CESifo WORKING PAPERS

9429
2021

November 2021

Health Improvements Impact Income Inequality

Rainer Kotschy

CESifo

Impressum:

CESifo Working Papers
ISSN 2364-1428 (electronic version)
Publisher and distributor: Munich Society for the Promotion of Economic Research - CESifo GmbH
The international platform of Ludwigs-Maximilians University's Center for Economic Studies and the ifo Institute
Poschingerstr. 5, 81679 Munich, Germany
Telephone +49 (0)89 2180-2740, Telefax +49 (0)89 2180-17845, email office@cesifo.de
Editor: Clemens Fuest
https://www.cesifo.org/en/wp
An electronic version of the paper may be downloaded
· from the SSRN website: www.SSRN.com
· from the RePEc website: www.RePEc.org
· from the CESifo website: https://www.cesifo.org/en/wp

CESifo *Working Paper No. 9429*

Health Improvements Impact Income Inequality

Abstract

This paper investigates whether and to what extent long-run trends in population health affected income inequality in the United States over the period 1960-2000. To isolate exogenous variation in health over time, the analysis exploits the sharp decline in cardiovascular disease mortality across states that originated from medical advances in the treatment and prevention of these diseases after 1960. The results demonstrate that health improvements contributed to rising income inequality through mechanisms related to education.

JEL-Codes: I140, I240, J110.

Keywords: population health, aging, Gini coefficient, skill-biased technical change.

Rainer Kotschy
Ludwig Maximilian University of Munich
Geschwister-School-Platz 1
Germany – 80539 Munich
rainer.kotschy@econ.lmu.de

November 2021

1 Introduction

Over the twentieth century medical advances led to a substantial increase in life expectancy in the United States. This rise created an opportunity for increasing human capital investment, high savings, and thus economic growth (Bloom et al., 2003; Cervellati et al., 2017; Hansen and Strulik, 2017; Kotschy, 2021). While there is ample empirical work on the long-run effects of health dynamics on economic development, little research concentrates on the potential consequences of these dynamics for economic inequality.

A link between health and inequality is conceivable through education. Health improvements increase an individual's survival probability at a given age of the life cycle, which raises the expected benefits of education relative to its costs and thus promotes educational attainment and lifetime wealth (Ben-Porath, 1967; Cervellati and Sunde, 2013; Sánchez-Romero et al., 2016; Strulik and Werner, 2016). Higher educational attainment, in turn, widens the income dispersion if technical change is skill-biased (Acemoglu, 1998; Acemoglu and Autor, 2011), if earnings differentials between skilled and unskilled work grow (Eckstein and Nagypal, 2004; Heathcote et al., 2010), or if earnings become an increasingly convex function of educational attainment (Lemieux, 2006). Population aging thanks to increased longevity further reinforces the inequalities in education, employment, and earnings, which tend to accumulate over the life cycle (Ashenfelter and Rouse, 1999; OECD, 2017).

This paper empirically examines whether and to what extent health improvements impacted income inequality in the United States over the period 1960–2000. The analysis exploits variation in health improvements that originated from medical advances in the treatment and prevention of cardiovascular diseases starting in the 1960s. These advances substantially reduced cardiovascular disease mortality and raised life expectancy at adult age. To isolate exogenous variation in health over time, the analysis applies a similar identification strategy as research by Hansen and Strulik (2017) and Kotschy (2021). Specifically, the empirical model combines state-level variation in cardiovascular disease mortality in 1960 with the unanticipated medical advances in the treatment and prevention of these diseases thereafter to predict the mortality decline for each state. The model then uses the predicted mortality decline as instrument for changes in life expectancy to estimate the effect of health on income inequality in a long-difference specification.

The results show that health improvements impacted income inequality: A 1-year increase in life expectancy at 30 raised the Gini coefficient of pre-tax, pre-transfer incomes by 1.8 points. In total, the health improvements between 1960 and 2000 account for an increase of 5.3 points in the income Gini coefficient. For comparison, this coefficient varied between 38 and 53 points over this period with increases of around 10 points in California, Connecticut, New Jersey, and New York. The results suggest that increased college education, an increased ratio of white-collar to blue-collar work, and, to a lesser extent, changes in labor supply were mechanisms through which health affected inequality. This evidence documents that dynamics in population health can be an important determinant of economic inequality and thus extends descriptive work by Deaton and Paxson (1994, 1998). Moreover, the evidence is consistent with theoretical predictions that rising college education has fueled income inequality through skill-biased technical change (Acemoglu, 1998; Acemoglu and Autor, 2011). Finally, the evidence confirms widening earnings differentials between white-collar and blue-collar work (Eckstein and Nagypal, 2004; Heathcote et al., 2010).

2 Assessment of the Effect of Health on Inequality

2.1 Empirical Strategy

To examine the effect of health improvements on economic inequality, I estimate a long-difference specification, which uses one observation in 1960 before medical advances revolutionized the treatment and prevention of cardiovascular diseases (pre treatment) and one observation in 2000 when the subsequent decline in cardiovascular disease mortality started to level off (post treatment). The corresponding outcome equation is given by

$$y_{s,t} = \alpha h_{s,t} + d_t w_s' \beta + \gamma_s + \delta_t + \varepsilon_{s,t}, \tag{1}$$

where $y_{s,t}$ denotes income inequality in state s and year t; $h_{s,t}$ denotes population health, which is measured by log life expectancy; w_s denotes the control variables for initial levels of income, education, and non-cardiovascular disease mortality, which are measured in 1960 and interacted with a post-1960 dummy indicator d_t that uses the year 1960 as reference category; γ_s and δ_t denote state-fixed and time-fixed effects; and $\varepsilon_{s,t}$ denotes the error term. This specification does not include contemporaneous control variables for reasons that will be discussed below.

Conceptually, the outcome equation relates changes in inequality between the pre-treatment and the post-treatment period to changes in log life expectancy between the pre-treatment and the post-treatment period to estimate the effect α of health improvements on economic inequality. Identification of α requires the changes in log life expectancy to be exogenous conditional on the covariates. However, reverse causality and omitted variables may confound this identification.

