The macro-level effect of religiosity on health

Herzer, Dierk

Veröffentlichungsversion / Published Version
Zeitschriftenartikel / journal article

Empfohlene Zitierung / Suggested Citation:
Herzer, D. (2022). The macro-level effect of religiosity on health. *Health Economics*, *31*(6), 993-1011. https://doi.org/10.1002/hec.4494

Received: 30 November 2020 | Revised: 21 December 2021 | Accepted: 6 February 2022

DOI: 10.1002/hec.4494

Health Economics WILEY

RESEARCH ARTICLE

The macro-level effect of religiosity on health

Dierk Herzer

Department of Economics, Helmut-Schmidt-University Hamburg, Hamburg, Germany

Correspondence
Dierk Herzer, Department of Economics, Helmut-Schmidt-University Hamburg, Holstenhofweg 85, 22043 Hamburg, Germany.
Email: herzer@hsu-hh.de

Funding information
Open access funding enabled and organized by Projekt DEAL

Abstract

An issue that has not yet been explored in the religiosity-health literature is the macro-level effect of religiosity on health—the effect of the religiosity of a society on the absolute health of the population of that society as a whole. We address this issue using two panel datasets: The first is a time-series cross-sectional panel dataset for 17 countries from 1925 to 2000. The second is a cross-sectionally dominated panel dataset of up to 92 countries for the period 1981–2016. Our main findings are as follows: first, religiosity has a significant negative causal effect on health at the macro level; second, a substantial part of this effect can be attributed to an indirect effect via public health expenditures; and third, changes in population health do not cause significant changes in societal religiosity.

KEYWORDS
health, life expectancy at birth, macro level, panel data, public health expenditures, religiosity

1 | INTRODUCTION

Numerous studies, summarized in several review articles (e.g., Koenig, 2012; VanderWeele, 2017; Zimmer et al., 2016), have found that people who are more religious have better health and lower mortality than people who are less religious. Three explanations for this phenomenon are that religion regulates behavior and therefore acts as a protective factor against many risk behaviors, including tobacco, alcohol, and drug use; that religion provides formal and/or informal networks of emotional, social and material support, sometimes including health services; and that religion provides meaning and is therefore an important resource for coping with stress and depression (e.g., Deaton, 2011).

While the positive effect of religiosity on health is well established at the individual or micro level, no studies have examined the macro-level effect of religiosity on health—the effect of the religiosity of a society on the absolute health of the population of that society as a whole. However, this macro-level effect could differ from its micro-level counterpart for at least two reasons.

One reason is that declining macro-religiosity could promote medical progress that benefits society as a whole. Arguments in support of this hypothesis are: (i) Studies at the individual level have shown that religiosity negatively affects attitudes toward science in general (e.g., Bénabou et al., 2015; McPhetres & Zuckerman, 2018) and biomedical research in particular (e.g., Brossard et al., 2009; Nisbet & Goidel, 2007). (ii) Studies have also shown that religious people are less likely to choose careers in science (e.g., Ecklund, 2010; Scheitle & Ecklund, 2017). And (iii) there is also evidence that

The manuscript contains original unpublished work and is not being submitted for publication elsewhere.

religiosity has a negative macro effect on the stock of technical knowledge, measured by total factor productivity (Herzer & Strulik, 2020).[1]

It could therefore be that a decline in the share of religious people leads to a better environment for medical research, and thus to improvements in medical technology, which have been shown to have a large impact on population health (e.g., Fonseca et al., 2021; Skinner & Staiger, 2015). Thus, a decline in religiosity at the macro level could have both a direct negative effect on population health—by increasing the proportion of non-religious individuals, whose health status is poor relative to religious individuals—and an indirect positive effect on population health—by increasing the health of both religious and non-religious people through better understanding of how to prevent and heal disease and through increasing the availability of better medicines, medical devices, and diagnostic techniques and treatments. Depending on which effect dominates, population health could improve or deteriorate as a result of a decline in societal religiosity.

A second reason why a decline in macro-religiosity could exert a positive total effect on population health is the following. If public health expenditures improve health, as several recent studies suggest (e.g., Bernet et al., 2020; Bokhari et al., 2007), then lower societal religiosity could have an additional, positive, indirect positive effect on population health, through a positive effect on public health expenditures. Main arguments in support of the hypothesis that an increase in the share of non-religious individuals leads to an increase in public health expenditure are: (i) There is a literature that suggests that Church and state are substitutes in providing social services (e.g., Gruber & Hungerman, 2007; Hungerman, 2005). Within this literature, Dills & Hernández-Julián, 2014, found evidence for the United States that a decline in religiosity leads to an increase in public social spending, particularly on health. Thus, declining religiosity could reduce the use of health services provided by religious institutions and, as a consequence, increase the demand for public health services. (ii) Scheve & Stasavage, 2006, developed a model in which it is assumed that social insurance and religious engagement are two alternative mechanisms that limit the costs of adverse life events. The predictions of their model are that religious individuals should prefer lower levels of social-insurance provision than those who are secular, and that, if policy outcomes reflect individual preferences, countries with higher levels of religiosity should have lower levels of welfare spending; both of these predictions are supported by their empirical analysis. To the extent that religion provides psychological benefits that help individuals to cope with adverse life events, as assumed in the model of Scheve & Stasavage, 2006, it improves individual health. Thus, it can be intuitively argued that religious individuals may prefer lower levels of public health expenditures because they are healthier, and that a growing share of non-religious individuals may therefore increase the demand for public health services.

Since all governments spend money on health services, it is reasonable to assume that governments generally respond to demands for public health services (for social, political or economic reasons), and since a decline in societal religiosity could lead to increased demand for public health services (as just discussed), it may be that public health expenditures increase in response to a decline in societal religiosity. An increase in public health expenditures that results in an increase in the supply of public health services, in turn, implies better access to health services for the population as a whole, including religious and non-religious people. Thus, an increase in public health expenditures—driven by the demand by non-religious individuals for public health services—may result in health improvements for both religious and non-religious people.

It is therefore possible that while individual religiosity has a positive effect on individual health, the effect of societal religiosity on population health is negative, not only because societal religiosity may have a negative effect on medical progress but also because societal religiosity may have a negative impact on public health expenditures. In other words, we hypothesize that although religious people tend to be healthier than less religious and non-religious people, a decline in societal religiosity can improve population health because such a decline may lead to improvements in medical technology and increases in spending on public health—factors that benefit the health of all (religious and non-religious) people.

It is perhaps useful to illustrate this with a simple example. Imagine a society consisting of 10 individuals. Suppose that nine of these 10 are religious, whereas one individual is not religious. Suppose further that, because religious people are healthier, the life expectancy of the religious society members is 30 years at age x, whereas the non-religious individual has a life expectancy of 25 years at age x, from which it follows that the average life expectancy at age x in this hypothetical society is 29.5 years. Now consider a situation where the number of religious people decreases to 5. Suppose that, in this situation, the life expectancies of religious and non-religious people increase, respectively, to 35 and 30 years—because the increase in the share of non-religious people promotes medical progress and results in an increase in public health expenditures, and because this leads to improvements in the health of both religious and non-religious people. Thus, in this situation, religious individuals (still) have better health than secular individuals, implying a positive effect of religiosity on health at the individual level. At the same time, however, the average life expectancy at age x increases from

29.5 to 32.5 years in this situation, implying that, in this example, societal religiosity has a negative impact on the health of the population as a whole. This macro-level effect of religiosity on health is the subject of this paper.

