political transition.* Demographic Research, 2004(Special Collection 2): p. 139–162.
46. Caminada, K. and K. Goudswaard, *Effectiveness of poverty reduction in the eu: a descriptive analysis.* Poverty & Public Policy, 2009. **1**(2).
47. Alesina, A., R. Baqir, and W. Easterly, *Public goods and ethnic divisions.* Quarterly Journal of Economics, 1999. **114**(4): p. 1243–1284.
48. Mackenbach, J.P., *Cultural values and population health: a quantitative analysis of variations in cultural values, health behaviours and health outcomes among 42 European countries.* Health Place, 2014. **28**: p. 116–32.
49. Hofstede, G., *Dimensionlizing cultures: the Hofstede model in context*, in *Online Readings in Psychology and Culture*. 2011: http://scholarworks.gvsu.edu/cgi/viewcontent.cgi?article=1014&context=orpc (accessed 01 December 2014).
50. Kraal, K., J. Roosblad, and J. Wrench, *Equal opportunities and ethnic inequality in European labour markets: discrimination, gender and policies of diversity*. 2014: Amsterdam University Press.
51. Mackenbach, J.P., *Income inequality and population health.* BMJ, 2002. **324**(7328): p. 1–2.
52. Rathmann, K., et al., *Macro-level determinants of young people's subjective health and health inequalities: a multilevel analysis in 27 welfare states.* Maturitas, 2015. **80**(4): p. 414–20.
53. Chauvel, L. and A.K. Leist, *Socioeconomic hierarchy and health gradient in Europe: the role of income inequality and of social origins.* Int J Equity Health, 2015. **14**(1): p. 132.
54. Clough-Gorr, K.M., M. Egger, and A. Spoerri, *A Swiss paradox? Higher income inequality of municipalities is associated with lower mortality in Switzerland.* Eur J Epidemiol, 2015. **30**(8): p. 627–36.
55. Rozer, J.J. and B. Volker, *Does income inequality have lasting effects on health and trust?* Soc Sci Med, 2016. **149**: p. 37–45.
56. Regidor, E., et al., *Low mortality in the poorest areas of Spain: adults residing in provinces with lower per capita income have the lowest mortality.* Eur J Epidemiol, 2015. **30**(8): p. 637–48.
57. Pickett, K.E. and R.G. Wilkinson, *Income inequality and health: a causal review.* Soc Sci Med, 2015. **128**: p. 316–26.
58. Mellor, J.M. and J. Milyo, *Reexamining the evidence of an ecological association between income inequality and health.* J Health Polit Policy Law, 2001. **26**(3): p. 487–522.
59. Avendano, M., *Correlation or causation? Income inequality and infant mortality in fixed effects models in the period 1960–2008 in 34 OECD countries.* Soc Sci Med, 2012. **75**(4): p. 754–60.
60. Jutz, R., *The role of income inequality and social policies on income-related health inequalities in Europe.* Int J Equity Health, 2015. **14**(1): p. 117.
61. Aittomaki, A., et al., *Household income and health problems during a period of labour-market change and widening income inequalities - a study among the Finnish population between 1987 and 2007.* Soc Sci Med, 2014. **100**: p. 84–92.
62. Mosquera, P.A., et al., *Income-related inequalities in cardiovascular disease from mid-life to old age in a Northern Swedish cohort: A decomposition analysis.* Soc Sci Med, 2016. **149**: p. 135–44.
63. Levin, K.A., et al., *National Income and Income Inequality, Family Affluence and Life Satisfaction Among 13 year Old Boys and Girls: A Multilevel Study in 35 Countries.* Soc Indic Res, 2011. **104**(2): p. 179–194.

64. van Doorslaer, E. and X. Koolman, *Explaining the differences in income-related health inequalities across European countries*. Health Economics, 2004. **13**(7): p. 609–628.
65. Kinge, J.M., L. Vallejo-Torres, and S. Morris, *Income related inequalities in avoidable mortality in Norway: A population-based study using data from 1994–2011*. Health Policy, 2015. **119**(7): p. 889–98.
66. Reuveny, R. and Q. Li, *Economic openness, democracy, and income inequality - An empirical analysis*. Comparative Political Studies, 2003. **36**(5): p. 575–601.
67. Sirowy, L. and A. Inkeles, *The Effects of Democracy on Economic-Growth and Inequality - a Review*. Studies in Comparative International Development, 1990. **25**(1): p. 126–157.
68. Gradstein, M. and B. Milanovic, *DOES LIBERTÉ = EGALITÉ? A survey of the empirical evidence on the links between political democracy and income inequality*. 2000: CESifo Working Paper No. 261, http://siteresources.worldbank.org/INTDECINEQ/Resources/Does_Liberte=Egalite.pdf (accessed 19 February 2016).
69. Lee, C.S., *Income inequality, democracy, and public sector size*. American Sociological Review, 2005. **70**(1): p. 158–181.
70. Acemoglu, D., et al., *Democracy, redistribution and inequality*. 2013: NBER working paper series, Working paper 19746, http://www.nber.org/papers/w19746 (accessed 19 February 2016)).
71. Bollen, K.A. and R.W. Jackman, *Political Democracy and the Size Distribution of Income*. American Sociological Review, 1985. **50**(4): p. 438–457.
72. Gradstein, M., B. Milanovic, and Y. Ying, *Democracy and income inequality: An empirical analysis*. 2001: CESifo Working Paper No. 411.
73. Knutsen, C.H., *Reinvestigating the reciprocal relationship between democracy and income inequality*. Review of Economics and Institutions, 2015. **6**(2): p. Article 1.
74. Muller, E.N., *Democracy, Economic-Development, and Income Inequality*. American Sociological Review, 1988. **53**(1): p. 50–68.
75. Judge, K., et al., *Health inequalities: a challenge for Europe*. 2006: An independent, expert report commissioned by, and published under the auspices of, UK Presidency of the EU. http://ec.europa.eu/health/ph_determinants/socio_economics/documents/ev_060302_rd05_en.pdf (accessed 16 Oct 2014).
76. Singh-Manoux, A., et al., *What does self rated health measure? Results from the British Whitehall II and French Gazel cohort studies*. J Epidemiol Community Health, 2006. **60**(4): p. 364–72.
77. Mackenbach, J.P., et al., *Trends in inequalities in premature mortality: a study of 3.2 million deaths in 13 European countries*. J Epidemiol Community Health, 2015. **69**(3): p. 207–17; discussion 205–6.
78. Barr, B., C. Bambra, and M. Whitehead, *The impact of NHS resource allocation policy on health inequalities in England 2001–11: longitudinal ecological study*. BMJ, 2014. **348**: p. g3231.
79. Bauld, L., K. Judge, and S. Platt, *Assessing the impact of smoking cessation services on reducing health inequalities in England: observational study*. Tob Control, 2007. **16**(6): p. 400–4.
80. Department of Health, *Fair society, healthy lives (the Marmot Review)*. 2010: London: Department of Health. http://www.ucl.ac.uk/whitehallII/pdf/FairSocietyHealthyLives.pdf (accessed 16 Oct 2014).
81. House of Commons, *Health inequalities*. 2009: London: Health Committee of the House of Commons. http://www.publications.parliament.uk/pa/cm200809/cmselect/cmhealth/286/28602.htm (accessed 02 Feb 2015).
82. Department of Health, *Tackling health inequalities: 10 years on*. 2009: London. http://webarchive.nationalarchives.gov.uk/20130107105354/http:/www.dh.gov.uk/prod_consum_dh/groups/dh_digitalassets/documents/digitalasset/dh_098934.pdf (accessed 02 Feb 2015).
83. Mackenbach, J.P., *Has the English strategy to reduce health inequalities failed?* Soc Sci Med, 2010. **71**(7): p. 1249–53; discussion 1254–8.