To address this identification challenge, the analysis applies an instrumental variables approach, which exploits quasi-experimental variation in the sharp decline in cardiovascular disease mortality across states between 1960 and 2000. Over this time, cardiovascular disease mortality declined by around 50 percent, while non-cardiovascular disease mortality stagnated (National Heart, Lung, and Blood Institute, 2012, p. 25). The decline originated from medical advances starting in the 1960s, which arrived through two channels. First, medical innovations introduced novel procedures and pharmaceuticals to treat cardiovascular diseases (CDC, 1999). Examples of these innovations are the cardiac pacemaker, the beta blocker, the portable defibrillator, chest compression, thrombolytic therapy, statins, and intravascular stents. Second, medical research identified the major risk factors for cardiovascular diseases improving diagnosis and prevention of these diseases (CDC, 1999). How much states benefited from these medical advances depended on their initial prevalence of cardiovascular diseases, which differed for cultural, environmental, and social reasons (CDC, 1999).

The empirical model combines the geographic variation in cardiovascular disease mortality with the timing of medical advances to isolate the exogenous variation in log life expectancy over time. Specifically, the model combines state-level variation in cardiovascular disease mortality in 1960 with the unanticipated medical advances in the treatment and prevention of these diseases thereafter to predict the mortality decline for each state. The model then uses the predicted mortality decline as instrument for changes in log life expectancy. The corresponding first stage is given by

$$h_{s,t} = \eta d_t z_s + d_t w_s' \theta + \kappa_s + \lambda_t + \xi_{s,t}, \tag{2}$$

where z_s denotes cardiovascular disease mortality in 1960, which is interacted with a post-1960 indicator d_t that uses the year 1960 as reference category; κ_s and λ_t denote state-fixed and time-fixed effects; and $\xi_{s,t}$ denotes the error term.

The first stage relates differences in the predicted mortality decline between the pre-treatment and post-treatment period to differences in log life expectancy between the pre-treatment and post-treatment period. Hence, it constitutes a differences-in-differences model, which measures the effect η of declining cardiovascular disease mortality on the increase in life expectancy exploiting different treatment intensities across states. In particular, the higher a state's initial cardiovascular disease mortality is, the higher is its predicted mortality decline, and the larger should be its improvement in life expectancy.

The predicted mortality decline is a valid instrument for log life expectancy if it fulfills several conditions. First, the predicted mortality decline must be sufficiently strongly correlated with changes in life expectancy. I assess this assumption with the first-stage F-statistic.

Second, the variation in cardiovascular disease mortality in 1960 must be exogenous conditional on the covariates. Specifically, the baseline specification controls for the initial levels of income, education, and non-cardiovascular disease mortality in addition to state-fixed and time-fixed effects. These covariates capture state-level selection in the adoption of health technologies that originates from initial differences in income, skilled labor supply, health environment, institutions, and culture.[1] Initial distributional differences play no role for the prevalence of cardiovascular disease mortality because socioeconomic status can affect health outcomes only if risk factors are known or access to treatment is available (Phelan and Link, 2013). The control variables are measured in 1960 and interacted with a post-1960 dummy indicator, so they can be incorporated in the empirical model, which corresponds to a first-difference specification because of the state-fixed effects. The empirical model does not include contemporaneous control variables because they can react endogenously to the treatment and would thus be bad controls.

To further undergird the exogeneity assumption, I present results for a placebo test that uses a counterfactual instrument based on geographic variation in infectious disease mortality. Because this decline was already completed at the time cardiovascular disease mortality started to fall, this test provides suggestive evidence that the estimation results for the predicted decline in cardiovascular disease mortality do not merely reflect geographic heterogeneity in development.

Third, conditional on the covariates, the instrument must affect inequality only through the first stage. This assumption is plausible because the instrument is specific to the health channel at the aggregate level. Nevertheless, I use set identification techniques by Conley et al. (2012) to show that small violations of the exclusion restriction do not affect inference with respect to the effect of health on inequality.

Fourth, trends in health and inequality would have been parallel across states in the absence of the medical advances. I provide support for the plausibility of this assumption by comparing pre-treatment trends in health and economic outcomes.

Under these assumptions, the model can be estimated with two-stage least squares.

[1]The empirical model controls for pre-treatment differences in the health environment, which may affect subsequent health dynamics (Aghion et al., 2011; Bloom et al., 2014). A concern in this context is that controlling for initial health might take up treatment variation and thus lead to multicollinearity (Acemoglu and Johnson, 2014). To address this concern, the model controls only for the non-treatment variation in initial health, that is, non-cardiovascular disease mortality. For more information, see the discussion in Kotschy (2021, pp. 6–7).

2.2 Data

I estimate the empirical model for a panel of 48 US states observed every 10 years from 1950 to 2000. Alaska and Hawaii are excluded because of missing data in early periods. The sample is restricted to the white population for reasons of data availability. Table A1 in the Appendix presents descriptive statistics.

Data on incomes, inequality, education, labor supply, and occupational structure are from US censuses (Ruggles et al., 2015) and refer to the working-age population 15–64. Income inequality is measured by the pre-tax, pre-transfer Gini coefficient constructed from inflation-adjusted data on (i) wage incomes or (ii) total incomes (including also social benefits and capital incomes). Values range from 0 (full equality) to 1 (full inequality). The data are collapsed to the state level at which information on health is available. Because this collapsing causes the variables to group at the state level, all regressions are weighted by the population size in 1960. The corresponding population figures are drawn from Hobbs and Stoops (2002).

Data on life expectancy are from US life tables (National Center for Health Statistics, 2017). There is considerable geographic variation in adult life expectancy. For example, life expectancy at age 30 in 1960 differed by almost three years between Nevada (where life expectancy was lowest) and North Dakota (where life expectancy was highest). Kotschy (2021) shows that estimation results can react sensitively to the age at which life expectancy is measured. To address this point, I estimate the model for life expectancy at age 30 and show that the qualitative results also hold for life expectancy at different ages.