The contributions of this paper are as follows: First, using two panel datasets, we are the first to examine how and to what extent societal religiosity, measured by the rate of church attendance and the share of religious people, respectively, affects population health, measured by life expectancy at birth. The first dataset is a time-series cross-sectional panel dataset with repeated observations for 5-year periods for 17 countries over a long period, from 1925 to 2000. The second is a cross-sectionally dominated panel dataset of up to 92 countries with between one and six annual observations per country between 1981 and 2016. Since each dataset has both advantages and disadvantages (which will become apparent later in this paper), we employ both to ensure the robustness of our findings, using different estimation methods. In both datasets we find evidence to suggest that societal religiosity has a negative effect on population health.

Second, while our primary objective is to examine the total effect of religiosity on health at the macro level, we also examine whether public health expenditures are a mechanism through which societal religiosity affects population health. Using a subsample of the second dataset, we find evidence for this mechanism: societal religiosity appears to have a negative indirect effect on population health via public health expenditures. However, while our estimates suggest that this indirect effect accounts for a substantial proportion of the total effect of religiosity on health, it is nevertheless much smaller than the total effect implied by our estimates. The conclusion that can be drawn from this is that there are also other channels through which population health is affected by religiosity. Medical technological progress may be such a channel, as argued above. Unfortunately, data on a direct measure of medical progress are not available for a sufficiently large number of countries to also examine the indirect effect of religiosity on health through medical progress.

Third, although the reverse effect of life expectancy on religiosity is not the focus of this paper, we test whether life expectancy is exogenous—not determined by religiosity. As argued by McCleary & Barro, 2006a, models that focus on the role of salvation and the after-life, such as Azzi & Ehrenberg, 1975,[2] predict that increased life expectancy should decrease religiosity. This prediction follows if the perceived probability of salvation depends on cumulated religious effort, including attendance at formal services and personal prayer, during one's lifetime, and if individuals postpone religious "outlays" (and thus religious intensity) until later in life, when they are closer to death. Under these conditions, increases in life expectancy imply that the share of the population that is concerned about what happens after death—and thus the share of the population that is religious—should decrease. In testing this prediction, McCleary & Barro, 2006a, found in a simple regression model with the church attendance rate as the dependent variable that this variable is positively correlated with the reciprocal of life expectancy at age one. However, this correlation does not necessarily imply that improvements in life expectancy cause declines in religiosity; it may also reflect the negative effect on life expectancy of religiosity. In fact, our results suggest that while societal religiosity affects life expectancy, life expectancy is not significantly affected by societal religiosity.

The rest of this paper is organized as follows: In Section 2, we examine the macro relationship between religiosity and health using our panel of 17 countries spanning the period 1925–2000. In Section 3, we examine the religiosity-health nexus using our panel of (up to) 92 countries across the years 1981–2016. Our conclusions are presented in Section 4.

2 | RESULTS FROM A PANEL OF 17 COUNTRIES FROM 1925 TO 2000

This section proceeds as follows: First, we present the basic empirical model and discuss some econometric issues. Then, we describe the data. Finally, we report and interpret the empirical results.

2.1 | Model and econometric issues

The basic model used in this analysis is a bivariate model of the form

$$\log(LIFE_{it}) = \beta \log(CHURCH_{it}) + c_i + \rho_i F_t + \varepsilon_{it} \quad (1)$$

where i and t are country and time indices; $LIFE_{it}$ is the most commonly used summary measure of population health, life expectancy at birth; and $CHURCH_{it}$ represents here our measure of societal religiosity, the rate of church attendance among children. Both variables are in logs so that the coefficient β can be interpreted as the elasticity of population health with respect to country-wide religiosity.[3] Finally, the c_i are unobserved time invariant country fixed-effects; and the term

$\rho_i F_t$ represents effects of unobserved time-varying common factors that, if not controlled, can induce cross-sectional dependence in the regression error, ε_{it}, and lead to inconsistent estimates.

Common factors, F_t, may be stationary or have stochastic and/or deterministic trends. We allow the common factors to have both unit roots and deterministic trends. To control for deterministic trend factors, we include a time trend whose coefficient is allowed to vary across countries, so that it can be interpreted as the country-specific effect of any omitted variables that follow a linear (or log-linear) trend.[4] To control for potentially stochastically non-stationary factors, we demean the data by subtracting the average value of $x_t = \left(\sum_{i=1}^{N} x_{it}\right)/N$ from each x_{it} in each period t, $x_{it} - \left(\sum_{i=1}^{N} x_{it}\right)/N$, and use the demeaned data in place of the original data (which is equivalent to using the residuals from regressions of each variable on time dummies).[5]

Since both the data (in logs) on life expectancy at birth and the data on church attendance exhibit trends that in most countries are not clearly linear (as shown in Figure 1), it is reasonable to assume that both $\log(LIFE_{it})$ and $\log(CHURCH_{it})$ have stochastically non-stationary components. If this assumption is correct, ε_{it} must (nevertheless) be stationary for Equation (1) to be a meaningful regression; in the terminology of Engle and Granger (1987), the requirement for the regression not to be spurious is that the two variables cointegrate.[6]

As is well known, and discussed in the Supplementary Appendix (Section A.2), cointegration between two (or more) variables implies not only the existence of a meaningful long-run relationship between these variables, but also that there are no relevant omitted variables in the relationship between these variables. Perman & Stern, 2003, therefore, argue that cointegration tests are also tests for misspecification of regression equations, which justifies the use of a bivariate model—if the variables are cointegrated. Nevertheless, to control for omitted variables that may be cointegrated with those in the model, we check the robustness of our results to the inclusion of additional variables. These variables include: education, measured by the log of average years of schooling in the population aged 15–64, $\log(YSCHOOL_{it})$, the level of economic development, measured by the log of real gross domestic product (GDP) per capita (PPP), $\log(GDPPC_{it})$, and fertility, measured by the log of the crude birth rate, $\log(CBR_{it})$.

