



Missing Poor in the U.S.

Mathieu Lefebvre¹ · Pierre Pestieau² · Gregory Ponthiere³

Received: 15 September 2023 / Accepted: 20 February 2024

© The Author(s), under exclusive licence to Springer Science+Business Media, LLC, part of Springer Nature 2024

Abstract

Given that poor individuals face worse survival conditions than non-poor individuals, one can expect that a steeper income gradient in mortality leads, through stronger income-based selection, to a lower poverty rate at the old age (i.e. the "missing poor" hypothesis). This paper uses U.S. state-level data on poverty at age 65+ and life expectancy by income levels to provide an empirical test of the missing poor hypothesis. Using average temperature as an instrument for mortality differentials, we show that instrumented changes in mortality differentials have a negative and statistically significant effect on old-age poverty: a 1 % increase in the mortality differential implies a 16 % decrease in the 65+ headcount poverty rate. Using those regression results, we compute hypothetical old-age poverty rates while neutralizing the impact of the income gradient in mortality, and show that correcting for heterogeneity in income-based selection effects modifies the comparison of old-age poverty prevalence across states.

Keywords Poverty · Measurement · Income gradient in mortality · Selection biases · Comparability

1 Introduction

The income gradient in mortality is a widely established stylized fact: empirical studies show, for various countries and epochs, that lower incomes are statistically related with higher mortality risks.¹ Those studies show that, although the strength of the income gradient in mortality varies across gender (it is stronger for men than for women) and also across

¹ See Pamuk (1985, 1988); Duleep (1986); Deaton and Paxson (1998); Backlund et al. (1999); Deaton (2003); Jusot (2004); Bossuyt et al. (2004); Van Oyen et al. (2005); Duggan et al. (2007); Cristia (2009); Salm (2012); Belloni et al. (2013); Kalwij et al. (2013); Chetty et al. (2016) and Milligan and Schirle (2018). One exception is Snyder and Evans (2006), who find a negative correlation between income and survival.

✉ Mathieu Lefebvre
mathieu-julien.lefebvre@univ-amu.fr

¹ Aix Marseille University, CNRS, AMSE, 5-9 Boulevard Maurice Bourdet, CS 50498, 13205 Marseille, France

² Université de Liège, CORE and PSE, 7, Bd Rectorat (B31), 4000 Liège, Belgium

³ UCLouvain, Hoover Chair in Economic and Social Ethics, Collège Dupriez, 3 Place Montesquieu, 1348 Louvain-La-Neuve, Belgium

countries, it remains true that poorer individuals have, on average, shorter lives than richer individuals.²

What are the implications of the income gradient in mortality for poverty measurement? In a pioneer work, Kanbur and Mukherjee (2007) argued that standard poverty measures tend, under income-differentiated mortality, to underestimate the actual extent of the poverty phenomenon. Actually, given that poor individuals face, on average, worse survival conditions than non-poor individuals, measures of old-age poverty are subject to a selection bias: these measure poverty only within the surviving population, which includes, due to excess mortality of the poor, a relatively smaller proportion of poor individuals.

Under income-differentiated mortality, low levels of measured old-age poverty may not be caused by a better situation of the old, but may be due to a steeper income gradient in mortality, leading to a stronger income-based selection. Making an analogy with Sen (1998) expression of "missing women", we can coin this hypothesis the "missing poor" hypothesis. According to that hypothesis, a steeper income gradient in mortality leads, through stronger income-based selection, to a lower poverty rate at the old age.

The missing poor hypothesis has important consequences for the comparison of old-age poverty measures across countries. When the strength of the income gradient in mortality varies across space, the missing poor phenomenon questions the comparability of old-age poverty measures across countries. If the missing poor hypothesis is correct, observed gaps in standard poverty measures may be due not to "true" differences in the prevalence of poverty, but may be due to income-based selection processes of unequal strengths across space. Lefebvre et al. (2018) stress that the measurement problems raised by income-based selection cannot be solved by adding another dimension, such as life expectancy, and shifting from one-dimensional poverty measures to two-dimensional poverty measures. Indeed, multidimensional poverty indicators are obtained for a given population, and, as such, are also subject to the selection biases induced by income-differentiated mortality. Lefebvre et al. (2018) also show that poverty measures that are sensitive to the income distribution (such as FGT measures) are impacted by the income gradient in mortality.

Whereas the existence of an income gradient in mortality is strongly established, the missing poor hypothesis has not yet been tested empirically. Actually, the mere existence of an income gradient in mortality does not necessarily imply that poverty measures are biased downwards. True, the income gradient in mortality leads to some selection effects, but the final impact of those selection effects on poverty is not obvious, since selection effects induce general equilibrium effects (e.g. variations in wages) that may make poverty measures take, in theory, either lower or higher levels than in the absence of selection effects.³

The goal of this paper is to test empirically the missing poor hypothesis, that is, to investigate the effect of the mortality differential between the poor and the non-poor on measures of old-age poverty. For that purpose, we use U.S. state-level data on old-age headcount poverty (age 65 and more) and on life expectancy at age 40 by income quartile for the years 2001 to 2014, to test the missing poor hypothesis, that is, to examine whether higher mortality differentials between the poor and the non-poor tend to reduce measured old-age poverty.⁴

² On the variation of the income gradient in mortality across gender and across European countries, see Lefebvre et al. (2018) and Lefebvre et al. (2019a).

³ See Lefebvre et al. (2019a) on how general equilibrium effects due to selection (i.e. excess mortality of the poor) can affect poverty measurement in an overlapping generations economy.

⁴ We use the headcount poverty index because it is the measure that is available at the US state and county levels. We could rely on other poverty measures if available. As stressed above, other measures are also affected by the missing poor hypothesis.

At this stage, it should be stressed that the empirical test of the missing poor hypothesis raises several difficulties. From an empirical perspective, the estimation of a relation between poverty and mortality differentials may not necessarily validate or invalidate the missing poor hypothesis. A first estimation problem may come from reverse causation, i.e. a higher prevalence of poverty leading to stronger mortality differentials between the poor and the non-poor. Reverse causation from poverty to mortality differentials is plausible in theory, and cannot be excluded *a priori*.⁵ Besides reverse causation, another estimation problem may come from omitted variable biases: the estimated relation between poverty and mortality differentials may be due to a third variable affecting both mortality differentials and poverty measures, so that the observed relation would not suffice to validate the missing poor hypothesis.⁶

In order to avoid drawing fallacious conclusions on the impact of mortality differential on poverty measurement, we adopt an empirical strategy that uses an instrumental variable for mortality differentials. For that purpose, we use state-specific data on weather that is assumed to be not affected by poverty measurement, but which affects mortality differentials within the population. More specifically, we use the average temperature as measured annually for each state as a source of exogenous variation of the mortality differential. Temperature is a plausible instrument for income-differentiated mortality in the US where changes in temperature have not been shown to have discernable effects on income growth (Gallup et al., 1999; Dell et al., 2009; Dell et al., 2012; Dell et al., 2014; Miguel et al., 2004; Hsiang, 2010; Barrios et al., 2015). However, it has been shown that there is a temperature-mortality relationship (Curriero et al., 2002; Deschenes and Greenstone, 2011; Deschenes and Moretti, 2009; Barreca, 2012), that self-protection expenditures are unequally distributed and mortality from waves of heat is higher in low income population (Barreca et al., 2016). However we cannot exclude that interstate mobility, in particular mobility from cold to warm states, is related to individual income or that temperature changes affect directly incomes through agricultural productivity. We thus conduct a series of robustness checks to confirm our results.

Anticipating on our results, we show that, when we carry out simple OLS regressions of mortality differentials on old-age poverty, the mortality differential has a negative and statistically significant effect on measured old-age poverty. Moreover, when we instrument for the mortality differential variable by means of a temperature indicator, we find that the instrumented changes in mortality differentials have a negative and statistically significant effect on 65+ headcount poverty rate: a 1 % increase in the mortality differential leads to a 16 % decrease in the 65+ headcount poverty rate. Those results, which are robust to various specifications concerning control variables but also to cross-sectional county-specific data, provide some empirical support for the missing poor hypothesis, in the sense that, in line with that hypothesis, a steeper income gradient in mortality lead to reduce the measured old-age poverty.

Empirical support for the missing poor hypothesis has implications concerning the comparability of old-age poverty measures across states. Indeed, given that the income-mortality gradient has a negative and significant effect on old-age poverty measures, one can question the meaningfulness of comparing old-age poverty rates across states with unequal income-mortality gradients. This limited comparability of old-age poverty rates across states motivates us to use our regression results to calculate hypothetical old-age poverty rates while neutralizing the impact of the income gradient in mortality on the measurement of old-age

⁵ For instance, in the presence of capacity constraints, a larger prevalence of poverty can reduce the capacity of pro-poor programs to achieve their goals, leading, *in fine*, to larger mortality differentials between the poor and the non-poor.

⁶ For instance, unemployment, by reducing accidents at work, may reduce mortality differentials, while, at the same time, contribute to increase poverty.

poverty. We show that correcting for income-based selection effects tends to increase old-age poverty rates by between 1 and 3 percentage points. It is also shown that the comparison of old-age poverty across states is affected by heterogeneity in the income gradient in mortality across states. For instance, on the basis of standard headcount poverty rates, old-age poverty is larger in Massachusetts than in Maryland. However, once the missing poor bias is corrected, the ranking is reversed, and old-age poverty is lower in Massachusetts than in Maryland. Similar reversals arise for other states (e.g. New Jersey vs Nevada, Wisconsin vs Wyoming, etc.), suggesting that the missing poor bias affects the comparison of old-age poverty across U.S. states.

This paper is related to the literature on the "missing poor" problem, such as Kanbur and Mukherjee (2007) and Lefebvre et al. (2013, 2018, 2019a). Those papers are mainly theoretical, and, when turning to data, propose to correct poverty measures for selection biases by constructing counterfactuals: hypothetical income distributions are constructed while assuming that all individuals enjoy the survival conditions of the top income class. Lefebvre et al. (2018) show how extending income profiles of the prematurely dead affects the measurement of poverty when poverty aversion is taken into account (i.e. FGT poverty measures) and Lefebvre et al. (2019a) look at the effect of income mobility on the missing poor hypothesis. This paper complements those studies by adopting an econometric approach, whose goal is to provide empirical tests of the missing poor hypothesis. Another related literature is the one on the income gradient in mortality, to which we referred above. Whereas that literature is descriptive, it documents a relation that is shown, in this paper, to have strong implications for poverty measurement. Finally, this paper is also related to the literature on mortality and development, such as Acemoglu and Johnson (2007); Hazan (2009) and Cervelatti and Sunde (2011). While those papers examine the link between life expectancy and GDP per capita, this paper studies another relation, between life expectancy differentials and poverty.