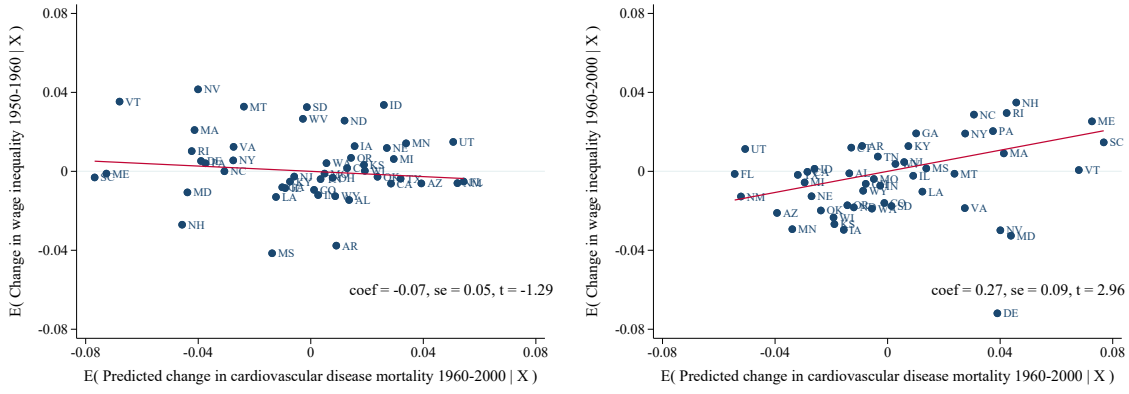
Age-adjusted mortality rates are from Grove and Hetzel (1968). The age adjustment weights the crude age-specific mortality rates with a pre-defined standard population, so that mortality rates can be compared even if states have different demographic structures. The age-adjusted mortality rates are expressed in deaths per 100 white persons.

3 Health Improvements Impact Income Inequality

3.1 The Effect of Health on Inequality

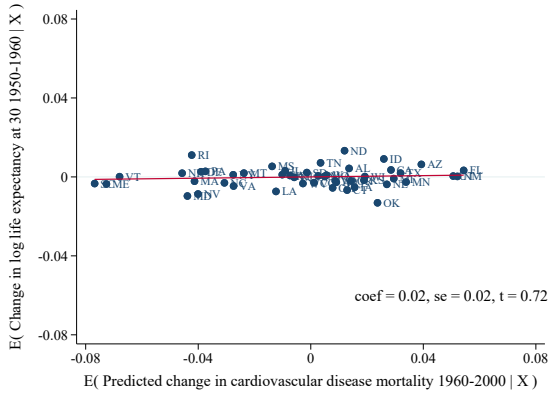
Figure 1 displays partial correlations for wage inequality in the pre-treatment and post-treatment period. Panels (a) and (b) show the reduced form, whereas Panels (c) and (d) show the first stage. Results are obtained from the baseline specification for the years 1950–1960 and 1960–2000. The predicted mortality instrument has no effect on the evolution of inequality and life expectancy before medical advances in the treatment and prevention of cardiovascular diseases arrived in the 1960s. This finding indicates that pre-treatment trends are parallel and thus suggests that the parallel-trends assumption is plausible. In contrast, the predicted mortality decline significantly correlates with increased income inequality and life expectancy post treatment. Similar results obtain for total income inequality—see Figure A1 in the Appendix.

The reduction in cardiovascular disease mortality had a sizable effect on adult life expectancy. Between 1960 and 2000, cardiovascular disease mortality halved from 0.4 to 0.2 deaths per 100 white persons. Combining this reduction with the first-stage estimate of 0.33 implies an increase in life expectancy at 30 of 6.6 percent, or 2.91 years ($\widehat{\Delta h} \approx 0.33 \cdot 0.20 \cdot 44.06 \approx 2.91$, where 44.06 is the mean of life expectancy at 30 in 1960).

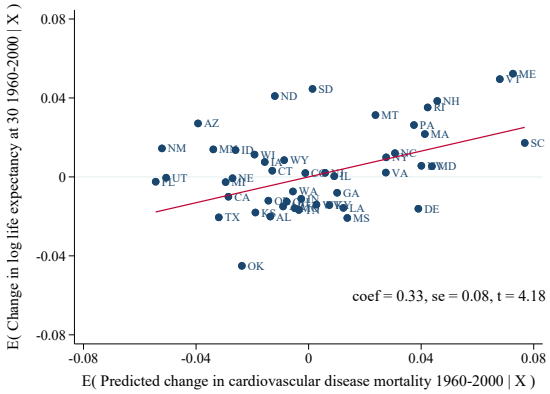


(a) Reduced Form 1950–1960

(b) Reduced Form 1960–2000



(c) First Stage 1950–1960



(d) First Stage 1960–2000

Figure 1: Partial Correlations for Wage Inequality Before and After Treatment

Table 1 reports the estimated effect of health on income inequality. Panel (a) presents the second stage, Panel (b) the reduced form, and Panel (c) the first stage. The dependent variable of the second stage is the wage Gini coefficient in the first three specifications and the total income Gini coefficient in last three specifications. The table presents results for a parsimonious specification that includes only fixed effects (columns 1 and 4), for an extended specification that also controls for initial income and initial education (columns 2 and 5), and for the baseline specification that additionally controls for initial non-cardiovascular disease mortality (columns 3 and 6). Because the dependent variable enters the model linearly, the estimates must be divided by 100 to obtain the marginal effects.

The results demonstrate that health improvements impacted income inequality. According to specification (3), a 1-percent increase in life expectancy at 30 is associated with an increase in inequality of around 0.8 Gini points. Therefore, a 1-year increase in adult life expectancy raised the Gini coefficient by 1.8 points ($\widehat{\Delta y} \approx 0.8 \cdot 1/44.06 \approx 0.018$). In total, the health improvements between 1960 and 2000 account for an increase in the Gini coefficient of around 5.3 points over this period ($\widehat{\Delta y} \approx 0.8 \cdot 0.33 \cdot 0.20 \approx 0.053$). A similar estimate obtains for total income inequality. All second-stage estimates—and likewise all reduced-form and first-stage estimates—are significant at conventional levels. The first-stage F-statistic takes values of around 18, which indicate a strong instrument considering the conventional threshold of 10.