Another advantage of the presence of cointegration is that endogeneity does not lead to inconsistency in the regression coefficients. However, although even the standard fixed-effects ordinary least squares (OLS) estimator is (super) consistent under panel cointegration even when the regressors are endogenous, it suffers from a second-order asymptotic bias in the presence of endogeneity, and, as a consequence, its t-statistic is not asymptotically standard normal. Hence,

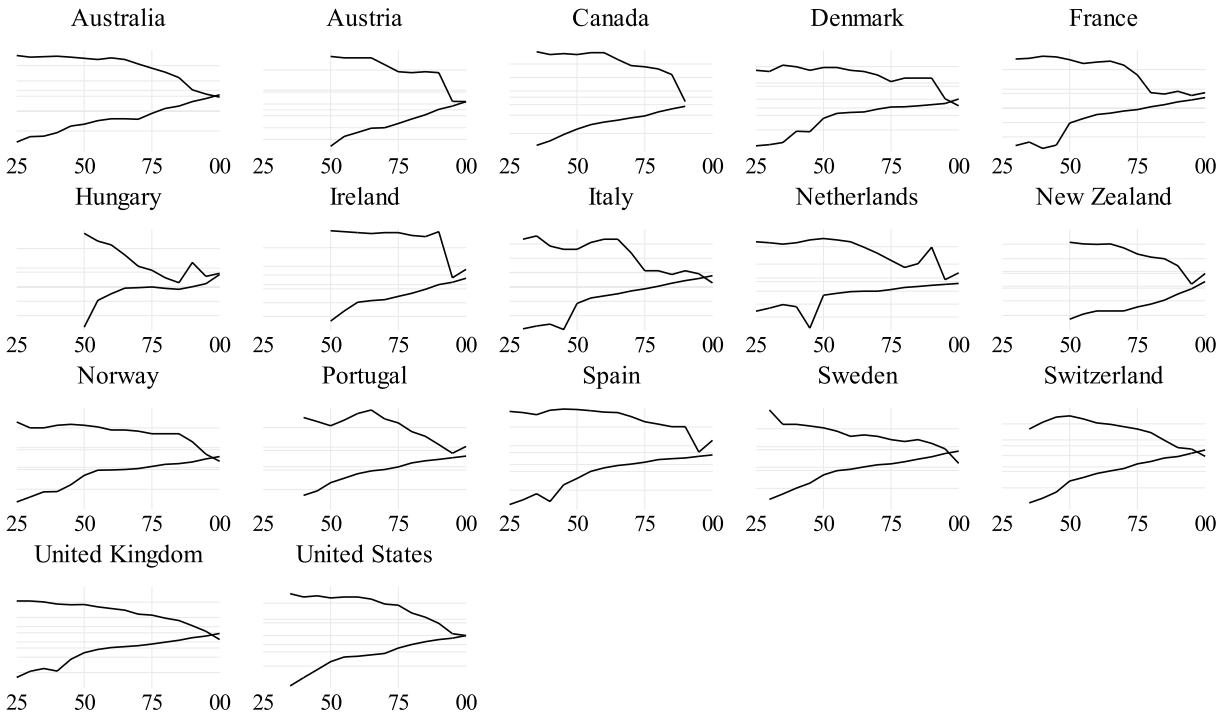

FIGURE 1 Log of life expectancy at birth and log of the rate of church attendance among children between 1925 and 2000. *Note*: Upper lines are church attendance rates of children (in logs); lower lines are life expectancies at birth (in logs)

we employ Pedroni's, 2001, panel fully modified OLS (PFMOLS) estimator, whose *t*-statistic is asymptotically standard normal even when the regressors are endogenous.[7]

Finally, an attractive feature of cointegration is that it implies long-run Granger causality in at least one direction (e.g., Granger, 1988).[8] However, cointegration says nothing about the direction of the long-run causal relationship between the variables. Therefore, we explicitly test for long-run Granger causality between $\log(LIFE_{it})$ and $\log(CHURCH_{it})$ using error correction models (discussed in more detail in Section 2.3.2).[9]

2.2 | Data

Because a detailed description of the data is provided in the Supplementary Appendix (Section A1), we only briefly describe the data and their sources here. The data on life expectancy at birth are from the Human Mortality Database (HMD).[10] The data on the rate of church attendance among children are from Iannaccone, 2003, who used retrospective questions in the 1991 and 1998 International Social Survey Program (ISSP) surveys to estimate weekly church attendance rates for 32 countries at 5 year intervals between 1925 and 1990 for respondents when they were 12 years old and their parents. We use the childhood rate of church attendance because it is reasonable to assume that respondents report their own church attendance during childhood more accurately than that of their parents and because the parental rate is based on a non-random sample of parents (as discussed in more detail in the Supporting Information). We updated these data to 2000 using information from the most recent (2008) ISSP survey. Combining our and Iannaccone's church attendance data with the life expectancy data from the HMD yields an unbalanced panel covering 17 predominantly Christian countries at five-year intervals between 1925 and 2000. Thus, the obvious advantage of this sample is that it allows us to examine the relationship between societal religiosity and population health over a long period of time (admittedly for a limited number of countries).

The series for the log of life expectancy at birth and the log of the rate of church attendance among children are plotted for each country in our panel in Figure 1 (which shows that while life expectancy has increased over time in all sample countries, church attendance has declined in all sample countries during the 20th century). Sources of data for the control variables are shown in Table A2 in the Supplementary Appendix.

Finally, before proceeding to the results, we note two things. First, there are several measures of religiosity at the societal or macro level (such as the percentage of people who believe in hell, the percentage of people who believe in heaven, and the percentage of people who pray regularly). If the term "societal religiosity" describes how religious a society is, the best available measure of societal religiosity is certainly the percentage of the population who identify themselves as religious. Unfortunately, data on this preferred measure are not available in the Iannaccone dataset, so that we are forced to rely on church attendance (among children) in the analysis in this section. The percentage of religious people in the population is used in the analysis in Section 3.

Second, church attendance is the most commonly used measure of religiosity. However, while Christians attend worship services as an integral part of their religious practice, other religions such as Judaism and Buddhism (but also Islam) are less, or not at all, organized around regular worship services. Consequently, non-Christians tend to have lower church attendance rates than Christians (e.g., Smith et al., 2002). Thus, church attendance does not adequately capture the religiosity of non-Christians and is therefore not a consistent measure of religiosity across countries with different religions.[11] For this reason, we include here only countries that are predominantly Christian,[12] as is common practice. The sample used in Section 3 includes all countries with available data.

2.3 | Results

The pre-tests for unit roots and cointegration, which are reported in the Supplementary Appendix (Section A2), suggest that the variables are non-stationary and cointegrated, as assumed in Equation (1). In this subsection, we provide estimates of the cointegrating relationship between life expectancy at birth and the rate of child church attendance and investigate the direction of causality between the two variables.

2.3.1 | Estimates of the long-run relationship between life expectancy at birth and the rate of child church attendance

Column 1 of Table 1 reports the PFMOLS estimate of Equation (1). The estimated coefficient on $\log(CHURCH_{it})$ is negative and significant at the 5% level. More specifically, the point estimate implies (if viewed causally) that, in the long run, a 1% decrease in the church attendance rate increases life expectancy by 0.0219%.

To get an idea of the magnitude of this effect, we multiply the coefficient on $\log(CHURCH_{it})$ by the average growth rate of the rate of Church attendance and dividing it by the average growth rate of life expectancy. The resulting value is −0.0867, implying that declining religiosity has been responsible for about 9% of the increase in life expectancy between 1925 and 2000. Thus, the PFMOLS point estimate implies a substantial, but not implausibly large, macro-level effect of religiosity on health.

In Column 2, we check the robustness of the negative relationship between $\log(LIFE_{it})$ and $\log(CHURCH_{it})$ to the inclusion of $\log(YSCHOOL_{it})$, $\log(GDPPC_{it})$, and $\log(CBR_{it})$. The coefficient on $\log(CHURCH_{it})$ remains negative and significant.

Turning briefly to the control variables, $\log(YSCHOOL_{it})$ is significant, but has an unexpected negative sign. The most likely reason for this is that the PFMOLS estimator requires that the regressors are not cointegrated. If this requirement is violated, the coefficient(s) on the cointegrated regressor(s) may be biased.[13] In contrast, the parameter estimate for income is significant with the expected sign. Finally, the coefficient on the fertility variable has the expected negative sign, but is not statistically significant.