This paper is organized as follows. Section 2 develops a model motivating our empirical explorations of the impact of mortality differentials on old-age poverty measurement. Section 3 presents the data. The estimation framework and the OLS estimates are presented in Section 4. The IV approach is developed in Section 5. Section 6 corrects old-age poverty measures for the missing poor effect quantified in Section 5. Section 7 examines the robustness of our results to alternative specifications. Section 8 explores potential mechanisms at work behind the missing poor phenomenon. Section 9 concludes.

2 Motivating theory

To frame the empirical analysis, this section develops a simple model describing how mortality differentials between the poor and the non-poor contribute to affect the measurement of poverty at the old age.

Let us consider an economy in which each cohort of individuals is a continuum of size 1.⁷ For simplicity, the human lifespan is composed of two periods. During the first period (young age), individuals earn some income level y^y . Survival to the second period (old age) arises with a probability $0 < \pi < 1$ that depends on the income level when being young. In case of survival, individuals enjoy an income y^o .

For the sake of simplicity, we assume that there exist two income levels, $\check{y}$ and $\hat{y}$, as well as a poverty line $z > 0$, with $\check{y} < z < \hat{y}$.

⁷ We make this hypothesis to use the law of large numbers.

At the young age, a fraction $p > 0$ of individuals has income $y^y = \check{y}$, and thus lies in poverty, whereas a fraction $1 - p$ has income $y^y = \hat{y}$, and thus escapes from poverty.

To keep the analysis simple, the survival rate to the old age takes only two levels, which depend on whether the income enjoyed at the young age is below or above the poverty line z :

$$\pi = \begin{cases} \check{\pi} & \text{if } y^y \leq z \\ \hat{\pi} & \text{if } y^y > z \end{cases}$$

We define the mortality differential as:

$$d \equiv \frac{\hat{\pi}}{\check{\pi}} > 1$$

The parameter d captures the strength of the income gradient in mortality. When the income-based mortality differential is low, d is close to 1, whereas when there is a strong mortality differential, d takes higher values.

There is some degree of income mobility between the young age and the old age. The mobility matrix M , defined conditionally on survival, is given by:

$$M \equiv \begin{pmatrix} \alpha & 1 - \alpha \\ \beta & 1 - \beta \end{pmatrix}$$

where $0 < \alpha < 1$ is the probability that an individual who is in poverty at the young age remains in poverty at the old age, whereas $0 < \beta < 1$ is the probability that an individual who is not poor at the young age falls into poverty at the old age. Given some hysteresis in income levels, the elements along the diagonal are higher than those outside the diagonal, implying that $\alpha > \beta$.

In our economy, the old-age headcount poverty rate H can be written as:

$$H = \frac{\alpha \check{\pi} p + \beta \hat{\pi} (1 - p)}{\check{\pi} p + \hat{\pi} (1 - p)} \quad (1)$$

Multiplying the numerator and the denominator by $\frac{1}{(1-p)\check{\pi}}$, the old-age headcount poverty rate can be rewritten as:

$$H = \frac{\frac{p}{1-p} \alpha + d \beta}{\frac{p}{1-p} + d} \quad (2)$$

This formula basically says that the old-age headcount poverty rate depends on (1) poverty at the young age (p); (2) income mobility (α and β); (3) the mortality differential between the non-poor and the poor (d).

Assuming, as an approximation, that the probability of falling into poverty at the old age β is low, and acknowledging that poverty at the young age is not a widespread phenomenon (i.e. p is low), so that $\frac{p}{1-p} \ll 1 < d$, one can, as a proxy, rewrite the old-age headcount poverty rate as:

$$H \approx \left(\frac{p}{1-p} \right) \frac{\alpha}{d} \quad (3)$$

Taking logs, we can rewrite the old-age headcount poverty rate as:

$$\log(H) \approx \underbrace{\log\left(\frac{p}{1-p}\right)}_{\text{initial poverty}} + \underbrace{\log(\alpha)}_{\text{effect of income mobility}} - \underbrace{\log(d)}_{\text{effect of mortality differential}} \quad (4)$$

This model suggests that, in theory, there are reasons to believe that a larger mortality differential d between the poor and the non-poor leads to lower measured old-age poverty, in line with the missing poor hypothesis. Having stressed this, it should be emphasized that this simple model abstracts from several forces that may be at work in real economies, and may invalidate the missing poor hypothesis. For instance, in real world economies, mortality differentials may affect the propensity to save, and, hence, capital accumulation, leading to changes in wages and in income levels. Thus, although this model points to a possible impact of mortality differentials on poverty measures, one needs to test empirically the missing poor hypothesis. This is the task of the next sections. For that purpose, our empirical strategy below is to estimate equations that are close to Eq. 4, and to test empirically whether the mortality differential has a statistically significant impact on the old-age poverty headcount ratio.

3 Data and descriptive statistics

We use recent estimates of the life expectancy by income across U.S. states, as released by the Health Inequality Project Chetty et al. (2016). For each state, we have access to data of life expectancy computed at age 40 and by quartile of income, for men and women separately. As shown on Fig. 1, men in the fourth quartile of income live on average 10 years longer than those in the first quartile, whereas the equivalent difference for women is 5 years. Note also that the first quartile life expectancy has, over the last 15 years, evolved in a smoother manner

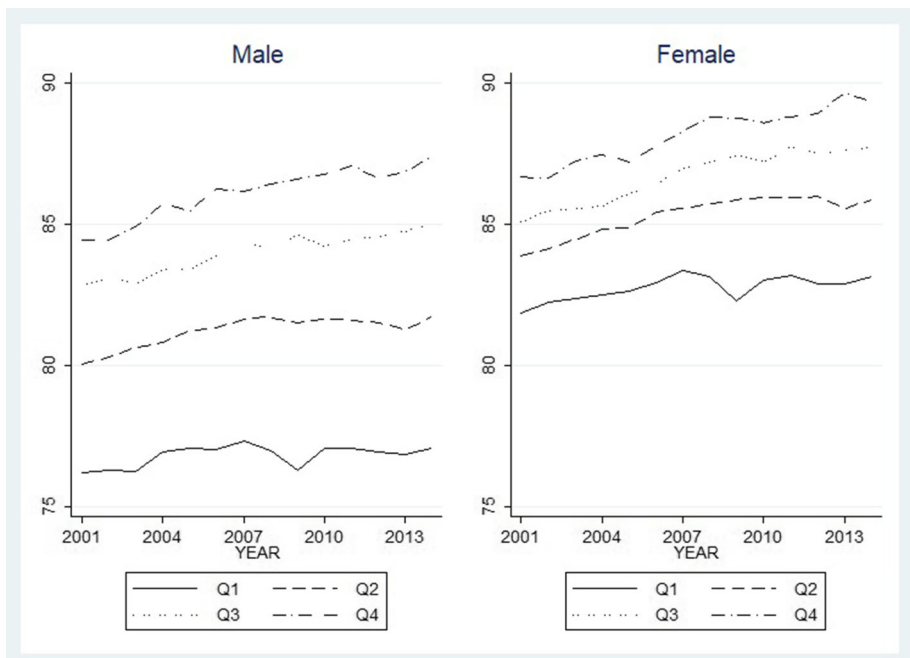


Fig. 1 Life expectancy at age 40 by income quartile (years), males and females, U.S., 2001–2014

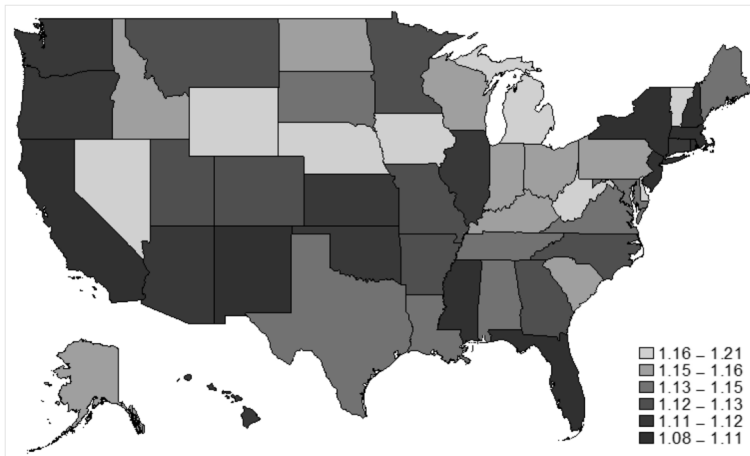


Fig. 2 Mortality differential in the U.S., males, 2014. Notes: Mortality differential is the ratio between the life expectancy at age 40 in the fourth and the first quartiles

than the life expectancy associated with the last income quartiles. The income gradient in mortality has become steeper over the last decade.

These estimates of life expectancy by income quartile are used to define the mortality differential variable as the ratio between the life expectancy in the fourth and the first quartiles. This indicator can be regarded as a measure of the strength of the income gradient in mortality. The larger that indicator is, the larger is the differential in survival conditions between the non-poor and the poor. Figures 2 and 3 show, respectively for men and women, mortality differentials across U.S. states in 2014. We observe a large heterogeneity, both for men and women, across U.S. states, with a lower differential on the West coast. The mortality differential is larger for men than women.