Table 1: Effect of Health on Income Inequality

Inequality in	wage incomes			total incomes		
	(1)	(2)	(3)	(4)	(5)	(6)
(a) Second stage						
Log life expectancy at 30	1.20***	0.79***	0.82***	1.03***	0.86***	0.89***
	(0.30)	(0.25)	(0.27)	(0.29)	(0.25)	(0.26)
(b) Reduced form						
Mortality × Post 1960	0.42**	0.26***	0.27***	0.36**	0.28***	0.28***
	(0.16)	(0.09)	(0.09)	(0.16)	(0.08)	(0.08)
(c) First stage						
Mortality × Post 1960	0.35***	0.33***	0.33***	0.35***	0.33***	0.32***
	(0.09)	(0.07)	(0.08)	(0.09)	(0.07)	(0.07)
First-stage F-statistic	13.9	20.1	17.5	13.9	21.5	18.8
Initial income	—	✓	✓	—	✓	✓
Initial education	—	✓	✓	—	✓	✓
Initial mortality	—	—	✓	—	—	✓
States	48	48	48	48	48	48
Observations	96	96	96	96	96	96

Note: This table reports estimates obtained from two-stage least squares. All regressions include state-fixed and time-fixed effects. In Panels (a) and (b), the dependent variable is the wage Gini coefficient (specifications 1–3) or the total income Gini coefficient (specifications 4–6). In Panel (c), the dependent variable is log life expectancy at 30. Estimates are weighted by the population size in 1960. Control variables are measured in 1960 and interacted with a post-1960 treatment indicator. Clustered standard errors are reported in parentheses. Asterisks indicate significance levels: * $p < 0.1$; ** $p < 0.05$; *** $p < 0.01$.

3.2 Additional Results and Robustness

Panel Results. The results in Table 1 document the long-run effect of health on inequality. Because data on incomes and life expectancy are available for every decade between 1950 and 2000, it is possible to trace how the effect of health on inequality evolved over this period. Table A2 in the Appendix shows the corresponding results that are obtained from a panel model in which cardiovascular disease mortality is interacted with year dummies to instrument for changes in log life expectancy. The reduced-form and first-stage estimates indicate that the effect of health on inequality has gradually accumulated between 1970 and 2000. The implied long-run effect is quantitatively close to the estimate obtained from the long-difference model. Both estimates are larger than estimates obtained from ordinary least squares, which are reported in Table A3 in the Appendix.

Measurement. A potential concern is that the estimated effect of health on inequality may depend on the age at which life expectancy is measured. On the one hand, health improvements may impact economic outcomes only if they take place sufficiently early in life. On the other hand, summary measures, such as life expectancy, only approximate population health at a given age. Table A4 in the Appendix demonstrates that the qualitative finding of a positive effect of health on inequality holds irrespective of the age at which life expectancy is measured. Moreover, the results show that the estimates decrease in magnitude as age increases, which is consistent with the notion that health improvements have larger effects if they take place earlier in life.

Placebo Test. The exogeneity assumption requires that the covariates account for any systematic geographic heterogeneity in development that influences the predicted mortality decline.

Although this assumption is untestable, a placebo test can provide suggestive evidence for its plausibility. If the predicted decline in cardiovascular disease mortality merely reflects geographical heterogeneity in development, a similar pattern should be observed if the predicted mortality decline is constructed based on a different set of diseases, which was not treated between 1960 and 2000. I construct a counterfactual instrument based on infectious disease mortality in 1935. This instrument exploits geographic variation in infectious disease mortality right before the introduction of antibiotics and other pharmaceuticals decimated mortality from pneumonia, influenza, typhoid, tuberculosis, measles, malaria, scarlet fever, diphtheria, whooping cough, and syphilis. By 1960 this decline was completed and infectious disease mortality has stagnated at low levels since then (Cutler et al., 2006, pp. 103–104). Table A5 in the Appendix shows that the predicted decline in infectious disease mortality has no predictive power for the increases in inequality and life expectancy between 1960 and 2000. Moreover, the estimated reduced-form and first-stage effects are quantitatively close to the baseline results when the empirical model includes the predicted declines in both cardiovascular disease and infectious disease mortality.

In addition, the placebo evidence in Table A5 suggests that health improvements at adult age were crucial for the increase in income inequality between 1960 and 2000. While the estimates are significant for the decline in cardiovascular diseases mortality, which predominantly boosted life expectancy at adult age, they are insignificant for the decline in infectious disease mortality, which also boosted life expectancy at child age.

Set Identification. The exclusion assumption requires that the predicted mortality instrument affects inequality only through the first stage. Just as exogeneity, this assumption is untestable. However, it is possible to use set identification techniques to gauge how slight violations of this assumption would affect inference. To this end, I apply the union of confidence intervals approach by Conley et al. (2012) to compute confidence bounds based on a support assumption with respect to the potential effect of the instrument on the dependent variable. Table A6 in the Appendix reports confidence bounds that impose violations of the exclusion restriction of up to 25 percent of the estimated reduced-form effects in the baseline results. Even under these violations, the bounds confirm a positive effect of health and inequality that is significant at the 95-percent level.

3.3 Potential Mechanisms for the Effect of Health on Inequality

What are potential mechanisms for the effect of health on income inequality? Hansen and Strulik (2017) find that improvements in cardiovascular health have raised college enrollment in the United States. Kotschy (2021) documents that these improvements have also shaped life-cycle incomes and that increased college education and changes in labor supply are potential explanations for this effect. Furthermore, Acemoglu (1998) and Acemoglu and Autor (2011) show theoretically that skill-biased technical change boosts income inequality by widening the college wage premium. Finally, Eckstein and Nagypal (2004) and Heathcote et al. (2010) report growing earnings differentials between skilled and unskilled work. Building on these insights, I test if college education, labor supply, and occupational structure are mechanisms for the effect of health on income inequality.