84. Mackenbach, J.P., *Can we reduce health inequalities? An analysis of the English strategy (1997–2010)*. Journal of Epidemiology and Community Health, 2011. **65**(7): p. 568–575.
85. Judge, K. and L. Bauld, *Learning from policy failure? Health action zones in England.* European Journal of Public Health, 2006. **16**(4): p. 341–343.
86. Exworthy, M., L. Berney, and M. Powell, *'How great expectations in Westminster may be dashed locally': the local implementation of national policy on health inequalities.* Policy and Politics, 2002. **30**(1): p. 79–96.
87. Hunter, D.J., et al., *Getting to grips with health inequalities at last?* British Medical Journal, 2010. **340**.
88. Horton, R., *What the UK government is (not) doing about health inequalities.* Lancet, 2002. **360**(9328): p. 186.
89. Macintyre, S., et al., *Using evidence to inform health policy: case studs.* British Medical Journal, 2001. **322**(7280): p. 222–225.
90. Bambra, C., et al., *A labour of Sisyphus? Public policy and health inequalities research from the Black and Acheson Reports to the Marmot Review.* J Epidemiol Community Health, 2011. **65**(5): p. 399–406.
91. Thomas, S., et al., *Population tobacco control interventions and their effects on social inequalities in smoking: systematic review.* Tob Control, 2008. **17**(4): p. 230–7.
92. Peretti-Watel, P., O. L'Haridon, and V. Seror, *Responses to increasing cigarette prices in France: how did persistent smokers react?* Health Policy, 2012. **106**(2): p. 169–76.
93. Department of Health, *Independent inquiry into inequalities in health (the Acheson report)*. 1998: London. https://www.gov.uk/government/publications/independent-inquiry-into-inequalities-in-health-report (accessed 16 Oct 2014).
94. Department of Health, *Tackling health inequalities: 2007 status report of the program for action*. 2007: London. http://webarchive.nationalarchives.gov.uk/+/www.dh.gov.uk/en/Publicationsandstatistics/Publications/DH_083471 (accessed 02 Feb 2015).
95. Department of Health, *Tackling health inequalities: A program for action*. 2003: London. http://www.dh.gov.uk/prod_consum_dh/groups/dh_digitalassets/@dh/@en/documents/digitalasset/dh_4019362.pdf (accessed 02 Feb 2015).
96. Department of Health, *Tackling health inequalities: Cross-cutting review*. 2002: London. http://www.dh.gov.uk/prod_consum_dh/groups/dh_digitalassets/@dh/@en/documents/digitalasset/dh_4068003.pdf (accessed 02 Feb 2015).
97. Ministry of Social Affairs and Health, *National Action Plan to Reduce Health Inequalities 2008–2011*. Publications of the Ministry of Social Affairs and Health Vol. 25. 2008, Helsinki.
98. Mackenbach, J.P. and K. Stronks, *A strategy for tackling health inequalities in the Netherlands.* BMJ, 2002. **325**(7371): p. 1029–32.
99. Mackenbach, J.P. and K. Stronks, *The development of a strategy for tackling health inequalities in the Netherlands.* Int J Equity Health, 2004. **3**(1): p. 11.
100. Muntaner, C., et al., *A Macro-Level Model of Employment Relations and Health Inequalities.* International Journal of Health Services, 2010. **40**(2): p. 215–221.
101. Muntaner, C., et al., *The Solution Space: Developing Research and Policy Agendas to Eliminate Employment-Related Health Inequalities.* International Journal of Health Services, 2010. **40**(2): p. 309–314.
102. Reeves, A., et al., *Do employment protection policies reduce the relative disadvantage in the labour market experienced by unhealthy people? A natural experiment created by the Great Recession in Europe.* Social Science & Medicine, 2014. **121**: p. 98–108.
103. Muntaner, C., et al., *Welfare state, labour market inequalities and health. In a global context: An integrated framework. SESPAS report 2010.* Gaceta Sanitaria, 2010. **24**: p. 56–61.