It should again be noted here that our estimates of the long-run relationship between church attendance and life expectancy control for unobserved common and country-specific time effects (as well as observed, time-varying factors that may affect both life expectancy and religiosity) and that our cointegration tests indicate that there is a non-spurious long-run relationship between church attendance and life expectancy. Thus, we can be quite confident that the estimated negative long-run relationship between church attendance and life expectancy at birth is not an artifact of the trends in the data.

Before closing this section, we note that the negative and significant long-run coefficient on church attendance is also robust to the use of alternative estimation techniques, as shown in the Supplementary Appendix (Section A3).

2.3.2 | Causality

To test the direction of long-run Granger causality, we follow common practice and employ a two-step procedure. In the first step, the PFMOLS estimate of the coefficient β is used to construct an error-correction term $EC_{it} = \log(LIFE_{it}) + 0.0219\log(CHURCH_{it}) - \hat{c}_i - \hat{\rho}_i F_t$. In the second step, this term (lagged one period) is used to estimate error-correction equations for $\Delta\log(LIFE_{it})$ and $\Delta\log(CHURCH_{it})$:

$$\Delta\log(LIFE_{it}) = c_{1i} + \alpha_1 EC_{it-1} + \sum_{j=1}^{k} b_{11j}\,\Delta\log(LIFE_{it-1}) + \sum_{j=1}^{k} b_{12j}\,\Delta\log(CHURCH_{it-1}) + \varepsilon_{1it} \quad (2)$$

TABLE 1 PFMOLS regression results

	(1)	(2)
$\log(CHURCH_{it})$	−0.0219[b] (−2.023)	−0.0173[a] (−2.849)
$\log(YSCHOOL_{it})$		−0.0929[a] (−5.417)
$\log(GDPPC_{it})$		0.0721[a] (8.522)
$\log(CBR_{it})$		−0.0061 (−0.620)
No. of obs.	221	221

Note: The dependent variable is $\log(LIFE_{it})$. The PFMOLS estimator was computed allowing for heterogeneous first-stage coefficients. The PFMOLS results are based on demeaned data to account for potential error cross-sectional dependence. The regressions include both individual fixed-effects and individual time trends. *t*- or *z*-statistics are in parentheses and are based on heteroskedasticity- and autocorrelation-consistent standard errors.

Abbreviations: GDP, gross domestic product; PFMOLS, panel fully modified OLS.

[a]Significant at the 1% level.

[b]Significant at the 5% level.

$$\Delta\log(CHURCH_{it}) = c_{2i} + \alpha_2 EC_{it-1} + \sum_{j=1}^{k} b_{21j}\,\Delta\log(LIFE_{it-1}) + \sum_{j=1}^{k} b_{22j}\,\Delta\log(CHURCH_{it-1}) + \varepsilon_{2it} \quad (3)$$

If α_1 is nonzero and α_2 is zero, then long-run Granger causality runs from $\Delta\log(CHURCH_{it})$ to $\Delta\log(LIFE_{it})$. If, in contrast, α_2 is nonzero and α_1 is zero, then long-run Granger causality runs from $\Delta\log(LIFE_{it})$ to $\Delta\log(CHURCH_{it})$. Finally, if both α_1 and α_2 are nonzero, then long-run Granger causality is bidirectional.

Table 2 presents the results of the causality tests based on a lag length of $k = 1$,[14] which suggest that long-run Granger causality is unidirectional from church attendance to life expectancy. In this context, we note that we also find no evidence of short-run Granger causality from life expectancy at birth to church attendance and no evidence of joint short- and long-run Granger causality from $\log(LIFE_{it})$ to $\log(CHURCH_{it})$: the t-statistic on the coefficient on $\Delta\log(LIFE_{it-1})$ in Equation (3) is only −0.377, and a Wald test does not reject the null hypothesis of joint insignificance of $\Delta\log(LIFE_{it-1})$ and EC_{it-1} in Equation (3) with a p-value of 0.751. This suggests that $\log(CHURCH_{it})$ is strictly exogenous with respect to $\log(LIFE_{it})$, implying that the estimated coefficient on $\log(CHURCH_{it})$ in Column 1 of Table 1 can be interpreted as the long-run effect of $\log(CHURCH_{it})$ on $\log(LIFE_{it})$ (e.g., Stock & Watson, 2011). We thus can be quite confident that this coefficient reflects the total effect of macro-religiosity on population health, which is our primary focus. Unfortunately, because of the unavailability of historical data on public health expenditures, we are not able to use a (system) cointegration approach to also examine the effect of religiosity on health through public health expenditures.[15] We examine this effect in the next section.

3 | RESULTS FROM A PANEL OF UP TO 92 COUNTRIES BETWEEN 1981 AND 2016

This section uses panel data for up to 92 countries between 1981 and 2016 to examine the macro-level effect of religiosity on health, including the exogeneity of religiosity, and to test whether public health expenditures act as a channel through which higher religiosity reduces the level of health. Following the procedure in the previous section, we first present our basic estimating equations and discuss some econometric issues. Then, we discuss the data used in this analysis. Finally, we report and discuss our results.

3.1 | Models and econometric issues

The basic empirical model used to estimate the macro-level effect of religiosity on health in this analysis is as follows:

$$\log(LIFE_{it}) = \beta\log(PERSON_{it}) + \gamma X_{it} + \lambda_t + c + \varepsilon_{it} \quad (4)$$

where $\log(LIFE_{it})$ is, as before, the log of life expectancy at birth and $\log(PERSON_{it})$ is the log of the proportion of the population who identify themselves as a religious person. The advantage of this measure of societal religiosity over the rate of church attendance is that it is also appropriate for non-Christian religions where attendance at places of worship is relatively unimportant (e.g., Deaton, 2011; McCleary & Barro, 2006b). In addition, as previously mentioned, it is logical to assume that this measure best captures the concept of societal religiosity.

TABLE 2 Long-run causality tests

	(1) Dep. var.: Δlog ($LIFE_{it}$)	(2) Dep. var.: Δlog ($CHURCH_{it}$)
Coefficient of the error-correction term, α_1	−0.702[a] (−8.144)	
Coefficient of the error-correction term, α_2		−0.261 (−0.416)
No. of obs.	204	204

Note: The results are based on demeaned data to account for error cross-sectional dependence. All regressions include individual fixed-effects. Numbers in parentheses are heteroskedasticity- and autocorrelation-consistent t-statistics.

[a]Significant at the 1% level.

X_{it} is a vector of control variables that includes the log of the (gross) secondary school enrollment ratio, log($SSCHOOL_{it}$), the log of real GDP per capita (PPP), log($GDPPC_{it}$), and the log of the fertility rate, log($FERT_{it}$). Since these control variables are not available for all countries and years in our sample (and therefore including these control variables reduces our sample size), we also estimate regressions without control variables.

To control for unobserved common factors (and thus for potential cross-sectional dependence), we include year fixed-effects, λ_t, in our model. Because including year fixed-effects requires at least two observations per year and thus results in an implicit loss of single-year observations (five in our case), we also estimate models without year fixed-effects.