Table 2: Mechanisms: College Education, Labor Supply, and Occupational Structure

Inequality in	wage incomes			total incomes		
	(1)	(2)	(3)	(4)	(5)	(6)
College graduation	0.54**			0.63**		
	(0.25)			(0.27)		
Hours worked per year		0.0008*			0.0007**	
		(0.0004)			(0.0003)	
Ratio of white-collar			0.04***			0.05***
to blue-collar work			(0.02)			(0.02)
First-stage F-statistic	14.4	4.1	19.5	15.4	7.1	21.2
States	48	48	48	48	48	48
Observations	96	96	96	96	96	96
Full controls	✓	✓	✓	✓	✓	✓

Note: This table reports second-stage estimates obtained from two-stage least squares. All regressions include state-fixed and time-fixed effects. The dependent variable is the wage Gini coefficient (specifications 1–3) or the total income Gini coefficient (specifications 4–6). Estimates are weighted by the population size in 1960. All specifications control for initial income, initial education, and initial non-cardiovascular disease mortality. Control variables are measured in 1960 and interacted with a post-1960 treatment indicator. Clustered standard errors are reported in parentheses. Asterisks indicate significance levels: * $p < 0.1$; ** $p < 0.05$; *** $p < 0.01$.

Table 2 reports second-stage results for a model in which health is replaced by either college graduation rates, hours worked per year, or the ratio of white-collar to blue-collar work. All three variables correlate positively and significantly with inequality. A 1-percentage-point increase in college graduation raises wage inequality by one-half of a Gini point. Likewise, increases of 70 hours worked per year or of 12 percentage points in the ratio of white-collar to blue-collar work lead to a similar rise in inequality. Hence, college education, labor supply, and occupational structure are mechanisms through which population health can affect income inequality. However, in the first stage, the instrument correlates less strongly with labor supply than with college education and occupational structure, suggesting that labor supply is less important than the other two mechanisms.[2]

Income growth is another potential mechanism through which health could affect inequality. States that grow faster could undergo larger economic transformations, such that they experience a stronger increase in inequality than states that are growing slower. To test this hypothesis, I estimate a model in which health is replaced by income per working-age person. The estimated correlation is positive but not significant at the 5-percent level—see Table A7 in the Appendix.

4 Conclusion

This paper demonstrates that health improvements have contributed to rising income inequality in the United States over the period 1960–2000. This finding provides novel evidence that long-run trends in population health can be an important determinant of economic inequality. The results suggest college education, occupational structure, and labor supply as mechanisms for the effect of health on inequality. This evidence is consistent with theoretical predictions

[2]These results are also consistent with an interpretation along the lines of agglomeration economies (see, for example, Kline and Moretti, 2013). Health improvements would start a self-reinforcing cycle in which higher supply of skilled labor would raise investment in technologies that rely on skilled labor, which in turn raises local demand for skills and amplifies inequality. In this context, the estimates should be interpreted as reduced-form effects.

that rising college education has fueled income inequality through skill-biased technical change. Moreover, the evidence confirms growing earnings differentials related to the increasing ratio of white-collar to blue-collar work.

Acknowledgments

The author wishes to acknowledge financial support from the Deutsche Forschungsgemeinschaft (German Research Foundation) through DFG Project 395413683. Moreover, the author thanks Uwe Sunde for helpful suggestions and comments.

References

Acemoglu, D. (1998): "Why Do New Technologies Complement Skills? Directed Technical Change and Wage Inequality," *Quarterly Journal of Economics*, 113, 1055–1089.

Acemoglu, D. and D. Autor (2011): "Skills, Tasks, and Technologies: Implications for Employment and Earnings," in *Handbook of Labor Economics*, ed. by D. Card and O. Ashenfelter, North Holland, Amsterdam, vol. 4B, chap. 12, 1043–1171.

Acemoglu, D. and S. Johnson (2014): "Disease and Development: A Reply to Bloom, Canning, and Fink," *Journal of Political Economy*, 122, 1367–1375.

Aghion, P., P. Howitt, and F. Murtin (2011): "The Relationship Between Health and Growth: When Lucas Meets Nelson-Phelps," *Review of Economics and Institutions*, 2, 1–24.

Ashenfelter, O. C. and C. E. Rouse (1999): "The Payoff to Education," *mimeo*, Princeton University, Princeton, NJ.

Ben-Porath, Y. (1967): "The Production of Human Capital and the Life Cycle of Earnings," *Journal of Political Economy*, 75, 352–365.

Bloom, D. E., D. Canning, and G. Fink (2014): "Disease and Development Revisited," *Journal of Political Economy*, 122, 1355–1366.

Bloom, D. E., D. Canning, and B. Graham (2003): "Longevity and Life-Cycle Savings," *Scandinavian Journal of Economics*, 105, 319–338.

CDC (1999): "Decline in Deaths from Heart Disease and Stroke – United States, 1900–1999," *Department of Health and Human Services, Centers for Disease Control and Prevention, Morbidity and Mortality Weekly Report*, 48, 649–656.

Cervellati, M. and U. Sunde (2013): "Life Expectancy, Schooling, and Lifetime Labor Supply: Theory and Evidence Revisited," *Econometrica*, 81, 2055–2086.

Cervellati, M., U. Sunde, and K. F. Zimmermann (2017): "Demographic Dynamics and Long-Run Development: Insights for the Secular Stagnation Debate," *Journal of Population Economics*, 30, 401–432.

Conley, T. G., C. B. Hansen, and P. E. Rossi (2012): "Plausibly Exogenous," *Review of Economics and Statistics*, 94, 260–272.

Cutler, D., A. Deaton, and A. Lleras-Muney (2006): "The Determinants of Mortality," *Journal of Economic Perspectives*, 20, 97–120.

Deaton, A. S. and C. H. Paxson (1994): "Intertemporal Choice and Inequality," *Journal of Political Economy*, 102, 437–467.

——— (1998): "Health, Income, and Inequality over the Life Cycle," in *Frontiers in the Economics of Aging*, ed. by D. A. Wise, Chicago, IL: University of Chicago Press, chap. 10, 431–462.

Eckstein, Z. and E. Nagypal (2004): "The Evolution of US Earnings Inequality: 1961–2002," *Federal Reserve Bank of Minneapolis Quarterly Review*, 28, 10–29.

Grove, R. and A. Hetzel (1968): *Vital Statistics Rates in the United States 1940–1960*, Washington, DC: U.S. Department of Health, Education, and Welfare, Public Health Service.