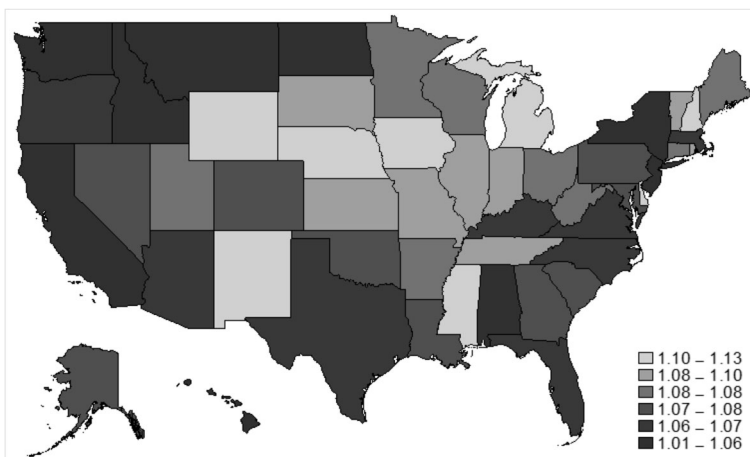


Fig. 3 Mortality differential in the U.S., females, 2014. Notes: Mortality differential is the ratio between the life expectancy at age 40 in the fourth and the first quartiles

Our aim is to relate the mortality differential to the level of poverty in each state. Data on poverty are obtained from the Current Population Survey of the U.S. Census Bureau. We use the standard headcount poverty measure, i.e., the percentage of population in poverty at the state level. The official poverty threshold is a measure of need that does not vary across states, but that is updated for inflation using the Consumer Price Index (CPI-U). The official poverty definition uses money income before taxes and does not include capital gains or non-cash benefits (such as public housing, Medicaid, and food stamps).

Given that we are interested in estimating the impact of mortality differentials on poverty measurement, we use headcount poverty rates at age 65+. These data are not available by sex. Figure 4 shows the headcount poverty rate for population aged 65+ by states for the U.S., in 2014. There exists a large variation in old-age poverty across U.S. states. Poverty of the elderly is, in general, larger in the South than in the North, but with some exceptions.

How can one interpret old-age poverty rates varying across U.S. states? A standard interpretation consists of stating that if a state exhibits a larger old-age poverty rate than another state, this accounts for a larger prevalence of old-age poverty in the former state than in the latter. However, this interpretation assumes *full comparability* of poverty rates across U.S. states, a questionable assumption. Actually, Figs. 2 and 3 show that the U.S. states are heterogeneous in the strength of the income gradient in mortality. Some states are characterized by a steep income gradient in mortality, which implies that the population reaching old age is strongly selected in terms of income. On the contrary, other states exhibit a weaker income/mortality relationship, implying smaller selection effects. This heterogeneity in the strength of selection effects across states creates interferences in the comparison of old-age poverty rates, which may limit the comparability of those measures.

One may wonder to what extent the differences in terms of old-age poverty rates on Fig. 4 are due to selection effects varying across states. If the missing poor hypothesis is correct, old-age poverty rates on Fig. 4 are not comparable across states. In states with higher mortality differentials between the poor and the non-poor, selection effects are stronger, and tend to

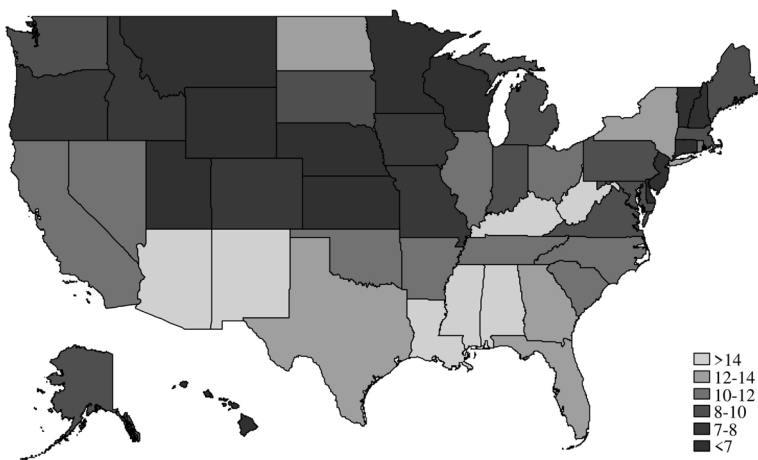


Fig. 4 Headcount poverty rate (%), age 65+, United States, 2014

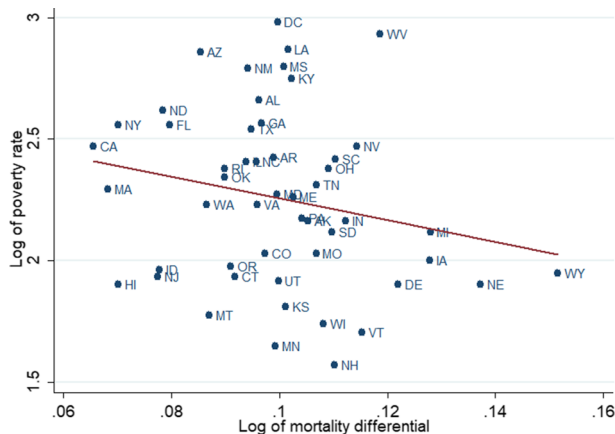


Fig. 5 Poverty and mortality differential, 2014. Notes: Mortality differential is the ratio between the life expectancy at age 40 in the fourth and the first quartiles

reduce old-age poverty rates, whereas, in states with lower mortality differentials, selection is weaker, which pushes old-age poverty rates up.

To compare poverty rates across states in a meaningful way, one must first test the missing poor hypothesis. As a starting point, Fig. 5 plots, for U.S. states, the old-age poverty rate (65+) against the measure of mortality differential in 2014.⁸ Figure 5 shows that there exists a decreasing relationship between the logarithm of the old-age poverty rate and the logarithm of the mortality differential.⁹ States with a higher mortality differential tend to exhibit a lower old-age poverty rate, in line with the missing poor hypothesis.

However, one should be cautious at this stage. First, Fig. 5 plots a relation between old-age poverty and mortality differential without any control. The plotted relation may be driven by a third variable, which impacts both old-age poverty and mortality differentials. Second, Fig. 5 does not allow us to claim that mortality differential has a *causal effect* on poverty. One cannot exclude the possibility of some reverse causality. For instance, under capacity constraints in pro-poor programs, higher poverty rates, by leading to more congestion, reduce the capacity of those programs to achieve their goals, and, hence, lead to higher mortality differentials. In that case, the decreasing relationship on Fig. 5 would be driven by the impact of poverty on mortality differentials, without any link with the missing poor hypothesis.

In order to tackle these issues, we will, in the rest of this paper, carry out econometric regressions of old-age headcount poverty rate on our indicator of mortality differential, by controlling for a series of covariates that may also explain the variation of poverty across states and years. This is done using OLS estimates (Section 4) as well as an instrumental variable approach (Section 5).

⁸ Since we only observe the poverty rate for men and women together, we rely on an aggregate measure of the mortality differential that corresponds to the weighted average of gender-specific differential, where the weights are the relative population size.

⁹ We use log values of our variables of interest because it makes the link with our theoretical model. On a more practical ground, it also allows to avoid that the results are too much impacted by extreme values. The same exercise is carried out for each year separately in the Appendix. Interestingly, the clear decreasing pattern is not observed for every year, making necessary the regression analysis where we control for various confounding factors.

4 OLS estimates

In order to test the missing poor hypothesis, that is, to test whether income-based mortality differentials affect old-age poverty, we use a state panel data model, controlling for fixed state characteristics and state differential in terms of longevity. In particular, we estimate the equation:

$$H_{st} = \beta MD_{st} + X'_{st}\gamma + \alpha_s + \delta_t + \varepsilon_{st} \quad (5)$$

where H_{st} is the log of the old-age headcount poverty rate in state s in year t , that is, the proportion of individuals (aged 65+) in state s in year t who have income below the poverty threshold, MD_{st} is the log of mortality differential and α_s and δ_t are state and year fixed effects. X_{st} is a set of control variables that we detailed below and ε_{st} is the idiosyncratic error term.

As control variables, we include the log of per capita GDP expressed in 2009 dollars, available at the Bureau of Economic Analysis for each state and year, and the unemployment rate obtained from the Bureau of Labor Statistics. It corresponds to the annual average. Both variables allow us to control for the state-specific economic circumstances and the macroeconomic cycle. We also include the log of state welfare spending per capita, as calculated by the Urban Institute-Brookings Institution Tax Policy Center, to account for the size of public services. Finally, we introduce the share of the state population aged 65 and more, the percentage of black population and an indicator of income inequality (the Gini index) obtained from the US Census Bureau.

Table 1 reports OLS regressions as presented in Eq. 5. All regressions are weighted using state population in the state-year cell and we use **clustered-robust the standard errors** at the state level. We present here the results using the weighted mortality differential as the main explaining variable.¹⁰

Column (1) of Table 1 shows the results of the base model where we observe a clear and significant relationship between the mortality differential and the level of old-age poverty. The negative coefficient shows that when the income gradient in mortality is stronger, the level of old-age poverty is reduced. This result is not much affected by the introduction of a series of covariates in column (2) to (4). The coefficient is between 1.5 and 2, depending on the econometric specification. When the differential increases by 1%, the level of poverty is reduced by 1.5% to 2% depending on specifications.

Although our panel approach allows us to control for state-specific characteristics and time-varying factors, the causal interpretation of the relationship between mortality differentials and old-age poverty rates can be questioned on the grounds of endogeneity issues. As we stressed in Section 1, it is indeed still possible that other variables which are varying across time and states are both related with the poverty rate and the life expectancy. It may also be the case that the direction of causation is reversed, and that it is the level of old-age poverty that explains the mortality differential. Thus the results of Table 1 do not allow us to conclude of any causal effect of the mortality differential on the level of measured old-age poverty. The goal of the next section is to develop an empirical strategy allowing us to identify and quantify the causal impact of mortality differential on poverty measurement.¹¹

¹⁰ Regressions based on the gender-specific differentials are presented in the appendix. Results are qualitatively similar.

¹¹ In order to ensure the robustness of our mean-based analysis and despite the fact that our model is not a high-dimensional one, we have also estimated a Lasso regression that adds the absolute value of magnitude of the coefficient as a penalty term to the loss function. This means that the coefficients can be "penalized" or shrunk to zero. This method is often used as a covariates selection method when there are numerous candidates.