Similarly, country fixed-effects models require at least two observations on each country, implying that countries with only one observation are implicitly dropped in the estimation of the slope parameters. In addition, if the number of time-series observations is small and the number of cross-sectional observations large, including country fixed-effects substantially reduces the number of degrees of freedom. Our panel dataset (discussed in more detail below) contains many countries but few time-series observations on each country, and there are several countries with only one observation in the dataset. Therefore, we do not include country fixed-effects in our core specification, but we also estimate models with country fixed-effects.[16]

We estimate our model using OLS. As is well known, the pooled OLS estimator (without country fixed-effects)[17] is inconsistent when the regressors are endogenous. To control for possible endogeneity, we run instrumental variables (IV) regressions where we instrument for (the log of) the proportion of religious persons with the percentage of the population aged 0–14, $POP0\_14_{it}$,[18] and/or the proportion of individuals who have used the Internet in the last three months, $INTERNET_{it}$.[19] The instruments are discussed in more detail later.

Finally, we examine the hypothesis that religiosity affects health through its effect on public health expenditures, $HEALTH_{it}$, measured as a percentage of GDP. To test this hypothesis, we estimate several regressions in which the dependent variables are log($HEALTH_{it}$) and log($LIFE_{it}$).

In the regressions with log($HEALTH_{it}$) as the dependent variable, we include log($PERSON_{it}$) as a regressor, so that the basic regression formula is:

$$\log(HEALTH_{it}) = \delta \log(PERSON_{it}) + \gamma X_{it} + \lambda_t + c + \varepsilon_{it} \tag{5}$$

where X_{it} is again a vector of control variables. The selection of the control variables is based on the existing literature on the determinants of government size (e.g., Alesina & Wacziarg, 1998; Ram, 2009). In the above equation, the control variables are: the log of the urbanization rate, measured by the logged percentage of the population living in metropolitan areas, log($URBAN_{it}$), the log of population size, log(POP_{it}), population density, measured by the number of people per square kilometer, $DENSITY_{it}$,[20] the log of trade openness, measured by the log of the percentage ratio of trade to GDP, log($OPEN_{it}$), the log of the age dependency ratio (defined as the ratio of people younger than 15 and older than 64 to the working-age population (15–64)), log(DEP_{it}), and the log of real GDP per capita (PPP), log($GDPPC_{it}$). Since the inclusion of the control variables reduces our sample size, we also run the regression without control variables.

In the regressions with log($LIFE_{it}$) as the dependent variable, we include log($HEALTH_{it}$) as a regressor and use the same control variables as in model (4). Thus, the basic specification is as follows:

$$\log(LIFE_{it}) = \gamma \log(HEALTH_{it}) + \gamma X_{it} + \lambda_t + c + \varepsilon_{it} \tag{6}$$

We estimate Equations (5) and (6) using both OLS and IV.[21] log($PERSON_{it}$) is instrumented with the same instruments as above ($INTERNET_{it}$ and $POP0\_14_{it}$). To instrument for log($HEALTH_{it}$), we use one and two period lagged values of public health expenditures as a percentage of GDP and the average of military expenditures as a percentage of GDP of a country's neighbors; the latter is a commonly used instrumental variable (e.g., Bokhari et al., 2007; Filmer & Pritchett, 1999).

3.2 | Data

Data on the control and instrumental variables are taken from the World Development Indicators (WDI);[22] the life expectancy data are also from the WDI; and the data to measure the percentage of religious persons come from six waves of the World Values Survey.[23] The percentage of people who responded "*Yes*" to the question "*Independently of whether you go to Church or not, would you say you are a religious person?*" is our measure of this variable. Putting together the data on $LIFE_{it}$ and $PERSON_{it}$ from these sources yields a dataset covering 92 countries and 27 years between 1981 and 2016.

TABLE 3 Countries in the 1981–2016 panel dataset

Country	Year(s) in which the WVS survey was conducted	Country	Year(s) in which the WVS survey was conducted	Country	Year(s) in which the WVS survey was conducted
Albania	1998	Haiti	2016	Poland	1997, 2005, 2012
Algeria	2002, 2014	Hong Kong	2005, 2013	Puerto Rico	1995, 2001
Argentina	1984, 1991, 1995, 1999, 2006, 2013	Hungary	1982, 1998, 2009	Romania	1998, 2005, 2012
Armenia	1997, 2011	India	1990, 1995, 2001, 2006, 2012	Russia	1990, 1995, 2006, 2011
Australia	1981, 1995, 2005, 2012	Indonesia	2001, 2006	Rwanda	2007, 2012
Azerbaijan	1997, 2012	Iran	2000, 2005	Saudi Arabia	2003
Bangladesh	1996, 2002	Iraq	2006, 2013	Serbia	1996, 2001, 2005
Belarus	1990, 1996, 2011	Italy	2005	Singapore	2001, 2012
Bosnia and Herz.	1998, 2001	Japan	1981, 1990, 1995, 2000, 2005, 2010	Slovak Rep.	1990, 1998
Brazil	2006, 2014	Jordan	2001, 2014	Slovenia	1995, 1998, 2005, 2011
Bulgaria	1997, 2005	Kazakhstan	2011	South Africa	1982, 1996, 2001, 2006, 2013
Burkina Faso	2007	Korea, Rep.	1982, 2001, 2006, 2010	Spain	1990, 1995, 2000, 2007, 2011
Canada	2000, 2005	Kyrgyz Rep.	2003, 2011	Sweden	1981, 1996, 1999, 2006, 2011
Chile	1990, 1996, 200, 2011	Latvia	1996	Switzerland	1989, 1996, 2007
China	1990, 2001, 2007, 2012	Lebanon	2013	Tanzania	2001
Colombia	1997, 2005, 2012	Libya	2013	Thailand	2007, 2013
Croatia	1996	Lithuania	1997	Trin. and Tobago	2006, 2010
Cyprus	2006, 2011	Malaysia	2006, 2011	Tunisia	2013
Czech Republic	1991, 1995	Mali	2007	Turkey	1990, 1996, 2001, 2007, 2011
Dom. Rep.	1996	Mexico	1981, 1990, 1996, 2000, 2005, 2012	Uganda	2001
Ecuador	2013	Moldova	1996, 2002, 2006	Ukraine	1996, 2006, 2011
Egypt	2001, 2008	Montenegro	1996, 2001	United Kingdom	2005
El Salvador	1999	Morocco	2001, 2007	United States	1981, 1995, 1999, 2006, 2011
Estonia	1996, 2011	Netherlands	2005, 2012	Uruguay	1996, 2006, 2011
Ethiopia	2007	New Zealand	1998, 2004, 2011	Uzbekistan	2013
Finland	1981, 1996, 2005	Nigeria	1990, 1995, 2000, 2011	Venezuela	1996, 2000
France	2006	N. Macedonia	1996, 2007	Vietnam	2000, 2005
Georgia	1996, 2008, 2014	Norway	1996, 2007	Yemen, Rep.	2010
Germany	1997, 2006, 2013	Pakistan	1990, 1998	Zambia	2005
Ghana	2007, 2011	Peru	1996, 2001, 2012	Zimbabwe	2000, 2010
Guatemala	2004	Philippines	1996, 2001, 2012		

These countries and the years for which we have data are reported in Table 3. The maximum number of observations per country in the dataset is six, most countries (35) have two observations, and 26 countries have only one observation (implying that including country fixed-effects is problematic for the reasons discussed above). Descriptive statistics on the variables for this panel sample are presented in Table A7 in the Supplementary Appendix.