Hansen, C. W. and H. Strulik (2017): "Life Expectancy and Education: Evidence from the Cardiovascular Revolution," *Journal of Economic Growth*, 22, 421–450.

Heathcote, J., F. Perri, and G. L. Violante (2010): "Unequal We Stand: An Empirical Analysis of Economic Inequality in the United States, 1967–2006," *Review of Economic Dynamics*, 13, 15–51.

Hobbs, F. and N. Stoops (2002): *Demographic Trends in the 20th Century*, Washington, DC: U.S. Census Bureau, Census 2000 Special Reports, Series CENSR-4, U.S. Government Printing Office.

Kline, P. and E. Moretti (2013): "Local Economic Development, Agglomeration Economies, and the Big Push: 100 Years of Evidence from the Tennessee Valley Authority," *Quarterly Journal of Economics*, 129, 275–331.

Kotschy, R. (2021): "Health Dynamics Shape Life-Cycle Incomes," *Journal of Health Economics*, 75, 102398.

Lemieux, T. (2006): "The "Mincer Equation" Thirty Years After Schooling, Experience, and Earnings," in *Jacob Mincer: A Pioneer of Modern Labor Economics*, ed. by S. Grossbard, New York, NY: Springer, chap. 11, 127–145.

National Center for Health Statistics (2017): *Vital Statistics of the United States*, Hyattsville, MD: U.S. Department of Health and Human Services, Centers for Disease Control and Prevention, National Center for Health Statistics, available online at: https://www.cdc.gov/nchs/products/life_tables.htm.

National Heart, Lung, and Blood Institute (2012): *Morbidity and Mortality: 2012 Chart Book on Cardiovascular, Lung, and Blood Diseases*, Bethesda, MD: National Institutes of Health.

OECD (2017): *Preventing Ageing Unequally*, Paris: OECD Publishing.

Phelan, J. C. and B. G. Link (2013): "Fundamental Cause Theory," in *Medical Sociology on the Move: New Directions in Theory*, ed. by W. C. Cockerham, New York, NY: Springer, 105–125.

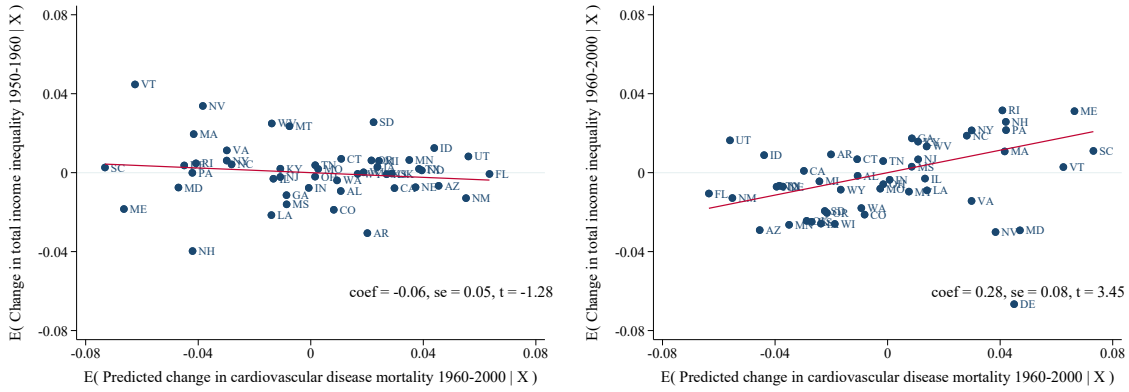
Ruggles, S., K. Genadek, R. Goeken, J. Grover, and M. Sobek (2015): *Integrated Public Use Microdata Series: Version 6.0 [Dataset]*, Minneapolis, MN: University of Minnesota.

Sánchez-Romero, M., A. Prskawetz, and H. d'Albis (2016): "Education, Lifetime Labor Supply, and Longevity Improvements," *Journal of Economic Dynamics and Control*, 73, 118–141.

Strulik, H. and K. Werner (2016): "50 is the New 30 – Long-Run Trends of Schooling and Retirement Explained by Human Aging," *Journal of Economic Growth*, 21, 165–187.

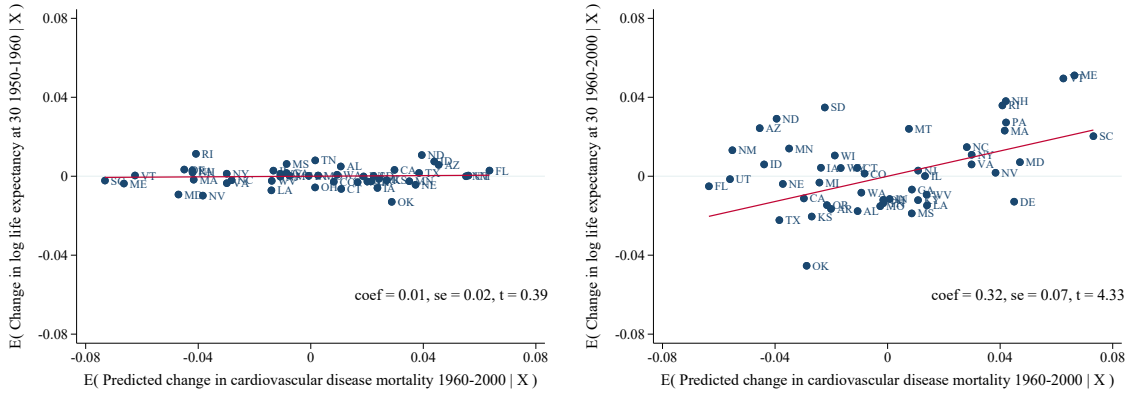
Appendix

Additional Figures



(a) Reduced Form 1950–1960

(b) Reduced Form 1960–2000



(c) First Stage 1950–1960

(d) First Stage 1960–2000

Figure A1: Partial Correlations for Total Income Inequality Before and After Treatment