Table 1 Results of OLS regressions

	(1) Basic	(2) + Controlling for Macroeconomic factors	(3) + Controlling for demographic factors	(4) + Controlling for welfare exp. and inc. distr.
Mortality differential	-2.114*** (1.036)	-1.702* (0.988)	-1.554* (0.820)	-1.536** (0.749)
GDP per cap.		-0.301 (0.239)	-0.311 (0.247)	-0.326* (0.273)
Unemployment rate		-0.008 (0.008)	-0.008 (0.007)	-0.008 (0.007)
%65+			0.005 (0.020)	-0.001 (0.018)
% black			-0.021 (0.026)	-0.020 (0.027)
Inequality				1.413** (0.836)
Welfare exp. per cap.				0.008 (0.193)
Constant	2.743*** (0.090)	5.906*** (2.524)	7.019*** (2.722)	6.544*** (2.774)
N	696	696	696	696
Adj. R-squared	0.587	0.590	0.590	0.592

Note: Clustered-robust standard errors at the state level are reported in parentheses. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$

5 IV estimates

In order to identify causal effects from mortality differentials to old-age poverty measures, we are completing the results obtained from OLS regressions by additional estimations relying on an instrumental variable, that is, a variable that is unrelated with our dependent variable but related with our explanatory variable. In our panel context, the challenge is to find a source of exogenous variation of the mortality differential across states.

In this section, we propose to exploit the yearly average temperature measured at the state level as a source of exogenous variation of the mortality differential. The underlying intuition is that the temperature level, by hitting poor and non-poor individuals in an asymmetric manner, is likely to impact the mortality differential, without affecting the old-age poverty rate. We collected temperature measures as provided by the National Centers for Environmental Information and the National Oceanic and Atmospheric Administration.¹² and the identification strategy exploits heterogeneity between states in heat levels.

In our case, it can be useful to see if we still observe a negative relationship between the mortality differential and the poverty rate. We have estimated the model with our variables and select the tuning parameters by cross-validation. We obtained 52 nonzero coefficients. Looking at the value of the coefficients, we find a negative effect of the mortality differential on poverty. The coefficient is - 0.632. Although these estimates are smaller than the OLS that confirm the negative relationship between both variables.

¹² In alternative specifications, we use the maximum temperature instead of the average one. See Section dedicated to robustness tests below.

The first-stage and reduced-form equations of the IV analysis are:

$$MD_{st} = \varphi_1 T_{st} + X'_{st} \pi_1 + \alpha_{1s} + \delta_{1t} + \mu_{1st} \quad (6)$$

$$H_{st} = \varphi_2 T_{st} + X'_{st} \pi_2 + \alpha_{2s} + \delta_{2t} + \mu_{2st} \quad (7)$$

where T_{st} is the average temperature observed in State s in year t . The covariates contained in X_{st} are the same as in the OLS regressions above.

Our identification strategy relies on two important conditions. A first condition is that temperature is an exogenous determinant of mortality, on which humans have, at least in the short run, no direct control. That hypothesis is supported by a voluminous literature showing that mortality is increasing with temperature Basu and Samet (2002); Deschenes and Greenstone (2011); Barreca et al. (2015, 2016).¹³ Moreover, the empirical literature has also established the existence of a substantial heterogeneity in the responses to extreme temperatures across the country. That heterogeneous response is partially due to the self-protection provided by the cooling from increased energy consumption Barreca et al. (2016). The self-protection expenditures, mainly air conditioning, is not equally distributed among the population, such that the mortality from wave of heat impacts differently people with unequal income, which may affect the mortality differential. In the light of this evidence, it is plausible to assume that the mortality differential is affected by the temperature. We show that this condition holds.

The second condition is that the exclusion restriction holds. That is the temperature affects poverty rates *only through* its effects on the mortality differential. That hypothesis would not be correct if, for example, regions with different weather conditions experience different economic situations. But this is unlikely to be the case when we consider U.S. states and the recent period of analysis, especially once we control for macroeconomic factors as well as state fixed effects and government transfers. Several studies showed that temperature effects exist concerning on a series of social and economic factors including agriculture, health, labour productivity or crime and social unrest (see Dell et al. 2012 for a review of papers dealing with these issues). In particular, the relationship between temperature and (national) income has been largely studied (see i.e. Gallup et al. 1999; Dell et al., 2009; Dell et al. 2012; Dell et al. 2014; Miguel et al. 2004). In cross-country comparison, one observes that hot countries tend to be poor (Nordhaus 2006). However, this correlation may be spurious and it may be that other correlated variables such as country's institutions or other economic factor are responsible for the economic prosperity (Sachs, 2003, Acemoglu et al, 2002, Rodrik et al, 2004). Dell et al. (2012), using yearly temperature and precipitation data for each country in the world over 50 years, show negative effects of higher temperatures on income growth but only in poor countries¹⁴. In industrialized (rich) countries, changes in temperature do not have a robust, discernable effect on income growth. These results are somehow confirmed by other studies Gallup et al. (1999); Dell et al. (2009, 2012, 2014) Barrios et al, 2015; Miguel et al. (2004) Hsiang, 2010¹⁵. In the U.S., Deschenes and Greenstone (2007) find no

¹³ Basu and Samet (2002) explain that hot temperatures are associated with excess mortality due to cardiovascular, respiratory, and cerebrovascular diseases. In the U.S., Deschenes and Greenstone (2011) find that each additional day of extreme heat raises the mortality rate. Similar results were reported by Barreca (2012), Curriero et al, (2002) and Deschenes and Moretti (2009).

¹⁴ In some studies, precipitation is included as an explaining factor of mortality. However, it was shown that precipitation has direct effect on economic growth in poor and rich countries, and thus can explain the level of poverty Barreca et al. (2015); Dell et al. (2012).

¹⁵ However, on the contrary to these studies, Burke et al (2015) find a non-linear cross-country relationship between temperature and economic growth. The country-level economic production is non-linear and concave in temperature, this both for rich and poor countries.

statistically significant relationship between weather and U.S. agricultural profit. Zivin and Neidell (2014) find an effect of temperature on labor productivity in industries with high exposure to weather but the effect is driven by health impact which goes in the sense of our use of temperature as a driver of mortality.

One exception is Deryugina and Hsiang (2014) that suggest for the US that, after taking into account many forms of adaptation and price changes, temperature affects contemporaneous income by reducing agricultural productivity and crops. However, their estimations are less important after controlling for rising prices and they find that temperature has no long lasting effects since there is no statistically significant evidence for temporal displacement of income. In a robustness section, we control for share of the agricultural sector to account for the possible effect of temperature on agricultural income. This does not affect our results. Another potential problem to our identification strategy concerns the state-to-state migration. In the US, especially at old age, many individuals move from cold states to warm states. If this movement is more likely among richer people, it may affect the level of poverty and thus show a direct relationship between poverty and climate conditions. Saying differently, warm states would have lower poverty because richer elderly moved in once retired but colder states thus stand with a higher proportion of lower income individuals that cannot move. We have collected data on the net state-to-state migration for each state to control for this potential issue. Although the data are only available since 2005 and thus reduce the size of our sample, results confirm our main findings. This is also presented in the robustness section. Finally our main results will be complemented by a series of robustness tests regarding the definition of the dependent variable, the sample of analysis and covariates.

5.1 Main results

For this strategy to hold, it must be the case that the temperature significantly explains the mortality differential and that, conditional on our observables, the temperature affects the poverty level only through its impact on mortality. The first-stage relationship underlying the 2SLS estimates are shown in Panel A of Table 2. The results suggest that the level of temperature has a strong predictive power for the mortality differential, and that this predictive power is robust to various specifications regarding control variables. We observe a decreasing relationship between average temperature and the mortality differential. The effect remains significant and similar once we introduce other explanatory variables.¹⁶

As stated above, this identification strategy is valid only if conditional on the controls included in the regression, the level of temperature has no effect on the level of poverty, other than the effect through the mortality differential. One advantage of using temperature data is that it is out of direct control. Our identification strategy thus relies on the unpredictable and presumably random year-to-year local variation in temperature, so concerns about omitted variables bias are unlikely to be important. Using our data, we do not find any correlation between poverty and temperature¹⁷ and in the section dedicated to robustness below, we also estimate our 2SLS model for subgroups of states according to their GDP and we still find a significant relationship for every groups of states.

This evidence makes us confident that our instrument is valid. However, in our first-stage estimations, our instrument appears somewhat weak (the F -statistic is between 5 and 6), suggesting that the 2SLS estimates may be somewhat biased toward OLS estimates

¹⁶ In Table 2, we only present the estimates of the variables of interest. See the Appendix for a complete table and the effect of the other covariates.

¹⁷ Results are available upon request.

Table 2 Results of IV regressions

	(1) Basic	(2) + Controlling for Macroeconomic factors	(3) + Controlling for demographic factors	(4) + Controlling for welfare exp. and inc. distr.
Panel A: First-stage				
Temperature	-0.081** (0.035)	-0.083** (0.036)	-0.078** (0.036)	-0.088** (0.035)
<i>F</i> – stat for excluded instruments	5.31	5.42	4.81	6.15
<i>P</i> – value <i>F</i> -statistic	0.022	0.020	0.029	0.013
Panel B: Second-stage				
Mortality differential	-16.398* (8.870)	-18.154* (9.316)	-19.108* (10.244)	-16.143** (8.099)
Interval				
<i>N</i>	678	678	678	678

Note: Standard errors are reported in parentheses. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$. Complete results are presented in the appendix

(Bound et al. 1995; Staiger and Stock, 1997). Though, the Anderson-Rubin test and Stock-Wright LM test reject the null hypothesis that the coefficient is null and indicate that the endogenous regressor is relevant. As an identification check, we estimate a “false experiment” specification in which future average temperature, which should be orthogonal to current mortality differential, is included as an additional explanatory variable, and we find that the coefficient estimate is close to zero and not significant (coefficient estimate -0.006 and standard error 0.035).¹⁸ Reduced-form regressions are presented in Table A.3 in the appendix. The results show that higher temperature is associated with significantly more poverty.

The 2SLS estimates are presented in Panel B of Table 2. The 2SLS coefficients are slightly higher than in the OLS model. Once we introduce control variables, the coefficient of the mortality differential increases a little to stabilize at a reduction of about 16% of the level of the old-age poverty rate for an increase of 1% of the mortality differential.