Public health expenditures (as a percentage of GDP) are also from the WDI. To complement the data from the WDI, we also use data on public health expenditures from the Government Finance Statistics.[24] Merging the data on $HEALTH_{it}$ and $PERSON_{it}$ yields a subsample of 85 countries.[25] Descriptive statistics on $HEALTH_{it}$ and $PERSON_{it}$ for this subsample are also presented in Table A7 in the Supplementary Appendix.

Finally, for the regressions of $\log(LIFE_{it})$ on $\log(HEALTH_{it})$, we use data for a subgroup of 81 countries;[26] these data cover the period 1981–2019 (the last year for which life expectancy data are currently available).

3.3 | Results

3.3.1 | Results of regressions with life expectancy as the dependent variable and the percentage of religious people as an independent variable

The results of Equation (4) are presented in Column 1 of Table 4. The control variables have the expected signs and are significant, with the exception of the log of the fertility rate, $\log(FERT_{it})$, which is nearly significant at the 10% level (the p-value is 0.115).

The measure of macro-religiosity, the log of the percentage of religious persons, is highly significant and has a negative coefficient. This coefficient implies that a 1% decrease in the percentage of religious people increases life expectancy by 0.0428%.

To evaluate the magnitude of this effect, consider a decrease in the percentage of religious people from 81.4%, the position of the United States in 1981%, to 67.0%, the position of the United States in 2011. This is predicted to increase life expectancy at birth by $((81.4 - 67) \times 0.0428 =)$ 0.616%. Life expectancy in the United States was 74.01 years in 1981 and 78.64 years in 2011, implying an increase in life expectancy by 6.256%. From this, it follows that about 10% of the increase in life expectancy between 1981 and 2011 can be attributed to the decline in religiosity. Thus, we again find a substantial, but not implausibly large, macro-level effect of religiosity on health.

The results in Column 1 are based on a relatively small number of observations (152). A potential concern could therefore be that these observations are not representative of all observations, and that the significant negative coefficient on $\log(PERSON_{it})$ is, therefore, due to the exclusion of relevant observations. To address this concern, we reestimate Equation (4) for the full sample, without the control variables. The results, which are presented in Column 2 of Table 4, show that the coefficient on $\log(PERSON_{it})$ remains negative and significant at the 1% level.

In Column 1 of Table 5, we report results without year and country fixed-effects. Column 2 of Table 5 shows results from a regression that includes country fixed-effects, but no year fixed-effects. Finally, Column 3 presents coefficient estimates from a model with both year and country fixed-effects. As can be seen, the coefficient estimates on $\log(PERSON_{it})$ are significant and negative across all specifications.[27]

TABLE 4 Year fixed-effects regression results

	(1)	(2)
$\log(PERSON_{it})$	−0.0428[a] (−3.547)	−0.1152[a] (−4.999)
$\log(SSCHOOL_{it})$	0.1186[a] (3.109)	
$\log(GDPPC_{it})$	0.0334[a] (3.573)	
$\log(FERT_{it})$	−0.0395 (−1.587)	
No. of obs.	152	217
No. of countries	74	92
Adjusted R^2	0.643	0.225

Note: The dependent variable is $\log(LIFE_{it})$. Both regressions include a constant term. Numbers in parentheses are t-statistics based on heteroskedasticity-consistent and degrees-of-freedom-adjusted standard errors.

Abbreviation: GDP, gross domestic product.

[a]Significant at the 1% level.

TABLE 5 Pooled and country fixed-effects models

	(1)	(2)	(3)
log($PERSON_{it}$)	−0.0379[a] (−4.317)	−0.0515[a] (−4.197)	−0.0309[b] (−2.380)
log($SSCHOOL_{it}$)	0.1088[a] (3.164)	0.0959[a] (4.082)	0.1103[a] (4.533)
log($GDPPC_{it}$)	0.0403[a] (5.004)	0.0647[a] (5.373)	0.0198 (1.104)
log($FERT_{it}$)	−0.0326 (−1.569)	−0.0032 (−0.162)	0.0238 (1.091)
Year fixed-effects	No	No	Yes
Country fixed-effects	No	Yes	Yes
No. of obs.	152	152	152
No. of countries	74	74	74
Adjusted R^2	0.652	0.980	0.985

Note: The dependent variable is log($LIFE_{it}$). The regression in column (1) includes a constant term. Numbers in parentheses are *t*-statistics based on heteroskedasticity-consistent and degrees-of-freedom-adjusted standard errors.

Abbreviation: GDP, gross domestic product.

[a]Significant at the 1% level.

[b]Significant at the 5% level.

3.3.2 | Controlling for endogeneity

We use IV and two-stage least squares (2SLS) to control for potential endogeneity. Instruments must be both strong and valid for the 2SLS method to produce consistent and efficient estimates. Such instruments are sufficiently strongly correlated with the potentially endogenous variables, but are uncorrelated with the error term and thus do not affect the dependent variable except through the endogenous variable.

The idea behind the use of $INTERNET_{it}$ as an instrument for log($PERSON_{it}$) is that the Internet provides opportunities to find information about people of other religions and non-religious people (and to interact with them personally), and that this has a negative effect on religiosity—an idea which is supported by the studies of Downey, 2014, and McClure, 2020, as well as by our first-stage regression of log($PERSON_{it}$) on $INTERNET_{it}$. More specifically, the coefficient on $INTERNET_{it}$ is negative and significant in the religiosity regression in Column 1 of Table 6, and the partial R^2 for this regression (conditioning on year effects) is 0.157, suggesting that $INTERNET_{it}$ is a strong instrument. In contrast, the coefficient on $INTERNET_{it}$ is insignificant in the life expectancy regression in Column 2. Although the latter is not a direct test of whether our instrument is valid, it provides at least suggestive evidence that $INTERNET_{it}$ can be used as a valid instrument for log($PERSON_{it}$).[28] Based on this instrument, we use 2SLS to estimate the effect of our religiosity measure on life expectancy. The results in Column 3 show that the IV estimate of this effect is negative and statistically significant, albeit less significant than the OLS effect. As expected, the first-stage *F*-test, also reported in Column 3, indicates that the instrument $INTERNET_{it}$ is highly correlated with the endogenous variable.