104. O'Campo, P., et al., *Social welfare matters: A realist review of when, how, and why unemployment insurance impacts poverty and health*. Social Science & Medicine, 2015. **132**: p. 88–94.
105. Ferrarini, T., K. Nelson, and O. Sjoberg, *Unemployment insurance and deteriorating self-rated health in 23 European countries.* Journal of Epidemiology and Community Health, 2014. **68**(7): p. 657–662.
106. Molnar, A., et al., *Protocol: Realist synthesis of the impact of unemployment insurance policies on poverty and health*. Evaluation and Program Planning, 2015. **48**: p. 1–9.
107. Burmaster, K.B., et al., *Impact of a private sector living wage intervention on depressive symptoms among apparel workers in the Dominican Republic: a quasi-experimental study.* Bmj Open, 2015. **5**(8).
108. Herttua, K., P. Makela, and P. Martikainen, *Changes in Alcohol-Related Mortality and its Socioeconomic Differences After a Large Reduction in Alcohol Prices: A Natural Experiment Based on Register Data.* American Journal of Epidemiology, 2008. **168**(10): p. 1110–1118.
109. Yongs, R., *Is the European Union supporting democracy in its neighbourhood ?* 2008, Spain: Priority Producción, Diseño y Comunicación S.L.
110. Pinzari, D., *EU democratization policies in the neighbourhood countries and Russia's reaction as a destabilizing factor. A comparative case study of Georgia and Moldova*. 2015: Bruges Political Research Papers 45/ 2015, https://www.coleurope.eu/research-paper/eu-democratization-policies-neighbourhood-countries-and-russias-reaction (accessed 19 February 2016).
111. Wilkinson, R. and K. Pickett, *The spirit level: Why more equal societies almost always do better*. 2009, United Kingdom: Allen Lane.
112. Giskes, K., et al., *Trends in smoking behaviour between 1985 and 2000 in nine European countries by education.* J Epidemiol Community Health, 2005. **59**(5): p. 395–401.
113. Dalstra, J.A., et al., *Trends in socioeconomic health inequalities in the Netherlands, 1981–1999.* J Epidemiol Community Health, 2002. **56**(12): p. 927–34.
114. Rana, C., V.O. Herman, and D. Stefaan, *Trends in social inequalities in obesity: Belgium, 1997 to 2004.* Preventive Medicine, 2009. **48**: p. 54–58.
115. Molarius, A., et al., *Educational level, relative body weight, and changes in their association over 10 years: an international perspective from the WHO MONICA Project*. Am J Public Health, 2000. **90**(8): p. 1260–8.
116. Helasoja, V., et al., *Trends in the magnitude of educational inequalities in health in Estonia, Latvia, Lithuania and Finland during 1994–2004.* Public Health, 2006. **120**(9): p. 841–53.
117. Scholes, S., et al., *Persistent socioeconomic inequalities in cardiovascular risk factors in England over 1994–2008: a time-trend analysis of repeated cross-sectional data.* BMC Public Health, 2012. **12**: p. 129.
118. Nagelhout, G.E., et al., *Trends in socioeconomic inequalities in smoking prevalence, consumption, initiation, and cessation between 2001 and 2008 in the Netherlands. Findings from a national population survey.* BMC Public Health, 2012. **12**: p. 303.
119. Imkampe, A.K. and M.C. Gulliford, *Increasing socio-economic inequality in type 2 diabetes prevalence--repeated cross-sectional surveys in England 1994–2006.* Eur J Public Health, 2011. **21**(4): p. 484–90.
120. Talala, K., et al., *Trends in socio-economic differences in self-reported depression during the years 1979–2002 in Finland.* Soc Psychiatry Psychiatr Epidemiol, 2009. **44**(10): p. 871–9.
121. Mackenbach, J.P., *Socialeconomic inequalities in health in high-income countries: the facts and the options*. In: Detels, R. Gulliford, M., Karim, Q.A., Tan, C.C., eds. Oxford Textbook of Global Public Health. 2015, Oxford: Oxford University Press.
122. Cavelaars, A.E., et al., *Educational differences in smoking: international comparison.* BMJ, 2000. **320**(7242): p. 1102–7.
123. Huisman, M., A.E. Kunst, and J.P. Mackenbach, *Educational inequalities in smoking among men and women aged 16 years and older in 11 European countries.* Tob Control, 2005. **14**(2): p. 106–13.

Summary

SUMMARY

European countries vary substantially in many ways, including their history, climate, political systems, welfare regimes, culture, and health. With regard to the latter, some European countries appear to have the best levels of population health world-wide, while others achieve much lower levels of population health. The between-country variation in health as well as differences in political and social developments in Europe offer good opportunities for research on macro-social determinants of population health. It helps us to gain deeper insights into why some countries are more successful than others in improving health. Although several factors have been suggested as potential determinants of between-country differences in population health, unanswered questions still remain. For example, being two of the most frequently studied potential determinants, democracy and income inequality have been consistently found to be associated with health in many cross-national studies. However, whether these associations have a causal interpretation is still controversial. Moreover, studies specifically assessing the potential impacts of democracy and income inequality on health in a European context, which would be important for policy makers in Europe, are limited in number.

Similarly, socioeconomic inequalities in health, which have been found in all European countries with available data, vary substantially in size across countries and have been a major challenge to public health in Europe. Although there is an increasing number of studies describing sizes and trends of socioeconomic inequalities in health in countries, an important question is how such inequalities can be tackled. Variations in the magnitude and trend in socioeconomic inequalities in health in Europe can be used for research on underlying determinants of health inequalities or the evaluation of strategies aimed at tackling health inequalities. Methodological challenges, however, such as establishing causal interpretations were only addressed in a limited number of studies.

Natural policy experiments, typically defined as "policies that are not under the control of the researchers, but which are amenable to research using the variation in exposure that they generate to analyze their impact" may hold great promise. Whereas the evaluation of natural policy experiments for population health becomes increasingly popular, the number of studies evaluating policy effects on health inequalities is still very small. Again, Europe's diversity offers a good but currently limited explored setting for the evaluation of policies aimed at tackling health inequalities.

The aim of this thesis was to contribute to the understanding of the potential relationship between social and political factors and population health and health inequalities from an international, European perspective, and to provide potential methods to improve the evaluation of the impacts of social and political factors on health and health inequalities.

Econometric regression technique was used as the main analytical tool due to its ability to reduce confounding and improve the causal inference of the relationships studied.

In the first part of this thesis, comprising chapters 2 and 3, we assessed whether democracy and income inequality were associated with population health, as indicated by life expectancy and cause-specific mortality rates in multiple European countries over decades.

In chapter 2, we examined whether democratization was associated with life expectancy and cause-specific mortality rates. Our sample covered 29 European countries in the period of 1960–1990 and 43 European countries in the period of 1987–2008. Country level characteristics and country fixed effects were controlled in the regressions. Democratization was measured by both "current democracy" (i.e. the democracy index measured in the same year as the population health, which captured short-term effect of democracy) and "cumulative democracy" (i.e. the weighted cumulative democracy index since an earlier year, which captured long-term and cumulative effect of democracy). Our results suggest that during these two periods democratization had overall positive effects on life expectancy in Europe. In the 1960–1990 period, current democracy was strongly associated with higher life expectancy. The positive effects of current democracy on total mortality were mainly the results from lower mortality from all heart diseases, pneumonia, liver cirrhosis, and suicide. In the 1987–2008 period, however, we found that current democracy was associated with lower, and cumulative democracy with higher life expectancy, particularly among men. The positive effects of cumulative democracy on total mortality were mainly the result from lower mortality from circulatory diseases, breast cancer, and external causes. Current democracy was associated with higher mortality from motor vehicle accidents in both periods, and also with higher mortality from cancer and all external causes in the second. That short-term changes in levels of democracy had positive effects in the first but not in the second period is probably due to the fact that democratization in Central and Eastern Europe in the second period was part of a complete system change which caused major societal disruptions.