Table 2 provides empirical support for the missing poor hypothesis: a 1 % rise in the strength of the income gradient in mortality tends to reduce the old-age headcount poverty rate by about 16 %. In states where poor individuals tend to die earlier than non-poor individuals, it is also the case that the measured old-age poverty rates are, *ceteris paribus*, lower. Thus the premature death of the poor pushes old-age poverty measures down. Table 2 suggests that this effect is far from negligible and, based on our identification strategy, we can be confident that this effect from the mortality differential to the measure of old-age poverty is a causal effect, which goes from the mortality/income gradient to the old-age headcount poverty measure, and not the other way around.

Our results tend also to question the – generally assumed – postulate of full comparability of poverty measures across space. If our calculations are correct, the strength of the income gradient in mortality varies across states, so that income-based selection effects also vary across states. This heterogeneity in the strength of selection effects across states tends to question the comparability of old-age poverty rates.

¹⁸ Results are available upon request.

5.2 Robustness check

In order to check the robustness of our results, we perform additional regressions and explore different variables and specifications.

Alternative samples A first step consists of considering alternative subsamples. For that purpose, we first look at the robustness of our results to dropping one state at a time. The coefficient on mortality differential ranges from -19.457 to -12.761 and remains significant at 95 percent confidence levels in all regressions¹⁹. We then check if our results are robust to redefining our sample. The sample of states that we used so far in our regressions is not perfectly balanced, since, for some years, data on poverty are not available for all states present in the mortality database. As shown in Table 4 (column (1)), focusing on the balanced panel ($N = 630$ instead of $N = 678$) does not affect our results: the point estimates are almost identical to the baseline results and the mortality differential, instrumented by our temperature variable, has still a negative and statistically significant impact on old-age measured poverty.

Another robustness check concerns the partition of our sample of U.S. states according to income and regions. One may wonder whether or not our results are driven by the fact that we are comparing states with unequal income levels. As a robustness check, we look at specific results when we consider poor states. States for which the GDP per capita is below the median GDP are classified as low income states. As shown in Table 4 (columns (2)), restricting our sample to low income states ($N = 326$) does not affect our results qualitatively. For the sake of robustness, we also estimate a separate model without the South and rural states and the correlation between the measured old-age poverty and the mortality differential remains negative and statistically significant (available upon request). As discussed in Section 5, it may be that some states experience migration flows such that it is difficult to relate the poverty of people aged 65+ to the mortality of those aged 40. This is the case when one considers the fact that high-income elderly tend to move to warmer states. This is not a easy task to clearly identify in which state they move but as first robustness check we estimate a model removing Florida, Arizona and New Mexico. These three states are probably the more important for that matter. As it is shown in column (3), it does not affect our results. As an additional test, we also control for the net migration rate, see below.

Measurement: mortality differential One may wonder whether our results are driven by how we measure the mortality differential. Various robustness checks were carried out. First, in order to avoid that our results are driven by states in which life expectancy of the richest is higher than in other states, we defined the mortality differential indicator differently, as the ratio between the life expectancy in the last two quartiles over the life expectancy in the first two quartiles ($(L_3 + L_4)/(L_1 + L_2)$). Second, an alternative mortality differential indicator was defined as the *difference* of life expectancies between the last quartile and the first quartile of the income distribution, in such a way as to have a measure of the absolute differential in survival conditions $(L_4 - L_1)$ ²⁰. In specifications (4) and (5) of Table 3, we report robustness tests when we change the definition of the mortality differential. In specification (4), we observe estimates very similar to the results presented in Table 2, both in terms of significance and value of the coefficient. Specification (5) displays rather different coefficients but significant at 10 % and of the same sign. Here the absolute definition of the differential changes the value of coefficient.

¹⁹ Results are available upon request.

²⁰ Although this indicator uses the same numbers as the one used in the benchmark case, these indicator convey different information.

Table 3 Robustness tests

	(1) Balanced panel	(2) Low income states	(3) without AZ, FL, NM states
Mortality differential	-16.796** (8.765)	-21.882** (10.895)	-14.656** (6.509)
N	630 (4)	326 (5)	638 (6)
	Dep. variable $(L_3 + L_4)/(L_1 + L_2)$	Dep. variable $L_4 - L_1$	With control for agriculture
Mortality differential	-24.473* (13.557)	-1.474* (0.914)	-16.073** (6.589)
N	678 (7)	678 (8)	678 (9)
	With control for insurance and health spending	With control for initial poverty	With control for net migration rate
Mortality differential	-17.105** (8.663)	-15.004*** (8.061)	-26.152** (12.348)
N	678	678	496

Note: First-stage regressions are presented in the Appendix. Standard errors are reported in parentheses. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$

Additional covariates One may suspect that our effects may be due to some (omitted) source of heterogeneity across states. Hence, to be sure that our regressions capture a significant relation between old-age measured poverty and the mortality differential and nothing else, it is important to add extra covariates in the analysis.

First, as explained above, it might be the case that actually the temperature affect the agricultural income that in turn affect the level of poverty. To control for this possibility, we include as a covariate the share of the agricultural income in the GDP. This allows to test if it affects directly the level of poverty. Another strategy would have to limit our sample to states that are less dependent on agriculture. Both gives similar results and as displayed in specification (6) in Table 3, the results are not affected. Second, one could argue that one needs to control for private insurance coverage and health expenditures. The addition of those covariates is motivated by the fact that health expenditures are quite sizeable at the old age, and that those expenditures increase strongly with the age.²¹ One may thus expect that the proportion of individuals who are covered by private insurance may affect the comparison of U.S. states. As shown in Table 3, controlling for private health insurance coverage in the state does not substantially affect our results: it is still the case that the mortality differential variable affects negatively the measured old-age poverty. Moreover, controlling also for the amount of public health spending in each state does not affect the robustness of our results.

Another robustness check consists in adding, among the covariates, a measure of *initial* poverty. The underlying intuition is that, as shown in equation (4), one can expect, in theory,

²¹ As shown by De Nardi et al. (2016), health spending more than double between the age of 70 and the age of 90. Medicare and the government pay for about 65 percent of it, about 20 percent is financed out-of-pocket, while 13 percent are paid by private insurance.

that the old-age poverty rate depends on poverty at earlier ages. In order to take this potential effect into account, we add, as an additional covariate, the measured poverty rate for population aged between 18 and 64 years, in each state, with a time lag of 20 years. As shown in Table 3 (specification (8)), even when controlling for initial poverty, it remains true that the mortality differential affects old-age poverty negatively.

Finally, in order to control for the heterogeneity in migration flows between states, we also include the net migration rate in each state as a control (sources: the U.S. Census Bureau).²² Results in specification (9) in Table 3 shows that our main results are not affected once with control for the differences in terms of migration between states.

5.3 Cross-county estimations

We complement our analysis with a cross-county analysis based on data on life expectancy by income at the county level. The Health Inequality Project also releases life expectancy computed at age 40 and by quartile of income for counties with population larger than 25,000 (see Chetty et al. 2016). Unfortunately, because there was too little data to get precise estimates separately in each year, these data are aggregated for the period 2001–2014.

In order to confirm our previous findings, we complete the results obtained from state-level analysis by additional estimations relying on county data using similar instrumental variable. We take the poverty level in 2014 in each county, as obtained from the Current Population Survey of the U.S. Census Bureau, and relate it to the mortality that prevailed during the period 2011–2014²³. Since data on mortality are only available for the period 2001–2014, we use the average temperature experienced in each county over the period 2001–2014 as an instrument for the mortality differential over the same period. Our identification strategy exploits heterogeneity between counties in heat levels. The first-stage and reduced-form equations of the IV analysis are:

$$MD_c = \theta_1 T_c + X'_c \pi_1 + \alpha_{1s} + \nu_{1c} \quad (8)$$

$$H_c = \theta_2 T_c + X'_c \pi_2 + \alpha_{2s} + \nu_{2c} \quad (9)$$

where MD_c and T_c are respectively the mortality differential and the average temperature during the time span 2001 to 2014 in a county c . α_{1s} and α_{2s} are state fixed effects. H_c is the headcount poverty level in 2014 in a county c and X_c contains a large set of contemporaneous controls. The higher number of observations of the county-level analysis allows us to control for many more dimension than the state-level. In particular, we control for economic characteristics in introducing the counties per capita income and the employment rate. In addition, we use a measure of industrial structure of the county as given by the share of working population in the manufacturing sector. We also use the rural-urban continuum code to measure the urban influence on poverty. Demographic characteristics are controlled through the introduction of the share of the county population aged 65+ and the percentage of non-black population as well as the share of school dropouts. We also include an indicator of income distribution, namely the Gini index. Finally, we control for social and political characteristics and include the local tax rate and the local government expenditure. All regressions are weighted using county population.

²² The net migration rate is the difference between the number of people coming into the state and the number of people leaving the state during the year, over the population of the state.

²³ Alternatively, we could use the average poverty over the 2001–2014 period. This empirical strategy gives similar results. The advantage of using the 2001–2014 period mortality differential to explain the poverty in 2014 is that it allows to take into account the possible time effect of mortality on poverty.

Table 4 displays the results. The first-stage regression confirms the validity of our instrument. The F -stat is much higher than in the state-level regressions. The level of temperature has a strong predictive power for the mortality differential, and as before, this predictive power is robust to various specifications regarding control variables. The effect of the mortality differential on the level of poverty is also similar to what we obtain with state data. When we control for the all set of covariates, the results shows a reduction of about 17% of the level of the old-age poverty rate for an increase of 1% of the mortality differential.

These results are robust to a series of robustness checks (presented in Appendix). First, the finding is robust to dropping one state at a time (we did not do the test with each county for sake of calculation) with coefficient ranging from -19.65 to -16.05 and remaining significant at 99% confidence level in all regressions. Second, we also conduct separate analyses using metro and non-metro counties to identify differences between rural and urban counties. We find similar results in both groups of counties although the effect is only significant at 10% when we consider the rural counties only. As for the state-level analysis, another important robustness check concerns the partition of our sample of U.S. counties into low and high income counties. In order to check that our results are not driven by this, we divide our sample into two samples: low-income and high-income states. Counties for which the GDP per capita is below the median are classified as low income states. Here also, we obtain similar results. Finally, we also test the effect of introducing specific control for medicare, the migration rate and the initial level of poverty and it does not affect our results.