Additional Tables

Table A1: Descriptive Statistics

Variable (units)	Obs.	Mean	Std. Dev.	Min.	Max.
Income inequality					
Wage Gini in 1960 (on scale 0–1)	48	0.43	0.02	0.39	0.50
Wage Gini in 2000 (on scale 0–1)	48	0.46	0.02	0.44	0.50
Total income Gini in 1960 (on scale 0–1)	48	0.46	0.02	0.42	0.49
Total income Gini in 2000 (on scale 0–1)	48	0.49	0.02	0.46	0.53
Adult health and mortality					
Life expectancy at 30 in 1960 (in years)	48	44.06	0.73	42.33	45.27
Life expectancy at 30 in 2000 (in years)	48	49.14	1.03	47.02	50.83
Log life expectancy at 30 in 1960 (in log points)	48	3.79	0.02	3.75	3.81
Log life expectancy at 30 in 2000 (in log points)	48	3.89	0.02	3.85	3.93
Cardiovascular disease mortality in 1960 (per 100 whites)	48	0.40	0.04	0.31	0.46
Cardiovascular disease mortality in 2000 (per 100 whites)	48	0.20	0.02	0.15	0.25
Non-cardiovascular disease mortality in 1960 (per 100 whites)	48	0.37	0.03	0.34	0.47
Non-cardiovascular disease mortality in 1960 (per 100 whites)	48	0.35	0.02	0.31	0.41
Income, education, labor supply, and occupational structure					
Log wage income per person in 1960 (in log points, wages in USD)	48	10.17	0.14	9.89	10.42
Log total income per person in 1960 (in log points, incomes in USD)	48	10.24	0.13	9.97	10.46
Share college graduation in 1960 (in percent)	48	7.26	1.45	4.61	10.07
Share college graduation in 2000 (in percent)	48	23.41	4.66	13.94	33.80
Hours worked per year in 1960 (in hours)	48	1887.70	51.14	1780.64	2041.93
Hours worked per year in 2000 (in hours)	48	1847.06	46.23	1710.19	1933.47
Ratio of white-collar to blue-collar work in 1960	48	1.34	0.29	0.83	1.86
Ratio of white-collar to blue-collar work in 2000	48	2.46	0.55	1.63	3.70

Note: Descriptive statistics for 48 US states (without Alaska and Hawaii) in the years 1960 and 2000. All numbers refer to the white population. Numbers for inequality, income, education, labor supply, and occupational structure refer to the working-age population aged 15–64.

Table A2: Panel Results: Effect of Health on Income Inequality

Inequality in	wage incomes			total incomes		
	(1)	(2)	(3)	(4)	(5)	(6)
(a) Second stage (2SLS)						
Log life expectancy at 30	1.46***	0.85***	0.88***	1.28***	0.97***	0.99***
	(0.30)	(0.26)	(0.27)	(0.29)	(0.28)	(0.29)
(b) Second stage (LIML)						
Log life expectancy at 30	1.50***	0.88***	0.91***	1.31***	0.98***	1.01***
	(0.33)	(0.28)	(0.29)	(0.32)	(0.29)	(0.31)
(c) Reduced Form						
Mortality × 1950	-0.15***	-0.07	-0.07	-0.12***	-0.06	-0.06
	(0.04)	(0.05)	(0.05)	(0.03)	(0.04)	(0.05)
Mortality × 1970	0.13**	0.03	0.02	0.12*	0.04	0.03
	(0.07)	(0.04)	(0.04)	(0.06)	(0.04)	(0.04)
Mortality × 1980	0.30***	0.16***	0.16***	0.27***	0.17***	0.17***
	(0.08)	(0.04)	(0.04)	(0.08)	(0.04)	(0.04)
Mortality × 1990	0.35***	0.14**	0.14**	0.33***	0.19***	0.20***
	(0.09)	(0.06)	(0.06)	(0.08)	(0.06)	(0.06)
Mortality × 2000	0.42**	0.26***	0.27***	0.36**	0.28***	0.28***
	(0.16)	(0.09)	(0.09)	(0.16)	(0.08)	(0.08)
(d) First Stage						
Mortality × 1950	-0.01	0.02	0.02	-0.01	0.01	0.01
	(0.03)	(0.02)	(0.02)	(0.03)	(0.02)	(0.02)
Mortality × 1970	0.13***	0.12***	0.11***	0.13***	0.12***	0.11***
	(0.04)	(0.03)	(0.04)	(0.04)	(0.03)	(0.03)
Mortality × 1980	0.20***	0.21***	0.20***	0.20***	0.20***	0.19***
	(0.05)	(0.05)	(0.05)	(0.05)	(0.05)	(0.05)
Mortality × 1990	0.28***	0.27***	0.26***	0.28***	0.26***	0.26***
	(0.05)	(0.07)	(0.07)	(0.05)	(0.07)	(0.07)
Mortality × 2000	0.35***	0.33***	0.33***	0.35***	0.33***	0.32***
	(0.09)	(0.07)	(0.08)	(0.09)	(0.07)	(0.07)
First-stage F-statistic	7.5	6.1	5.1	7.5	6.2	5.1
Initial income	—	✓	✓	—	✓	✓
Initial education	—	✓	✓	—	✓	✓
Initial mortality	—	—	✓	—	—	✓
States	48	48	48	48	48	48
Observations	288	288	288	288	288	288

Note: This table reports estimates for ten-year panel data over the time period 1950–2000. All regressions include state-fixed and time-fixed effects. Second-stage results are reported for two-stage least squares in Panel (a) and for heteroskedasticity-robust limited information maximum likelihood (LIML) in Panel (b). In Panels (a) to (c), the dependent variable is the wage Gini coefficient (specifications 1–3) or the total income Gini coefficient (specifications 4–6). In Panel (d), the dependent variable is log life expectancy at 30. Estimates are weighted by the population size in 1960. Control variables are measured in 1960 and interacted with year dummies; the year 1960 serves as reference category, so the corresponding dummy is omitted. Clustered standard errors are reported in parentheses. Asterisks indicate significance levels: * $p < 0.1$; ** $p < 0.05$; *** $p < 0.01$.