As an additional instrument, we use the percentage of the population aged 0–14, $POP0\_14_{it}$. The idea is that religious organizations provide services that especially benefit adults with young children (including moral guidance and instruction, child care, social events, and networks of like-minded parents), and that these services provide a positive incentive for parents (and their children) to participate in religious activities and thereby lead to increased religiosity. If this is correct, then it is reasonable to hypothesize that the higher the percentage of the population aged 0–14, the higher the percentage of religious people. The first-stage regression results in Column 4 of Table 6 support this hypothesis: $POP0\_14_{it}$ is significant and positive in the regression equation for our religiosity variable. Since $INTERNET_{it}$ remains negative and significant in this regression, and the partial R^2 is 0.166, we conclude that the instruments are sufficiently strong in explaining the variation of log($PERSON_{it}$), a conclusion which is also supported by the first-stage *F*-test reported in Column 6. In Column 6, we also report Wooldridge's, 1995, heteroscedasticity-robust score test of overidentifying restrictions. With a *p*-value of 0.127, this test does not reject the null that the excluded instruments are exogenous to the log of life expectancy at birth. However, this test (like other tests of overidentifying restrictions) assumes that at least one instrument is valid. To provide suggestive evidence that both instruments are valid, we (again) include the instruments as regressors in the OLS model. The results of this exercise are reported in Column 5: the coefficients on $INTERNET_{it}$ and $POP0\_14_{it}$ are statistically insignificant, suggesting that both variables are valid instruments.[29] Using these variables as instruments, we again find a negative effect of log($PERSON_{it}$) on log($LIFE_{it}$), although this effect is only significant at

TABLE 6 IV analysis

	IV analysis using $INTERNET_{it}$ as an instrument for log($PERSON_{it}$)			IV analysis using $INTERNET_{it}$ and $POP0\_14_{it}$ as instrumental variables		
	(1)	(2)	(3)	(4)	(5)	(6)
	First-stage reg. Dep. var.: log ($PERSON_{it}$)	Dep. var.: log ($LIFE_{it}$)	IV regression Dep. var.: log ($LIFE_{it}$)	First-stage reg. Dep. var.: log ($PERSON_{it}$)	Dep. var.: log ($LIFE_{it}$)	IV regression Dep. var.: log ($LIFE_{it}$)
$INTERNET_{it}$	−0.0093[a] (−8.531)	0.0004 (0.898)		−0.0056[a] (−3.354)	0.0005 (1.102)	
$POP0\_14_{it}$				0.0163[a] (3.538)	0.0030 (1.155)	
log($PERSON_{it}$)		−0.0383[a] (−2.657)	−0.0760[b] (−2.084)		−0.0405[a] (−2.749)	−0.0631[c] (−1.660)
log($SSCHOOL_{it}$)		0.1185[a] (3.076)	0.1243[a] (3.392)		0.1206[a] (3.136)	0.1219[a] (3.410)
log($GDPPC_{it}$)		0.0263[b] (2.527)	0.0319[a] (3.530)		0.0295[a] (2.724)	0.0325[a] (3.6740)
log($FERT_{it}$)		−0.0426 (−1.633)	−0.0218 (−0.797)		−0.0933[c] (−1.894)	−0.0289 (−1.0475)
Wooldridge's (1995) robust score test of overidentifying restrictions (p-value)						0.127
Robust F-test of instrument relevance (p-value)			0.001			0.005
Wooldridge's (1995) robust score test of exogeneity (p-value)			0.330			0.579
Wooldridge's (1995) robust regression-based test of exogeneity (p-value)			0.371			0.615
No. of obs.	149	149	149	149	149	149
No. of countries	74	74	74	74	74	74
Adjusted R^2	0.200	0.647	0.636	0.260	0.648	0.643

Note: All regressions include a constant term. All regressions include year fixed-effects. Numbers in parentheses are t-statistics based on heteroskedasticity-consistent and degrees-of-freedom-adjusted standard errors.

Abbreviation: GDP, gross domestic product.

[a]Significant at the 1% level.

[b]Significant at the 5% level.

[c]Significant at the 10% level.

the 10% level, as can be seen in Column 6 of Table 6.[30] In addition, we find that the IV estimates are somewhat larger (in absolute value) than the OLS estimates.

The most likely explanation for the discrepancy between the OLS and IV results is that 2SLS is less efficient than OLS when the explanatory variables are exogenous and that $\log(PERSON_{it})$ is exogenous. In fact, Wooldridge's, 1995, heteroskedasticity-robust score and regression-based tests of exogeneity reported in Columns 3 and 6 show that the null hypothesis of exogeneity of $\log(PERSON_{it})$ is not rejected. Therefore, we prefer the OLS estimates in Table 5.[31]

As we describe in the Supplementary Appendix (Section A7), we also estimate 2SLS models with $\log(LIFE_{it})$ and $\log(PERSON_{it})$ as the dependent (and independent) variables using the heteroskedasticity-based instrumental variable estimator proposed by Lewbel (2012)—an estimator that constructs internal instruments using the available regressors by exploiting model heteroskedasticity. Using this estimator we again find similar results, in that the instrumented religiosity variable has a negative and significant coefficient in the life expectancy regression, whereas instrumented life expectancy has no significant relationship with $\log(PERSON_{it})$.

3.3.3 | Public health expenditures as dependent and independent variable

We now examine whether religiosity affects health through its effect on public health expenditures. To answer this question, we first estimate Equation (5), using OLS and 2SLS. The OLS estimates with and without controls are reported in Columns 1 and 2 of Table 7, respectively. The coefficient on $\log(PERSON_{it})$ is negative and significant in both regressions. Similarly, the 2SLS estimate for the coefficient on $\log(PERSON_{it})$ is significantly negative, as can be seen in Column 3. However, while the instruments pass the tests of overidentifying restrictions and instrument relevance, the exogeneity tests do not reject the exogeneity of $\log(PERSON_{it})$. The implication is that OLS is preferable to 2SLS and that we can be quite confident that the estimated coefficients on $\log(PERSON_{it})$ in Columns 1 and 2 reflect mainly the effect of societal religiosity on public health expenditures, rather than vice versa.[32]

For brevity, we do not discuss the results for the control variables in detail here, but note that the coefficients of the control variables are largely consistent with the literature on the determinants of government expenditure. The only exception is the coefficient on $\log(OPEN_{it})$, which, in contrast to previous estimates in the literature, is negative and

TABLE 7 Public health expenditures as the dependent variable

	OLS		2SLS
	(1)	(2)	(3)
$\log(PERSON_{it})$	−0.2038[b] (−2.005)	−0.4664[a] (−2.901)	−0.5773[b] (−2.224)
$\log(URBAN_{it})$	−0.1199[c] (−1.789)		−0.1243[c] (−1.663)
$\log(POP_{it})$	−0.1610[a] (−4.152)		−0.1680[a] (−4.005)
$DENSITY_{it}$	−0.0001[b] (−2.384)		−0.0001[a] (−2.300)
$\log(OPEN_{it})$	−0.3613[a] (−3.539)		−0.3729[a] (−3.560)
$\log(DEP_{it})$	0.1071 (0.334)		0.4399 (1.159)
$\log(GDPPC_{it})$	0.5224[a] (7.849)		0.4973[a] (6.892)
Wooldridge's (1995) robust score test of overidentifying restrictions (*p*-value)			0.177
Robust *F*-test of instrument relevance (*p*-value)			0.000
Wooldridge's (1995) robust score test of exogeneity (*p*-value)			0.103
Wooldridge's (1995) robust regression-based test of exogeneity (*p*-value)			0.138
No. of obs.	137	159	136
No. of countries	71	85	71
Adjusted R^2	0.664	0.076	0.627

Note: The dependent variable is $\log(HEALTH_{it})$. All regressions include a constant term. All regressions include year fixed-effects. The instruments are $INTERNET_{it}$ and $POP0\_14_{it}$. Numbers in parentheses are *t*-statistics based on heteroskedasticity-consistent and degrees-of-freedom-adjusted standard errors.