6 Towards comparable poverty measures

Whereas our results - the non-rejection of the missing poor hypothesis - question the comparability of old-age poverty measures across states or counties, this section proposes to use

Table 4 Cross-county IV regressions

	(1)	(2)	(3)	(4)
Panel A: First-stage				
Temperature	-0.068*** (0.018)	-0.072*** (0.018)	-0.066*** (0.120)	-0.063*** (0.018)
F-stat for excluded instruments	14.13	15.96	13.48	12.42
P-value F-statistic	0.000	0.000	0.000	0.000
Panel B: Second-stage				
Mortality differential	-26.051*** (7.709)	-19.937*** (5.610)	-20.739*** (6.313)	-17.297*** (5.643)
Economic controls	Yes	Yes	Yes	Yes
Demographic controls	No	Yes	Yes	Yes
Inequality control	No	No	Yes	Yes
Social and Pol. controls	No	No	No	Yes
N	1150	1150	1150	1150

Note: Standard errors are reported in parentheses. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$

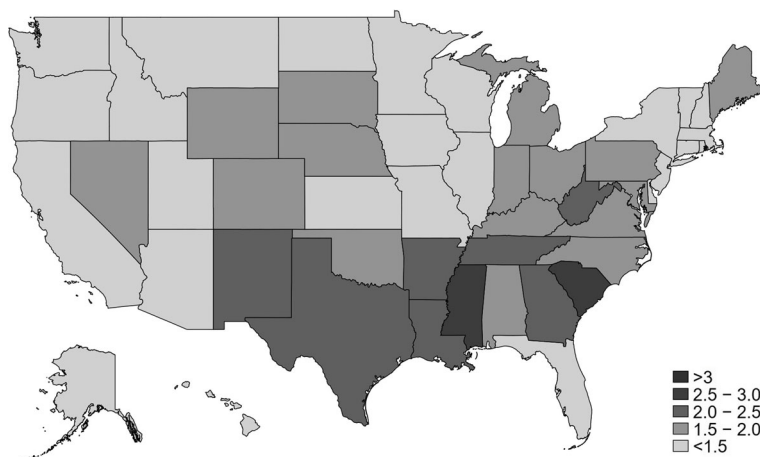


Fig. 6 Differential between the old-age poverty rate corrected for the missing poor bias and the standard old-age poverty rate (in percentage points), U.S. states, 2014

the results of regressions of Section 5 to build old-age poverty measures that are *comparable across states*.²⁴

In this section, we simulate hypothetical old-age poverty rates, where the impact of mortality differential is, by construction, neutralized. We consider a hypothetical case where the income mortality differential variable takes, in each state, the value of 1, so that poor and non-poor individuals face the same survival conditions. Then, we use our regressions results from Section 5 to calculate the *hypothetical poverty rates that would have prevailed provided income-based selection effects were cancelled out*. The hypothetical old-age poverty rate differs from the standard old-age poverty rate on an extent that depends on (1) the strength of the income gradient in mortality is in that state; (2) the estimated causal impact of the income gradient in mortality on measured old-age poverty.

Figure 6 shows, across the U.S. states, the size of the differential between, on the one hand, the hypothetical old-age poverty rate (correcting for the missing poor bias), and, on the other hand, the standard old-age headcount poverty rate. In each state, the hypothetical old-age poverty rates are *higher* than the standard old-age poverty rates. This result is not surprising, since it was shown above that there exists a significant income gradient in mortality in U.S. states, and that this income gradient in mortality has a negative and statistically significant impact on old-age poverty measures. Relying on standard (uncorrected) measures of old-age poverty tends to minor the prevalence of poverty at the old age, by focusing on measures that are subject to income-based selection effects.

The magnitude of the differentials between the standard and the hypothetical old-age poverty rates varies across states, from about 1 percentage point to more than 3 percentage points. The gap is larger in states where the mortality differential variable takes higher values, that is, in states where the income gradient in mortality is stronger. The differential between the simulated and the standard old-age poverty rate can be interpreted, for a given state, as a measure of the missing poor effect, that is, the impact of income-based selection effects on the measurement of old-age poverty.

²⁴ We use state-level regressions because they constitute baseline analysis in which we can control for both endogeneity and fixed unobserved heterogeneity.

The varying size of the missing poor bias has important consequences regarding the comparison of old-age poverty across states. The ranking of states in terms of old-age poverty is affected by correcting selection effects. To see this, Fig. 7 shows, state by state, the standard and the corrected old-age poverty rates. Figure 7 makes appear the occurrence of a large number of rank reversals between states (in terms of old-age poverty).

Take the example of Massachusetts (MA) and Maryland (MD). Standard headcount old-age poverty measures indicate that old-age poverty is slightly larger in Massachusetts than in Maryland. However, the mortality differential between the poor and the non-poor is much lower in Massachusetts than in Maryland. When computing hypothetical old-age poverty measures while neutralizing for the impact of the income gradient in mortality, it appears that old-age poverty is actually larger in Maryland than in Massachusetts. Another example of such a reversal of ranking across states is given by the comparison of old-age poverty in New Jersey (NJ) and Nevada (NV). Whereas old-age poverty is, on the basis of standard headcount ratios, higher in New Jersey than in Nevada, the correction of the selection bias leads to reverse that ranking. Another case of ranking reversal is given by the comparison of Wisconsin (WI) with Wyoming (WY). Those ranking reversals suggest that correcting for the missing poor bias affects the comparison of old-age poverty across U.S. states.

In sum, the missing poor phenomenon has an important impact on the measurement of old-age poverty in the U.S., and on the comparison of U.S. states in terms of old-age poverty. Whereas the missing poor bias is present in all states, it affects poverty measurement in varying degrees across states, depending on the strength of the income gradient in mortality. Correcting for the missing poor bias affects the ranking of states in terms of old-age poverty.

7 Mechanisms

Although the existence of a negative causal effect of the income-based mortality differential on the measured old-age poverty is a result on its own, one may want to know more about the

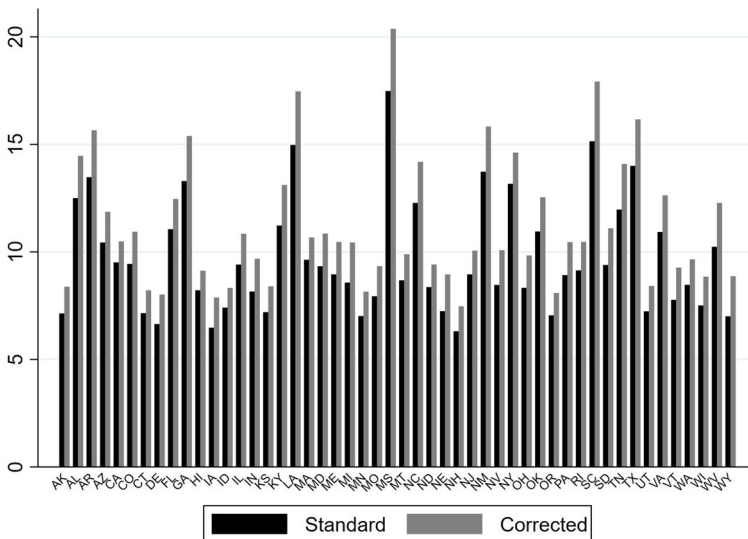


Fig. 7 Standard old-age poverty rates and corrected old-age poverty rates (without mortality differentials), in percents, U.S. states, 2014

mechanisms at work behind that effect. Which mechanisms lie behind the missing poor bias? Through which channels do poor individuals "disappear" from old-age headcount poverty statistics?

There are potentially many mechanisms behind the missing poor phenomenon. One may think about policy variables, such as differences across states in terms of access to health care or to health care insurance. Despite the federal Medicaid program, states with lower measured old-age poverty may be characterized by a more restricted access to health care or health insurance (leading to a higher mortality differential). But other explanations may consist of differences across states in terms of unhealthy lifestyles. States with lower measured old-age poverty could be also characterized by a higher prevalence of risky lifestyles (leading, also, to a higher mortality differential).²⁵

Another difficulty lies in the fact that some mechanisms may not be captured by macro databases. For instance, a recent microeconomic study focussing on renal diseases and access to transplantation in France showed that, despite the universal coverage of the French health social insurance system, there remain, at different stages of the medical process, significant inequalities in treatment across groups, depending on their socioeconomic status Baudelot et al. (2016). The existence of such effects in the U.S. cannot be excluded *a priori*, and macro data can hardly allow us to identify these effects.

A third difficulty is that one has also to be careful about issues of reverse causation: it cannot be excluded, in theory, that some channels work in both directions, from mortality differential to poverty measurement and *vice versa*. Thus one has to be extremely cautious in interpreting potential evidence supporting some particular mechanisms that may be at work.

In the light of those difficulties, one cannot reasonably claim to be able to identify a particular mechanism; instead, this section will present some empirical elements that may be interpreted as "clues" supporting, or, alternatively, rejecting, mechanisms at work behind the missing poor phenomenon.

7.1 Insurance coverage

A natural candidate mechanism behind the missing poor phenomenon consists in inequalities in health care access. Note that, if the missing poor phenomenon were driven by unequal access to health care (between the poor and the non-poor), we would expect that, among states with the lowest measured old-age poverty, we would also find, *ceteris paribus*, a larger proportion of individuals with limited access to health care. This would entail a higher mortality differential because poor people with limited access to health care would experience shorter life expectancy

In the US, poor individuals are eligible to the Medicaid program and older individuals (65+) are eligible to Medicare. These programs ensure that they have access to health care even if they are not insured elsewhere. Although this seems to rule out *de facto* a "health care channel" for the missing poor phenomenon, there is however variation in the coverage by Medicaid across states. Indeed, Medicaid is administered by states and they have flexibility to determine what populations and services to cover. Because of this flexibility, there is important heterogeneity across state Medicaid programs, and as a result, the share of state residents covered by the program. The percentage of the population covered by Medicaid ranges from 10% to more than 30%, which is far from the difference in poverty rates between states.

²⁵ Note that it is hard, without further analysis, to separate those two possible mechanisms, since both mechanisms lead to income-differentiated mortality.