Table A3: Ordinary Least Squares: Effect of Health on Income Inequality

Inequality in	wage incomes			total incomes		
	(1)	(2)	(3)	(4)	(5)	(6)
Log life expectancy at 30	1.04***	0.32**	0.32**	0.98***	0.40***	0.40***
	(0.15)	(0.14)	(0.14)	(0.15)	(0.13)	(0.13)
Initial income	—	✓	✓	—	✓	✓
Initial education	—	✓	✓	—	✓	✓
Initial mortality	—	—	✓	—	—	✓
States	48	48	48	48	48	48
Observations	96	96	96	96	96	96

Note: This table reports results obtained from ordinary least squares. All regressions include state-fixed and time-fixed effects. The dependent variable is the wage Gini coefficient (specifications 1–3) or the total income Gini coefficient (specification 4–6). Estimates are weighted by the population size in 1960. Control variables are measured in 1960 and interacted with a post-1960 treatment indicator. Clustered standard errors are reported in parentheses. Asterisks indicate significance levels: * $p < 0.1$; ** $p < 0.05$; *** $p < 0.01$.

Table A4: Robustness: Life Expectancy Measured at Different Ages

	Life expectancy measured at						
	birth	20	30	40	50	60	mean 20–60
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
(a) Wage inequality							
Log life expectancy	1.56***	1.05***	0.82***	0.64***	0.56***	0.39***	0.71***
	(0.53)	(0.36)	(0.27)	(0.21)	(0.20)	(0.12)	(0.23)
First-stage F-statistic	12.8	17.4	17.5	17.7	18.4	19.5	20.1
States	48	48	48	48	48	48	48
Observations	96	96	96	96	96	96	96
Full controls	✓	✓	✓	✓	✓	✓	✓
(b) Total income inequality							
Log life expectancy	1.74***	1.17***	0.89***	0.69***	0.54***	0.42***	0.63***
	(0.55)	(0.36)	(0.26)	(0.20)	(0.17)	(0.11)	(0.22)
First-stage F-statistic	13.3	17.9	18.8	19.1	27.6	22.1	20.1
States	48	48	48	48	48	48	48
Observations	96	96	96	96	96	96	96
Full controls	✓	✓	✓	✓	✓	✓	✓

Note: This table reports second-stage estimates obtained from two-stage least squares. All regressions include state-fixed and time-fixed effects. The dependent variable is the wage Gini coefficient in Panel (a) or the total income Gini coefficient in Panel (b). Estimates are weighted by the population size in 1960. All specifications control for initial income, initial education, and initial non-cardiovascular disease mortality. Control variables are measured in 1960 and interacted with a post-1960 treatment indicator. Clustered standard errors are reported in parentheses. Asterisks indicate significance levels: * $p < 0.1$; ** $p < 0.05$; *** $p < 0.01$.

Table A5: Placebo Test: Infectious Diseases versus Cardiovascular Diseases

Inequality in	wage incomes		total incomes	
	(1)	(2)	(3)	(4)
(a) Reduced form				
Infectious disease mortality × Post 1960	-0.02	0.05	-0.08	0.00
	(0.06)	(0.06)	(0.06)	(0.05)
Cardiovascular disease mortality × Post 1960	—	0.27***	—	0.28***
	—	(0.10)	—	(0.09)
(b) First stage				
Infectious disease mortality × Post 1960	-0.16	-0.09	-0.16	-0.07
	(0.10)	(0.10)	(0.10)	(0.10)
Cardiovascular disease mortality × Post 1960	—	0.31***	—	0.30***
	—	(0.07)	—	(0.07)
States	48	48	48	48
Observations	96	96	96	96
Controls	✓	✓	✓	✓

Note: This table reports reduced-form and first-stage estimates obtained from ordinary least squares. All regressions include state-fixed and time-fixed effects. The dependent variable is the wage Gini coefficient (specifications 1–3) or the total income Gini coefficient (specifications 4–6). Estimates are weighted by the population size in 1960. Infectious disease mortality is measured in 1935 and interacted with a post-1960 treatment indicator. Controls for education and income are measured in 1960 and interacted with a post-1960 treatment indicator. Clustered standard errors are reported in parentheses. Asterisks indicate significance levels: * $p < 0.1$; ** $p < 0.05$; *** $p < 0.01$.

Table A6: Set Identification: Relaxing the Exclusion Restriction

	Assumed portion of reduced-form effect that instrument has on dependent variable				
	5%	10%	15%	20%	25%
	(1)	(2)	(3)	(4)	(5)
(a) Wage inequality					
Confidence bounds	[0.23, 1.19]	[0.20, 1.24]	[0.17, 1.28]	[0.13, 1.32]	[0.10, 1.36]
(b) Total income inequality					
Confidence bounds	[0.29, 1.22]	[0.26, 1.26]	[0.23, 1.31]	[0.19, 1.35]	[0.16, 1.40]
States	48	48	48	48	48
Observations	96	96	96	96	96
Full controls	✓	✓	✓	✓	✓

Note: This table reports 95-percent confidence bounds for long-difference specifications with two observations, one in 1960 and one in 2000. Confidence bounds are computed from the union of confidence intervals following Conley et al. (2012). The specifications impose violations of the exclusion restriction of 5 up to 25 percent of the reduced-form effect. All regressions include state-fixed and time-fixed effects. The dependent variable is is the wage Gini coefficient in Panel (a) or the total income Gini coefficient in Panel (b). Estimates are weighted by the population size in 1960. All specifications control for initial income, initial education, and initial non-cardiovascular disease mortality. Control variables are measured in 1960 and interacted with a post-1960 treatment indicator. Standard errors are clustered at the state level.

Table A7: Mechanism: Income Growth

Inequality in	wage incomes	total incomes
	(1)	(2)
Income per working-age person	0.43*	0.27*
	(0.22)	(0.14)
First-stage F-statistic	4.9	8.9
States	48	48
Observations	96	96
Full controls	✓	✓

Note: This table reports second-stage estimates obtained from two-stage least squares. All regressions include state-fixed and time-fixed effects. The dependent variable is the wage Gini coefficient (specification 1) or the total income Gini coefficient (specification 2). Estimates are weighted by the population size in 1960. All specifications control for initial income, initial education, and initial non-cardiovascular disease mortality. Control variables are measured in 1960 and interacted with a post-1960 treatment indicator. Clustered standard errors are reported in parentheses. Asterisks indicate significance levels: * $p < 0.1$; ** $p < 0.05$; *** $p < 0.01$.