In chapter 3, we explored whether changes in income inequality were associated with improvements in population health as measured by life expectancy and cause-specific mortality rates in 43 European countries between 1987 and 2008. We showed that statistically significant associations between income inequality and many mortality indicators were found in pooled cross-sectional regressions, indicating higher mortality in countries with larger income inequalities. However, once the country fixed effects were added, all associations between income inequality and mortality indicators became statistically insignificant, except for mortality from external causes and homicide among men, and cancers among women. The statistically significant results for homicide and cancers disappeared after further adjustment for indicators of democracy, education, transition to national independence, armed conflicts, and economic freedom. Our findings suggested that cross-sectional associations between income inequality and mortality seem to reflect the confounding effects of other country characteristics. In a European context, national levels of income inequality did not have an independent effect on mortality and may not able to causally explain variations in population health.

In the second part of this thesis, containing chapters 4 to 7, we focused on socioeconomic inequalities in Europe, with a special emphasis on the impact of health policies implemented in Europe on socioeconomic inequalities in health.

In chapter 4, we provided an overview of trends in socioeconomic inequalities in self-assessed health among men and women aged 30–79 years in seventeen European countries between 1990 and 2010. Declining trends in prevalence of less-than-good self-assessed health were observed in many countries, particularly in Southern and Eastern Europe and the Baltic States. In all countries, less-than-good self-assessed health was more prevalent in lower educational and manual groups. For all countries together, absolute inequalities in self-assessed health were mostly constant, whereas relative inequalities increased. Almost no country consistently experienced a statistically significant decline in either absolute or relative inequalities. After referring to national policies tackling health inequalities in some countries, we found that countries with national policies aimed at tackling health inequalities have not fared better, in terms of inequalities in self-assessed health, than countries without such policies.

In chapter 5, we explored the potential methods for evaluating policy effects in reducing health inequalities. We identified seven methods mostly originating in econometrics for the evaluation of natural policy experiments: regression adjustment, propensity score matching, difference-in-differences, fixed effect analysis, instrumental variable, regression discontinuity and interrupted time-series. Using a fictitious numerical example, we explored how to apply these methods to assess policy effects on the magnitude of health inequalities. We found that all methods can be used to assess the impact of policies on health inequalities, either by doing a stratified analysis or by including interaction terms (the latter, however, was impossible with propensity score matching). The choice of the methods depends on data availability and the nature of the confounders in the analysis (whether observable or time-variant). However, the low external validity of results from instrumental variable and regression discontinuity makes them less desirable for assessing policy effect on population-level health inequalities.

In chapter 6, we empirically evaluated the effectiveness of an ambitious programme pursued by the English government to reduce health inequalities. Trends in health inequalities in England both over time and between countries were compared. We showed that although more favourable trends in some health indicators were observed among low-educated people after the implementation of the English strategy, trends in health inequalities in 2000–2010 in England were not more favourable than those observed in the period 1990–2000, and changes in trends of health inequalities after 2000 in England were also not statistically significantly more favourable than those seen in the other countries. Hence, our results did not suggest a favourable effect of the English strategy.

In chapter 7, we investigated the impact of price and non-price related population-wide tobacco control policies on smoking by socioeconomic group in nine European countries between 1990 and 2007. We demonstrated that the price of popular cigarettes and non-price policies were negatively associated with smoking among men, and the price of the cheapest cigarettes was negatively associated with smoking among women. While these favourable effects were generally in the same direction for all socioeconomic groups, they were larger and statistically significant in lower socioeconomic groups only. Our findings suggested that tobacco control policies as implemented in the nine European countries, have probably helped to reduce the prevalence of smoking in the total population, particularly in lower socioeconomic groups, thereby contributing to a reduction of inequalities in smoking.

Overall, we demonstrated that democracy is beneficial for country-level population health, while income inequality is not a causal determinant of the international variations in public health in Europe; the effects of national policies tackling health inequalities might be limited in several European countries, and the tobacco control policies which are not implemented with a focus on tackling health inequalities may have the potential to narrow health inequalities. From this perspective, policies that promote democracy should be encouraged not only from the view of political justice, but also from the aspect of improving general health, and we need to allow democratization the time that is needed to realize important and persistent health benefits. In contrast, although reducing income inequalities is important for creating equality of opportunity, it should not be regarded as an effective strategy to improve population life expectancy and reduce mortality in Europe. As regarding to socioeconomic inequalities in health, policy makers need to develop more effective policies in reducing inequalities in health, especially the policies that will be implemented at a sufficient scale, with proven effectiveness and entry-points which are important determinants of health inequalities. One potential policy area can be considered is tobacco control. Future research need to focus more on providing evidence for the effectiveness of policies in reducing socioeconomic inequalities in health and choose the outcomes that are directly related to the policies that we are interested in. The methods we identified can be useful tools to evaluate policy effects on health and health inequalities. These methods need be considered by future research with the aim to improve the causal interpretations of the potential social and political determinants of health and health inequalities, and build a good evidence base for policy makers and stakeholders.

Samenvatting

SAMENVATTING

Europese landen verschillen onderling aanzienlijk in hun geschiedenis, klimaat, politieke systemen, welvaartregimes, en cultuur. Ook wat betreft de volksgezondheid bestaan er belangrijke verschillen. De verschillen in macrosociale factoren en gezondheid tussen Europese landen bieden goede mogelijkheden voor onderzoek naar de samenhang tussen beide. Dit helpt om meer inzicht te krijgen in de vraag hoe de volksgezondheid van landen kan worden verbeterd. Hoewel verschillende macrosociale factoren zijn geopperd als potentiële determinanten van verschillen in volksgezondheid tussen landen, zijn er nog steeds onbeantwoorde vragen. Democratie en inkomensongelijkheid zijn bijvoorbeeld vaak in verband gebracht met gezondheid in internationaal vergelijkende studies. Het is echter controversieel of deze verbanden causaal zijn.