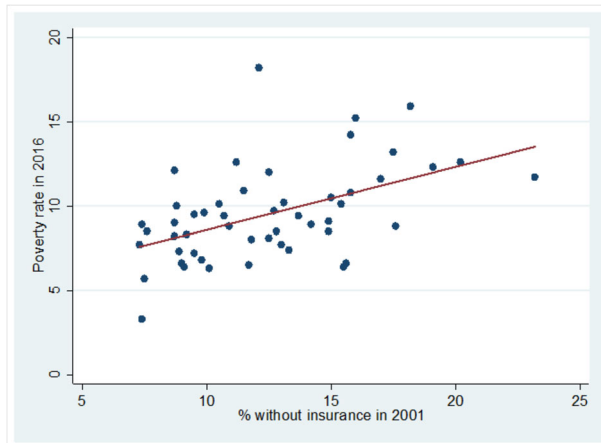


Fig. 8 Percentage of uninsured individuals (2001) and old-age measured poverty (2016)

In order to check that restricted access health care within the population is not related to the prevalence of old-age poverty, we can use the proportion of uninsured individuals in the population as a proxy measure of the extent to which access to health is limited in a given state. These individuals are considered as uninsured, in the sense that they are aged less than 65 (and thus non eligible to Medicare), not covered by Medicaid and not covered by private insurance. Using that proxy variable, the story behind the missing poor problem would go as follows: in states where a larger proportion of individuals are uninsured, access to health care would be more limited, which would lead to a lower measured old-age poverty through a selection effect.

When studying the relation between insurance coverage and old-age measured poverty, one must be cautious about the possibility of reverse causality (i.e. from poverty to insurance coverage). In order to minimize the problems raised by reverse causality, our analysis uses lagged variables for insurance coverage. Figure 9 plots old-age poverty rates by states in 2016 against the proportion of the uninsured population by state in 2001.²⁶ The underlying intuition is that, although poverty in 2016 may affect insurance coverage in 2016, it is less likely to influence it 15 years before.

As shown in Fig. 9, the relation between the (lagged) percentage of uninsured individuals in the population by states and the measured old-age poverty is increasing. Thus Fig. 9 does not support the existence of a "health insurance channel" behind the missing poor phenomenon. If limited insurance coverage were the mechanism driving the missing poor phenomenon, then there would be a *decreasing* relation between the proportion of uninsured individuals and the measured old-age poverty, unlike what appears on Fig. 9.

7.2 Causes of death

Another approach consists of shifting the analysis from the study of overall mortality to the study of mortality by causes. Actually, studying mortality by causes is an indirect way to account for interstate heterogeneity in unhealthy lifestyles. If the missing poor phenomenon were driven by inequalities across states in terms of lifestyles, then it would also be the case

²⁶ Sources: U.S. Census Bureau, 1-year American Community Surveys, 2018.

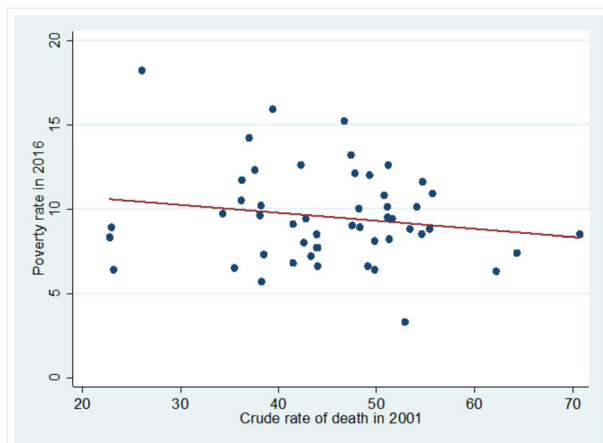


Fig. 9 Respiratory related deaths, in per 100,000 (2001) and old-age measured poverty (2016), in percent

that, among states with the lowest measured old-age poverty, we would find, *ceteris paribus*, a larger proportion of deaths related to causes associated to unhealthy lifestyles.

In order to examine the plausibility of the lifestyle channel, this section plots crude death rates by causes in 2001 against measured old-age poverty in 2016, for three distinct causes of death: alcohol related death, diabete-related death, and respiratory-related death.²⁷ As in the previous section, we use here time lags to try to minimize problems due to reverse causality.

Figure 10 shows a decreasing relationship between the (lagged) respiratory-related crude death rate and the measured old-age poverty. This decreasing relationship is not, of course, a proof that the missing poor phenomenon is driven by the "respiratory channel", but at least this constitutes a clue suggesting that this channel does not seem incompatible with the missing poor phenomenon. However, Fig. 11, which shows an increasing relation between (lagged) alcohol-related death and old-age poverty, leads to the rejection of that "alcohol mechanism". Similarly, Fig. 12 rejects the "diabete mechanism".²⁸

8 Conclusions

Although the income gradient in mortality is widely documented, there are few empirical attempts to test empirically the consequences of that gradient on the measurement of poverty. In particular, one may expect, in line with the missing poor hypothesis, that regions with a steeper income gradient in mortality exhibit lower old-age poverty measures, because of a stronger income-based selection, which excludes the poor from the population under study.

²⁷ Our emphasis on alcohol-related deaths is motivated by empirical studies showing that this source of death is correlated with the socio-economic status (see Romelsjö and Lundberg (1996); Makela (1999) and Van Oers et al. (1999)). Our focus on respiratory diseases is motivated by empirical studies relating poverty to tobacco consumption, such as Flint and Novotny (1997) and Hausteijn (2006).

²⁸ We also carried out other tests concerning potential mechanisms, while also using a 15-year lag between the explanatory variable (measured in 2001) and the old-age poverty rate (measured in 2016). We found that health expenditures per capita are positively related to old-age poverty measured 15 years later, whereas for nursing homes per capita we found a decreasing relationship with old-age poverty measured 15 years later. Those additional tests are available upon request.



Fig. 10 Alcohol related deaths (2001) in per 100,000 and old-age measured poverty (2016) in percent

The goal of this paper was to provide an empirical test of the missing poor hypothesis. For that purpose, we used U.S. state-level and county-level data on poverty at age 65+ and life expectancy by income levels, and we tried to estimate the impact of variations in the mortality differential (measured by the ratio of the life expectancy of the first and the fourth income quartiles) on the headcount poverty rate at age 65+. We showed that the strength of the income gradient in mortality has a negative and statistically significant effect on the old-age headcount poverty rate, and that this effect is robust to various specifications concerning control variables. In order to address potential reverse causation issues, we also developed an IV approach, by instrumenting the mortality differential variable by means of temperature statistics. Our IV regressions showed that a 1 % rise in the instrumented mortality differential causes a 16 % decrease in the old-age headcount poverty rate.

In the light of those results, one can seriously question the comparability of standard old-age poverty rates. A lower measured old-age poverty may be due not to a lower prevalence



Fig. 11 Diabetes related deaths (2001) in per 100,000 and old-age measured poverty (2016) in percent

of poverty, but to a stronger income-based selection, which excludes the poor from the old-age. In order to deal with this comparability problem, we proposed, in a second stage, to calculate hypothetical old-age poverty rates while neutralizing the income-mortality gradient (i.e. assuming that survival conditions are similar for all income quartiles). This thought experiment allowed us to show that neutralizing selection effects can lead to a rise of old-age headcount poverty rates by between 1 and 3 percentage points. Correcting for the missing poor problem was also shown to significantly affect the comparison of states in terms of old-age poverty. The ranking of states in terms of standard old-age poverty rates may differ from the ranking of states in terms of hypothetical old-age poverty rates corrected for selection effects.

Regarding possible mechanisms, our analysis suggests that some mechanisms are more plausible than others. Unequal health insurance coverage could hardly be the driving force behind the missing poor bias, but some dimensions of lifestyles, leading to respiratory-related deaths, may play some role. But the identification of the mechanisms behind the missing poor phenomenon is complex, and would require another paper.

Although our analysis provides robust empirical support for the missing poor hypothesis, our results should be interpreted with caution. Actually, our empirical test of the missing poor hypothesis relies on highly aggregated, macro datasets. As such, our analysis does not provide direct, microeconomic evidence supporting the missing poor hypothesis, but, rather, some form of *indirect* macroeconomic evidence compatible with that hypothesis. In some sense, this paper provides a "first pass" for the missing poor hypothesis.

Having stressed this, our results point nonetheless to a real problem for the measurement of poverty. If our calculations are correct, the missing poor hypothesis is true, and standard old-age poverty measures are biased downwards due to selection bias associated to mortality differentials between the poor and the non-poor. Moreover, income gradients in mortality varying across states tend to question the comparability of standard poverty measures at the old age.

Acknowledgements The authors would like to thank two anonymous referees for their insightful suggestions. They thank also participants of seminars at the University of Freiburg and Aix-Marseille School of Economics as well as participants at the 19th Journées Louis-André Gerard-Varet, the 8th meeting of the Society for the Study of Economic Inequality and the 20th International Conference on Public Economic Theory. This work was supported by the French National Research Agency Grant ANR-17-EURE-0020, and by the Excellence Initiative of Aix-Marseille University - A*MIDEX.