In alle Europese landen met beschikbare gegevens bestaan sociaaleconomische gezondheidsverschillen. Onderzoek heeft aangetoond dat deze verschillen variëren in omvang tussen landen. Het verkleinen van sociaaleconomische gezondheidsverschillen is een grote uitdaging in Europees volksgezondheidsbeleid. Variaties in de omvang en ontwikkeling van sociaaleconomische gezondheidsverschillen over de tijd in Europa kunnen worden gebruikt voor onderzoek naar de onderliggende determinanten van gezondheidsverschillen, en voor de evaluatie van strategieën gericht op de aanpak van gezondheidsverschillen. Tot op heden kwamen methodologische uitdagingen, zoals het vaststellen van causale interpretaties, nog maar in een beperkt aantal studies aan bod.

Natuurlijke beleidsexperimenten, gedefinieerd als "beleidsmaatregelen die niet onder de controle staan van de onderzoekers, maar die kunnen worden gebruikt voor onderzoek dankzij de variatie in blootstelling die zij genereren", kunnen veelbelovend zijn. De evaluatie van natuurlijke beleidsexperimenten voor de volksgezondheid krijgt steeds meer belangstelling van onderzoekers, maar het aantal studies dat het effect van deze experimenten op sociaaleconomische ongelijkheden in gezondheid evalueert, is nog steeds erg klein.

Het doel van dit proefschrift was (a) bij te dragen aan het begrijpen van de mogelijke relatie tussen sociale en politieke factoren enerzijds en de volksgezondheid en gezondheidsverschillen anderzijds vanuit een Europees perspectief, en (b) te voorzien in methoden om de evaluatie van de impact van sociale en politieke factoren op de volksgezondheid en gezondheidsverschillen te verbeteren. Dit is gedaan door toepassing van methoden en technieken die eerder al in de economie en econometrie werden toegepast, en vaak het vermogen hebben rekening te houden met zowel gemeten als ongemeten verstorende variabelen. Als gevolg hiervan wordt het aannemelijker dat een gevonden verband een causaal verband is.

In het eerste deel van dit proefschrift, bestaande uit de hoofdstukken 2 en 3, werd onderzocht of de mate van democratie en inkomensongelijkheid in Europese landen is geassocieerd met de volksgezondheid.

In hoofdstuk 2 werd het verband onderzocht tussen democratisering en de levensverwachting en doodsoorzaakspecifieke sterfte. Gegevens waren beschikbaar van 29 Europese landen in de periode van 1960 tot 1990, en van 43 Europese landen in de periode van 1987 tot 2008. In "fixed-effects"regressieanalyses werd rekening gehouden met gemeten en ongemeten (voor zover constant over de tijd) kenmerken van landen, die zowel van invloed zouden kunnen zijn op de mate van democratie als de volksgezondheid. De mate van democratisering werd gemeten in hetzelfde jaar als de indicatoren van de volksgezondheid) ("huidige democratie"), en kon zo het korte termijn effect van de mate van democratisering vastleggen. Het effect werd ook bepaald met behulp van een cumulatieve index, waarin ook gegevens van een eerder jaar zaten. Onze resultaten suggereren dat gedurende deze twee periodes, democratisering over het algemeen positieve effecten had op de levensverwachting in Europa. In de periode van 1960 tot 1990 bleek de index voor de "huidige democratie" sterk geassocieerd met een hogere levensverwachting. De daling in de sterfte bij een toenemende maat voor democratie was vooral het resultaat van een lagere sterfte aan hart- en vaatziekten, longontsteking, levercirrose en zelfmoord. In de periode van 1987 tot 2008 was de maat voor de huidige democratie geassocieerd met een lagere levensverwachting, terwijl de cumulatieve democratie was geassocieerd met een hogere levensverwachting, vooral onder mannen. De positieve effecten van de cumulatieve democratie op de totale sterfte waren vooral het gevolg van lagere sterfte aan hart- en vaatziekten, borstkanker en externe oorzaken. De maat voor de huidige democratie werd geassocieerd met een hogere sterfte door ongevallen met motorvoertuigen in beide periodes, maar ook met hogere sterfte aan kanker en alle externe oorzaken in de tweede periode (van 1987 tot 2008). Veranderingen op korte termijn in het niveau van democratie hadden positieve effecten in de eerste periode, maar niet in de tweede periode. Dit is waarschijnlijk te wijten aan het feit dat democratisering in Centraal en Oost-Europa in de tweede periode onderdeel was van een complete systeemverandering, die belangrijke maatschappelijke problemen veroorzaakte.

In hoofdstuk 3 werd onderzocht of veranderingen in inkomensongelijkheid waren geassocieerd met verbeteringen in de levensverwachting en de doodsoorzaakspecifieke sterfte in 43 Europese landen tussen 1987 en 2008. In een analyse op basis van alle gegevens werden statistisch significante verbanden gevonden tussen inkomensongelijkheid en veel oorzaken van sterfte, waarbij de sterfte hoger was in landen met grotere inkomensverschillen. Echter, zodra rekening werd gehouden met gemeten en tijdsonafhankelijke ongemeten verstorende variabelen waren alle verbanden niet langer statistisch significant, met uitzondering van sterfte door externe oorzaken en doodslag bij mannen, en sterfte door kanker bij vrouwen. Deze statistisch significante resultaten waren niet langer significant als ook rekening werd gehouden met indicatoren voor de mate van democratie, onderwijs, de overgang naar nationale onafhankelijkheid, gewapende conflicten en economische vrijheid. Onze bevindingen suggereerden dat verbanden tussen inkomensongelijkheid en sterfte het gevolg zijn van de verstorende effecten van andere landspecifieke kenmerken. Binnen Europa blijkt de mate van inkomensongelijkheid geen onafhankelijk effect te hebben op sterfte.

In het tweede deel van dit proefschrift, bestaande uit de hoofdstukken 4 tot 7, werd de variatie in sociaaleconomische ongelijkheid in Europa onderzocht, met een bijzondere nadruk op de gevolgen van geïmplementeerd gezondheidsbeleid in Europa op sociaaleconomische gezondheidsverschillen.

In hoofdstuk 4, werden trends in sociaaleconomische ongelijkheid in de "ervaren gezondheid" tussen 1990 en 2010 onderzocht van 30-79 jarige mannen en vrouwen in zeventien Europese landen. Dalende trends in de prevalentie van een minder-dan-goede ervaren gezondheid werden waargenomen in veel landen, met name in Zuid- en Oost-Europa en de Baltische Staten. In alle landen was de prevalentie van een minder-dan-goede ervaren gezondheid hoger bij lager opgeleide mensen en bij mensen met een lager beroepsniveau. In alle landen bleek de trend in absolute sociaaleconomische verschillen in ervaren gezondheid stabiel, terwijl de relatieve ongelijkheid veelal toenam. In vrijwel geen enkel land werd een statistisch significante daling in zowel absolute als relatieve ongelijkheden gevonden. Landen met specifiek beleid gericht op het terugdringen van sociaaleconomische gezondheidsverschillen, verging het niet beter dan landen zonder dergelijk beleid.

In hoofdstuk 5 werden de potentiële methoden voor de evaluatie van natuurlijke beleidsexperimenten onderzocht gericht op het verkleinen van sociaaleconomische gezondheidsverschillen. Hiervoor werden zeven methoden geïdentificeerd, veelal met een oorsprong in de econometrie: 1) regressie analyse met correctie voor verstorende variabelen, 2) 'propensity score matching', 3) 'difference-in-differences', 4) 'fixed effects' analyse, 5) instrumentele variabele analyse, 6) 'regression discontinuity' en 7) 'interrupted time series'. Met behulp van een fictief rekenvoorbeeld werd onderzocht hoe deze methoden kunnen worden toegepast om het effect van natuurlijke beleidsexperimenten op de omvang van gezondheidsverschillen te vast te stellen. Uit de studie bleek dat alle methoden kunnen worden gebruikt voor dit doel, hetzij in een gestratificeerde analyse of door het opnemen van statistische interactietermen. Dit laatste was alleen onmogelijk met de 'propensity score matching' methode. De keuze van de methode hangt af van de beschikbare data, en van de aard van de verstorende variabelen in de analyse (gemeten, ongemeten, variërend over de tijd of niet). De lage externe validiteit van de resultaten verkregen met behulp van de instrumentele variabele analyse en de 'regression discontinuity' maakt deze methoden minder geschikt voor het beoordelen van het effect van natuurlijke beleidsexperimenten op sociaaleconomische gezondheidsverschillen.

In hoofdstuk 6 werd de effectiviteit van het grote en ambitieuze Engelse programma ter bestrijding van sociaaleconomische gezondheidsverschillen geëvalueerd. Hiervoor werden trends in gezondheidsverschillen in Engeland voor en na de implementatie van het programma onderzocht, en werden deze vergeleken met trends in dezelfde periode in 4 landen variërend in beleid op dit terrein. Na de implementatie van de Engelse strategie werden gunstige ontwikkelingen in sommige gezondheidsindicatoren waargenomen onder laagopgeleide personen. Ondanks deze gunstigere trends bleken ontwikkelingen in gezondheidsverschillen tussen 2000 en 2010 in Engeland niet gunstiger dan de ontwikkelingen waargenomen in de periode van 1990 tot 2000. Veranderingen in de ontwikkeling van gezondheidsverschillen na 2000 in Engeland waren ook niet gunstiger dan de trends waargenomen in de andere landen. Onze resultaten wijzen er dus niet op dat de Engelse strategie er in is geslaagd de sociaaleconomische gezondheidsverschillen voor de gemeten gezondheidsindicatoren te verkleinen.

In hoofdstuk 7 werd de impact van prijs- en "niet-prijs" (zoals een rookverbod of beperking op openbare plaatsen en op het werk, een verbod op reclame en promotie van tabaksartikelen, labels met gezondheidswaarschuwingen op tabaksartikelen, en stoppen-met-roken diensten) gerelateerd tabaksontmoedigingsbeleid op sociaaleconomische verschillen in roken in negen Europese landen tussen 1990 en 2007 onderzocht. De prijs van de meest populaire sigaretten en niet-prijsgerelateerd antirookbeleid werden negatief geassocieerd met roken onder mannen, en de prijs van de goedkoopste sigaretten werd negatief geassocieerd met het roken onder vrouwen. Hoewel deze gunstige effecten in het algemeen dezelfde richting hadden voor alle sociaaleconomische groepen, waren de effecten groter en alleen statistisch significant voor de lagere sociaaleconomische groepen. Onze bevindingen suggereren dat anti-rookbeleid, zoals geïmplementeerd in negen Europese landen, waarschijnlijk heeft bijgedragen aan het terugdringen van rookgedrag voor de gehele populatie, vooral in de lagere sociaaleconomische groepen, en zo ook heeft bijgedragen aan de vermindering van de ongelijkheden in roken tussen sociaaleconomische groepen.

Samenvattend hebben de studies aangetoond dat democratie gunstig is voor de volksgezondheid, terwijl inkomensongelijkheid geen causale determinant is van internationale verschillen in de volksgezondheid in Europa. Er werd ook aangetoond dat in verschillende Europese landen de effecten van nationaal beleid gericht op het bestrijden van gezondheidsongelijkheden beperkt blijken te zijn. Anti-rookbeleid, dat niet primair is geïmplementeerd om ongelijkheden in gezondheid te bestrijden, lijkt wel te kunnen bijdragen aan het verkleinen van de gezondheidsverschillen. Het lijkt er op dat beleid gericht op de bevordering van democratie ook vanuit een volksgezondheidsperspectief moet worden gestimuleerd. Hierbij moet wel worden benadrukt dat het tijd kost voordat meer democratie de volksgezondheid zichtbaar verbetert. Hoewel het verminderen van inkomensongelijkheid belangrijk is voor het creëren van gelijke kansen, moet dit niet worden beschouwd als een effectieve strategie om de levensverwachting te verbeteren of sterfte te verminderen in Europa. Met betrekking tot het verder terugdringen van sociaaleconomische gezondheidsverschillen moet effectiever beleid worden ontwikkeld. Zulk beleid moet op voldoende grote schaal worden geïmplementeerd, en aangrijpen op determinanten waarvan het causale effect is vastgesteld. Een belangrijk beleidsdomein hierbij is de bestrijding van tabaksgebruik. Toekomstig onderzoek zou zich meer moeten richten op het leveren van bewijs voor de effectiviteit van beleid in het verminderen van sociaaleconomische gezondheidsverschillen, en op het kiezen van uitkomsten die rechtstreeks verband houden met het beleid waarin we geïnteresseerd zijn. De geïdentificeerde methoden moeten als belangrijk worden beschouwd in toekomstig onderzoek naar sociale en politieke determinanten van gezondheid en gezondheidsverschillen.