References

- Acemoglu, D., Johnson, S.: Disease and development?: the effect of life expectancy on economic growth. *J. Polit. Econ.* **115**(6), 925–985 (2007)
- Aldy, J.E.: An environmental kuznets curve analysis of U.S state-level carbon dioxide emissions. *The J. Environ. Development* **14**(1), 48–72 (2005)
- Anderson, H.: Air pollution and mortality: A history. *Atmos. Environ.* **43**(16), 142–152 (2009)
- Andrews, I., Stock, J.H., Sung, L.: Weak Instruments in IV Regression: Theory and Practice. *Annual Rev. Econ.* **11**(1), 727–753 (2019)
- Apergis, N., Christou, C., Gupta, R.: Are there Environmental Kuznets Curves for US state-level CO2 emissions?. *Renewable and Sustainable Energy Reviews* **69**(C), 551–558 (2017)
- Atasoy, B.: Testing the environmental Kuznets curve hypothesis across the U.S.: Evidence from panel mean group estimators. *Renewable and Sustainable Energy Reviews* **77**(C), s 731–747 (2017)
- Backlund, E., Sorlie, P., Johnson, N.: A comparison of the relationships of education and income with mortality: the national longitudinal mortality study. *Soc. Sci. Med.* **49**(10), 1373–1384 (1999)

- Barreca, A., Clay, K., Deschênes, O., Greenstone, M., Shapiro, J.: Convergence in Adaptation to Climate Change: Evidence from High Temperatures and Mortality, 1900–2004. *Amer. Econ. Rev.* **105**(5), 247–251 (2015)
- Barreca, A., Clay, K., Deschênes, O., Greenstone, M., Shapiro, J.: Adapating to Climate Change: The remarkable decline in the US Temperature-mortality relationship over the twentieth century. *J. Polit. Econ.* **124**(1), 105–160 (2016)
- Basu, R., Samet, J.: Relation between elevated ambient temperature and mortality: a review of the epidemiologic evidence. *Epidemiologic Review.* 190–202 (2002)
- Baudelot, C., Caillé, Y., Godechot, O., Mercier, S.: Renal diseases and social inequalities in access to transplantation in France. *Population-E* **71**(1), 23–51 (2016)
- Beelen, R., Raaschou-Nielsen, O., Stafoggia, M., et al.: Effects of long-term exposure to air pollution on natural-cause mortality: an analysis of 22 European cohorts within the multicentre ESCAPE project. *The Lancet* **383**, 785–795 (2014)
- Belloni, M., Alessie, R., Kalwij, A., Marinacci, C.: Lifetime income and old-age mortality risk in Italy over two decades. *Demogr. Res.* **29**(45), 1261–1298 (2013)
- Bossuyt, N., Gadeyne, S., Deboosere, P., Van Oyen, H.: Socio-economic inequalities in healthy expectancy in Belgium. *Public Health* **118**, 3–10 (2004)
- Cervelatti, M., Sunde, U.: Life expectancy and economic growth?: the role of the demographic transition. *J. Econ. Growth* **16**(2), 99–133 (2011)
- Chan, M., Weinhold, R., Thoams, R., Gohlke, J., Portier, C.: Environmental predictors of US county mortality patterns on a national basis. *PLoS One* **10**(12) (2015)
- Chetty, R., Stepner, M., Abraham, S., Lin, S., Scuderi, B., Turner, N., Bergeron, A., Cutler, D.: The Association between income and life expectancy in the United States, 2001 – 2014. *The Journal of the American Medical Association*, April 11, 2016, 315 (14) (2016)
- Cristia, J.P.: Rising mortality and life expectancy differentials by lifetime earnings in the United States. *J. Health Econ.* **28**(5), 984–995 (2009)
- De Nardi, M., French, E., Bailey Jones, J., McCauley, J.: Medical spending of the US elderly. *Fisc. Stud.* **37**(3–4), 717–747 (2016)
- Deaton, A.: Health, inequality and economic development. *J. Econ. Literature* **41**, 113–158 (2003)
- Deaton, A., Paxson, C.: Aging and inequality in income and health. *Amer. Econ. Rev.* **88**, 248–253 (1998)
- Dell, M., Jones, B., Olken, B.: Temperature and income: reconciling new cross-sectional and panel estimates. *Amer. Econ. Rev.: Papers And Proceedings* **99**(2), 198–204 (2009)
- Dell, M., Jones, B., Olken, B.: Temperature Shocks and Economic Growth: Evidence from the Last Half Century. *Amer. Econ. J. Macroeconomic* **4**(3), 66–75 (2012)
- Dell, M., Jones, B., Olken, B.: What do we learn from the weather? The new climate-Economy literature. *J. Econ. Literature* **52**(3), 740–798 (2014)
- Deryugina, T., Hsiang, S.: Does the environment still matter? daily temperature and income in the United States, NBER Working Paper. No. 20750 (2014)
- Deschenes, O., Greenstone, M.: Climate Change, Mortality, and Adaptation: Evidence from Annual Fluctuations in Weather in the US. *Am. Econ. J. Appl. Econ.* **3**(4), 152–185 (2011)
- Deschenes, O., Moretti, E.: Extreme weather events, mortality and migration. *Rev. Econ. Statistics* **91**(4), 659–81 (2009)
- Di, Q., Dominici, F., Schwartz, J.D.: Air pollution and mortality in the Medicare population. *N. Engl. J. Med.* **376**, 2513–2522 (2017)
- Duggan, J., Gillingham, R., Greenlees, J.: Mortality and lifetime income: evidence from U.S. Social Security Records. IMF Working Paper, 07/15 (2007)
- Duleep, H.O.: Measuring the effect of income on adult mortality using longitudinal administrative record data. *J. Human Resources* **21**(2), 238–251 (1986)
- Flint, A., Novotny, T.: Poverty status and cigarette smoking prevalence and cessation in the United States, 1983–1993: the independent risk of being poor. *Tob. Control* **6**, 14–18 (1997)
- Forastiere, F., Stafoggia, M., Tasco, C., Picciotto, S., Agabiti, N., Cesaroni, G., Schwartz, C.: Socioeconomic status, particulate air pollution and daily mortality: differential exposure or differential susceptibility. *Am. J. Ind. Med.* **50**(3), 208–216 (2007)
- Gallup, J., Sachs, J., Mellinger, A.: Geography and Economic Development. *Int. Reg. Sci. Rev.* **22**(2), 179–232 (1999)
- Haustein, K.O.: Smoking and poverty. *Eur. J. Prev. Cardiol.* **13**(3), 312–318 (2006)
- Hazan, M.: Longevity and Lifetime Labor Supply: Evidence and Implications. *Econometrica* **77**(6), 1829–1863 (2009)
- Jusot, F.: Mortalité et revenu en France: construction et résultats d’une enquête cas-témoins. *Santé, Société et Solidarité* **3**(2), 173–186 (2004)

- Kalwij, A., Alessie, R., Knoef, M.: The association between the individual income and the remaining life expectancy at the age of 65 in the Netherlands. *Demography* **50**(1), 181–206 (2013)
- Kanbur, R., Mukherjee, D.: Premature mortality and poverty measurement. *Bull. Econ. Res.* **59**(4), 339–359 (2007)
- Lefebvre, M., Pestieau, P., Ponthiere, G.: Measuring poverty without the Mortality Paradox. *Soc. Choice Welfare* **40**(1), 285–316 (2013)
- Lefebvre, M., Pestieau, P., Ponthiere, G.: FGT old-age poverty measures and the Mortality Paradox: theory and evidence. *Rev. Income Wealth* **64**(2), 428–458 (2018)
- Lefebvre, M., Pestieau, P., Ponthiere, G.: Missing poor and income mobility. *J. Comp. Econ.* **47**, 330–366 (2019)
- Lefebvre, M., Pestieau, P., Ponthiere, G.: Premature mortality and poverty measurement in an OLG economy. *J. Popul. Econ.* **32**(2), 621–664 (2019)
- Makela, P.: Alcohol-related mortality as a function of socio-economic status. *Addiction* **94**, 867–886 (1999)
- Miguel, E., Satyanath, S., Sergenti, E.: Economic Shocks and Civil Conflict: An Instrumental Variables Approach. *J. Polit. Econ.* **112**(4), 725–53 (2004)
- Milligan, K., Schirle, T.: The evolution of longevity: Evidence from Canada. NBER Working Paper No. 24929 (2018)
- Milojevic, A., Niedzwiedz, C., Pearce, J., Milner, J., Mackenzie, I., Doherty, R., Wilkinson, P.: Socioeconomic and urban-rural differentials in exposure to air pollution and mortality burden in England. *Environ. Health* **16**(1), 104 (2017)
- Pamuk, E.R.: Social class inequality in mortality from 1921 to 1972 in England and Wales. *Popul. Stud.* **39**, 17–31 (1985)
- Pamuk, E.R.: Social class inequality in infant mortality in England and Wales from 1921 to 1980. *Eur. J. Popul.* **4**, 1–21 (1988)
- Qian, D., Wang, Y., Zanobetti, A., Wang, Y., Koutrakis, P., Choirat, C., Dominici, F., Schwartz, J.: Air Pollution and Mortality in the Medicare Population. *N. Engl. J. Med.* **376**, 2513–2522 (2017)
- Romelsjö, A., Lundberg, M.: The changes in the social class distribution of moderate and high alcohol consumption and of alcohol-related disabilities over time in Stockholm County in Sweden. *Addiction* **91**, 1307–1323 (1996)
- Salm, M.: The effect of pensions on longevity: evidence from Union Army veterans. *Economic Journal* **121**, 595–619 (2012)
- Sen, A.K.: Mortality as an indicator of economic success and failure. *Economic Journal* **108**, 1–25 (1998)
- Snyder, S., Evans, W.: The effect of income on mortality: evidence from the social security notch. *Rev. Econ. Stat.* **88**(3), 482–495 (2006)
- Soytas, U., Sari, R., Ewing, B.: Energy consumption, income, and carbon emissions in the United States. *Ecol. Econ.* **62**(3–4), 482–489 (2007)
- Staiger, D., Stock, J.H.: Instrumental Variables Regression with Weak Instruments. *Econometrica* **65**(3), 557–586 (1997)
- Van Oers, J.A., Bongers, I.M., Van De Goor, L.A., Garretsen, H.F.: Alcohol consumption, alcohol-related problems, problem drinking, and socioeconomic status. *Alcohol Alcohol.* **34**, 78–88 (1999)
- Van Oyen, H., Bossuyt, N., Deboosere, P., Gadeyne, S., Abatith, E., Demarest, S.: Differential inequity in health expectancy by region in Belgium. *Social Prevention Medicine* **50**(5), 301–310 (2005)
- Wang, Y., Kloog, I., Coull, B.A., Kosheleva, A., Zanobetti, A., Schwartz, J.D.: Estimating causal effects of long-term PM_{2.5} exposure on mortality in New Jersey. *Environmental Health Perspective* **124**, 1182–1188 (2016)

Publisher's Note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Springer Nature or its licensor (e.g. a society or other partner) holds exclusive rights to this article under a publishing agreement with the author(s) or other rightsholder(s); author self-archiving of the accepted manuscript version of this article is solely governed by the terms of such publishing agreement and applicable law.