Acknowledgements

ACKNOWLEDGEMENTS

Whenever I think back, I marvel at how right the decision of moving to the Netherlands was. It became an unforgettable and enjoyable experience in my life. I have gained a lot of "first times" in the four years' journey: the first time of going abroad and being a minority, the first time of working in an international project, the first time of presenting at an international conference, the first time of publishing in an international journal, and also my first times of skiing, curling, hiking and kayaking. Soon, I will (and hope to) have my first time of being called Dr. Hu. Let me take this opportunity to express my gratitude towards the very people who made it so.

I cannot thank my promotor professor Johan Mackenbach enough for offering the opportunity to work at MGZ. It seemed impossible to get a research position in a European project for a person without any public health background or working experience in Europe. Dear Johan, thanks a lot for your trust from the beginning, for your support and for your extremely valuable feedback on my research in the past years. I have learned a lot from you, not only your endless knowledge in this field, but also your efficiency and concentration. I feel so lucky having you as my promotor.

I would also like to thank my co-promotor dr. Frank van Lenthe, whose contribution to the completion of this dissertation is invaluable. Dear Frank, thank you for all the mentoring and encouragement, for checking my manuscripts down to details, and for all our Friday meetings. I greatly appreciate the time and the energy you invested in guiding me through my PhD. You were always considerate and patient, which made writing the PhD thesis a great experience rather than a torture. Thank you for being such a great co-promotor.

Further, I would like to express my sincere gratitude to the members of my inner doctoral committee Prof. van Doorslaer, Prof. Stronks and Dr. Avendaño Pabon for taking the time to read and assess my dissertation. I am also indebted to Prof. Polder, Prof. Putters and Dr. van Dalen for agreeing to be part of the plenary committee.

I would like to thank all international project partners of the DEMETRIQ project. I learned a lot from the collaboration and meetings, which allowed me to meet many of you in person and to establish further cooperation. I enjoyed the fruitful discussions related to our research. Thanks also to all co-authors for their contributions to the publications.

I am grateful that I worked with Rasmus, Gwenn, Rianne and Karen, members of the small Rotterdam team of the DEMETRIQ project, for all the ups and downs of every stage of research that we went through together. Rasmus, thanks for walking me around the department the day of my interview. I appreciated your guidance in the beginning of my PhD journey. Later on you moved to Italy, but our collaboration continued. Gwenn, thanks for your great efforts in the data collection. Rianne, you were always kind and supportive. I guess harmonizing the data from DEMETRIQ will be an unforgettable experience for both of us. Karen, many thanks for being such a great colleague and friend. You were always sharing your thoughts, your wisdom, and your positive energy with me. I cherish our memories outside work, eating, chatting and laughing. Also thanks for your help with turning Google-translated Dutch into readable Dutch in my Samenvatting.

Dear Joost and David, thank you for agreeing to be my paranymphs. It is amazing to have support from two clever and handsome gentlemen at this final step. Joost, I really appreciate your smartness and humour. Chatting with you is such a delightful thing and you are always so encouraging. David, we met on the day of my interview and I was very impressed by your kindness. All the pleasant chats, coffee breaks and dinners in the past years reinforced this impression. I deeply cherish all the warm feelings both of you have given to me.

Thanks also to my lovely roomies. Alex and Reinier, you deserve a special mention, as you have accompanied me for almost my whole PhD journey. Thank you for explaining all Dutch things, for always inviting me for lunch and walk, and for sharing happiness and sadness. You made the hard journey much easier. Thanks Valerie, Kirsten, Annemieke and Sonja for your company. I really miss our room's picnic in het park. Also thanks Kwadwo, Jegatharsan, Anouk, Shanty and Rita who welcomed me to the new office. I enjoyed our lunch breaks a lot. You were one of the reasons I looked forward to going to work every day in the last months.

I thank all other MGZ colleagues as well for all the positive interactions over the years. Giorgia, thanks for your friendship and support. I wish I can visit you in Prague in the near future (then our little ones can meet each other). Frederick, thanks for all the nice chats over drinks and dinners. Marielle, thank you for giving me all kinds of good advices about taking care of the little one. Raquel and Camila, I really miss the days we went to pubs for the World Cup. Adi, thank you for your sweetness and emotional support during some difficult moments. Thanks also go to Ivana, Amy, Tiago, Kevin, Maggie, Tanja, Wilma and Astrid, for all the chats, drinks, PubQuiz in the Irish pub and other work or non-work related advices. Many thanks to the ICT team and the secretaries. A big "thank you" goes to Astrid who has always been willing to help me. I would also like to thank Caspar and Gerard for their statistical support.

I would like to thank all my Chinese friends. Special thanks go to Rui and Tianshi for helping me to set up my life in Rotterdam; Qing and Yon for all the nice diners and company; Junwen and Yunyu for your friendship and super delicious fish; Shuo and Hao for all the healing with food and hiking. I also thank Nana, Minghui, Baoyue, Lu, Guannan, Wei and Qiye, and my friends in EUR and VU, Yuan Gu, Chen Li, Yajie Sun, Wenqian Huang and Xiao Yu for all the fun we had together. Thanks for making me feel home here.

My parents, many thanks for the support over the years. You are not only my parents, but also my best friends. I feel so lucky that I can be your daughter. I love you.

Bo, a special big "thank you" goes to you. Thanks for always being there, which is very important to me. You know my ups and downs the best, and you always support me no matter what decision I make. Although I am not sure which country will be our next stop, I wish I can be always by your side and explore the world together with you. Mieke, each of your smiles is a wonderful gift for me. Your arrival went side by side with the completion of this dissertation, thus it is not surprising to find that your favourite thing is reading picture books. If you want to pursue a PhD later in your life, as a mother, I would advise you to think carefully before really doing it.

About the author

ABOUT THE AUTHOR

Yannan Hu was born in the province of Hunan (China) on April 9th 1988. She spent her childhood in Huangshi (a small city in the Hubei province, in the middle of China). After finishing secondary school in 2005, she enrolled at Huazhong University of Science and Technology (located in Wuhan, the capital city of Hubei Province), where she obtained a Bachelor's Degree in Finance. From 2009, she continued her study on economics and econometrics at Shanghai University of Finance and Economics. During her study, she worked for three months as a survey assistant in the project "School Type and Education of Migrant Children", funded by China Natural Science Foundation. In 2010, Yannan was selected as a manager assistant in Hua An Fund Management Co., Ltd. (Shanghai branch) for a three-month internship, the main work of which was data analysis. In the same year, she spent two months working in China Credit Information Service (P.R.C), Ltd. (Shanghai branch) as an intern, where she helped in promoting its business database and market research. In 2011, Yannan obtained a Master degree in Labor Economics. Her master thesis was entitled: "Social Networks, Income, and the Income Inequality in China". After that, she started to work as an accountant in the audit department of KPMG, Shanghai.

In 2012, together with her husband, Yannan moved to the Netherlands, where she was appointed as a researcher at the Department of Public Health of the Erasmus Medical Center in Rotterdam. Here, she was involved in the large European research project – DEMETRIQ, which aimed at developing methodologies to reduce inequalities in the determinants of health. Her research work resulted in hers PhD thesis. Currently, she remains interested in social and political determinants of health and health inequalities, as well as the application of econometric methods in epidemiology and public health.

List of publications

LIST OF PUBLICATIONS

Papers in this thesis

Mackenbach JP, **Hu Y**, Looman C.W.N. Democratization and life expectancy in Europe, 1960–2008. Social Science & Medicine, 2013. 93: p. 166–175.

Hu Y, van Lenthe FJ, Mackenbach JP. Income inequality, life expectancy and cause-specific mortality in 43 European countries, 1987–2008. European Journal of Epidemiology, 2015. 30: p. 615–625.

Hu Y, van Lenthe FJ, Borsboom GJ, Looman CWN, Bopp M, Burström B, Dzúrová D, Ekholm O, Klumbiene J, Lahelma E, Leinsalu M, Regidor E, Santana P, de Gelder R, Mackenbach JP. Trends in socioeconomic inequalities in self-assessed health in 17 European countries between 1990 and 2010. Journal of Epidemiology and Community Health, 2016. Published online first: doi:10.1136/jech-2015–206780.

Hu Y, van Lenthe FJ, Hoffmann R, van Hedel K, Mackenbach JP. Assessing the impact of natural policy experiments on socioeconomic inequalities in health: how to apply commonly used quantitative analytical methods? (submitted)

Hu Y, van Lenthe FJ, Judge K, Lahelma E, Costa G, de Gelder R, Mackenbach JP. Did the English strategy reduce inequalities in health? A difference-in-differences analysis comparing England with three other European countries. (submitted)

Hu Y, van Lenthe FJ, Platt S, Bosdriesz JR, Lahelma E, Menvielle G, Regidor E, Santana P, de Gelder R, Mackenbach JP. The impact of tobacco control policies on smoking among socioeconomic groups in nine European countries, 1990–2007: a fixed effects study. (submitted)

Other publications

Hoffmann R, **Hu Y**, de Gelder R, Menvielle G, Bopp M, Mackenbach JP. The impact of increasing income inequalities on educational inequalities in mortality - An analysis of six European countries. International Journal for Equity in Health, 2016.15(1): p. 103

Hu Y. Eugen Bohm-Bawerk and the Austrian school of economic thought. Market Modernization, 2008. 6: p. 393–394 (published in Chinese).

Hu Y. The application of the neomercantilism in China. China Market, 2008. 26: p.72–73 (published in Chinese).

Hu Y. Suggestions for the labour-intensive enterprises in the period of labour shortage in China. Special Zone Economy, 2008. 8: p.218–219 (published in Chinese).

Phd portfolio

PHD PORTFOLIO

PhD student	Yannan Hu	PhD period	2012–2016
Erasmus MC	Department of Public Health	Promotor	Prof.dr. Johan P. Mackenbach
		Supervisor	Dr. Frank J. van Lenthe

PhD training	Year	Workload (ECTS)
Courses		
Advanced Medical Writing and Editing	2014	0.7
Medical Demography	2013	1.1
Causal Inference	2013	0.7
Methods of Public Health Research	2012	0.7
Social Epidemiology	2012	0.7
Systematic Literature Retrieval in PubMed	2012	0.3
Presentations		
Did the English Strategy Reduce Health Inequalities? European Public Health Conference, Glasgow, Scotland	2014	1.0
STATA Introduction. Social Epidemiology Methods Meeting, Erasmus MC, Rotterdam, the Netherlands	2014	1.0
Did the English Strategy Reduce Health Inequalities? DEMETRIQ End of Project Conference, Leiden, the Netherlands	2014	1.0
Quantitative Methods for the Evaluation of Natural Policy Experiments (NPE) for Their Impact on Health Inequalities. DEMETRIQ End of Project Conference, Leiden, the Netherlands	2014	1.0
Trends in Socioeconomic Inequalities in Self-assessed Health in 16 European Countries Between 1990 and 2010. DEMETRIQ End of Project Conference, Leiden, the Netherlands	2014	1.0
Did the English Strategy Reduce Health Inequalities? -A Trend Comparison of England and Other Four European Countries. DEMETRIQ Work Package Leader Meeting, Rotterdam, the Netherlands	2014	1.0
Democratization and Life Expectancy in Europe, 1960–2008. Social Epidemiology Methods Meeting, Erasmus MC, Rotterdam, the Netherlands	2013	1.0
Introduction to Panel Data Analysis. Clubmeth Meeting, Erasmus MC, Rotterdam, the Netherlands	2012	1.0
International Conferences		
European Public Health Conference, Glasgow, Scotland	2014	1.0
DEMETRIQ End of Project Conference, Leiden, the Netherlands	2014	0.5
DEMETRIQ Work Package Leader Meeting, Rotterdam, the Netherlands	2014	0.5
DEMETRIQ Policy Evaluation Methodologists Expert Panel Meeting, Liverpool, England	2013	0.5
DEMETRIQ Project Board Meeting, Liverpool, England	2013	0.5

Seminars and Workshops		
Research Seminars of the Department of Public Health, Erasmus MC, Rotterdam, the Netherlands	2012–2016	2.0
Social Epidemiology Methods Meetings, Erasmus MC, Rotterdam, the Netherlands	2013–2014	0.5
Clubmeth Meetings, Erasmus MC, Rotterdam, the Netherlands	2012–2013	0.5
Rotterdam Health Economics Seminars, Erasmus University, Rotterdam, the Netherlands	2013	0.3
Netspar Theme Conference - Health and Inequality Across the Life Cycle, Erasmus University, Rotterdam, the Netherlands	2012	0.3

DEMETRIQ = Developing Methodologies to Reduce Inequalities in the Determinants of Health
Clubmeth = Methodology Club in Public Health