Abbreviations: GDP, gross domestic product; OLS, ordinary least squares.

[a]Significant at the 1% level.

[b]Significant at the 5% level.

[c]Significant at the 10% level.

TABLE 8 Public health expenditures as a regressor

	OLS	2SLS	GMM-IV
	(1)	(2)	(3)
log($HEALTH_{it}$)	0.0075[b] (2.089)	0.0674[a] (5.222)	0.0142[b] (2.317)
[Long − run effect of log($HEALTH_{it}$)]			[0.1222[a] (3.312)]
log($LIFE_{it-1}$)			0.8835[a] (16.444)
log($GDPPC_{it}$)	0.0452[a] (14.593)	0.0418[a] (11.879)	−0.0037 (−1.221)
log($SCHOOLS_{it}$)	0.0600[a] (7.758)	0.0496[a] (5.477)	−0.0029 (−0.305)
log($FERT_{it}$)	−0.0571[a] (−7.545)	−0.0584[a] (−6.765)	−0.0088 (−1.160)
Wooldridge's (1995) robust score test of overidentifying restrictions (*p*-value)		0.1240	
Robust *F*-test of instrument relevance (*p*-value)		0.0000	
Wooldridge's (1995) robust score test of exogeneity (*p*-value)		0.0198	
Wooldridge's (1995) robust regression-based test of exogeneity (*p*-value)		0.0182	
AR2 (*p*-value)			0.594
HANSEN (*p*-value)			0.912
No. of instruments			38
No. of obs.	1279	1072	1278
No. of countries	81	79	81
Adjusted R^2	0.736	0.728	

Note: The dependent variable is log($LIFE_{it}$). All regressions include a constant term. All regressions include year fixed-effects. The instruments in the 2SLS model are one and two period lagged values of public health expenditures as a percentage of GDP and the average of military expenditures as a percentage of GDP of a country's neighbors. The long-run effect of log($HEALTH_{it}$) on log($LIFE_{it}$) is the coefficient on log($HEALTH_{it}$) divided by one minus the coefficient on log($LIFE_{it-1}$). AR2 is the Arellano-Bond test for second-order autocorrelation in differenced residuals. HANSEN is the Hansen test of overidentifying restrictions. Numbers in parentheses are *t*-statistics. Numbers in parentheses are *t*-statistics based on heteroskedasticity-consistent and degrees-of-freedom-adjusted standard errors.

Abbreviation: GDP, gross domestic product; GMM-IV, generalized methods of moments instrumental variable; OLS, ordinary least squares; SLS, stage least squares.

[a]Significant at the 1% level.

[b]Significant at the 5% level.

significant at the 1% level. A possible explanation for the negative coefficient could be that governments substitute health expenditures with other public expenditures in response to increased trade openness.

Next, we estimate Equation (6) to examine the relationship between public health expenditures and life expectancy at birth. The first column of Table 8 shows the OLS results. The coefficient on log($HEALTH_{it}$) is positive and significant, which is consistent with several recent studies that find evidence of health improvements due to increases in public health expenditures (e.g., Bernet et al., 2020; Bokhari et al., 2007). Of course, we have to be cautious about causal inferences. Therefore, to account for the likely endogeneity of public health expenditures, we also estimate Equation (6) using 2SLS. In addition, as a robustness check, we estimate a dynamic version of Equation (6) using the standard generalized methods of moments instrumental variable (GMM-IV) estimator of Blundell & Bond, 1998,[33] which is applicable here given the relatively large number of (repeated) observations in this particular sample. The 2SLS and GMM-IV results are reported in Columns 2 and 3, respectively.[34] As can be seen, the instruments pass the specification tests, and the exogeneity tests in Column 2 indicate that public health expenditures are endogenous. Therefore, it can be concluded with some confidence that the IV coefficients reflect a significant positive effect of public health expenditures on life expectancy at birth. It is perhaps needless to say that the estimates in Columns 2 and 3 are not directly comparable because of different specifications: As in Tables 4 to 7, the results in Column 2 are based on a static model, whereas the results in Column 3 are based on a dynamic model. Overall, the combined results in Tables 7 and 8 support our hypothesis that religiosity reduces life expectancy via its negative effect on public health expenditures.

It should be noted that we also experimented with the Lewbel (2012) IV approach to estimate Equations (5) and (6), but found that the generated instruments do not pass all required tests.

In order to roughly gauge the effect of religiosity on life expectancy at birth through its effect on public health expenditures, we consider the product of the coefficient on log($HEALTH_{it}$) in Column 2 of Table 8 and the coefficient on

log($PERSON_{it}$) in Column 1 of Table 7, −0.0137. Comparing this value with that of the coefficient on log($PERSON_{it}$) in Column 1 of Table 4 (−0.0428), we find that the indirect effect accounts for about 32% of the total effect of religiosity on health. While this this percentage is substantial, it nevertheless means that most of the total effect cannot be attributed to the indirect effect of religiosity on life expectancy at birth that operates through public health expenditures. Thus, there are certainly other channels through which societal religiosity affects population health: for example, through medical technological progress, as discussed in the Introduction.[35]

Lastly, a word of caution: It is a well-known fact that causal effects are difficult to identify—particularly using macro data. Therefore, the possibility that the significant associations found in this study reflect, to some degree, reverse causality from life expectancy (and public health expenditures) to religiosity should not be completely ruled out. However, given that (i) our long-run (and short-run) Granger causality tests suggest that temporal causality runs from church attendance to life expectancy but not from life expectancy to church attendance, (ii) we find a negative and significant coefficient for the percentage of religious people in both our OLS and IV regressions, (iii) our exogeneity tests and our Lewbel (2012) regressions suggest that life expectancy does not affect the percentage of religious people, and (iv) our analysis supports the hypothesis that religiosity affects life expectancy through public health expenditures, we are convinced that even if there is some reverse causality here, this bias is not large enough to overturn our conclusions regarding the effect of societal religiosity on population health.

4 | CONCLUSIONS

In this study, we examined the macro-level effect of religiosity on health using two panel datasets. Surprisingly, this issue has not yet been explored in the literature, which has focused on the effect of religiosity on health at the micro level.

Three main conclusions can be drawn from our results: First, although (or because) religious people tend to be healthier and live longer than less religious and non-religious people, as micro-level studies have shown, religiosity appears to have a negative effect on health at the macro level. We proposed two hypotheses to explain this finding: that societal religiosity negatively affects population health through adverse effects on medical technological progress, and that macro-religiosity negatively impacts population health through negative effects on public health expenditures.

Second, while we were not able to test the first hypothesis, our results support the second hypothesis—that societal religiosity has a negative effect on population health via public health expenditures. This effect appears to account for a substantial proportion of the total effect of societal religiosity on population health. However, our evidence also suggests that most of the total effect cannot be explained by this effect. From this it can be concluded that there are also other channels through which health is affected by religiosity at the macro level. A plausible candidate for such a channel is medical technological progress.

Third, the idea, suggested by McCleary & Barro, 2006a, that increased life expectancy reduces religiosity is not supported by our study, which suggests that societal religiosity affects life expectancy, but not vice versa.

Finally, from a public policy perspective, the finding of a negative effect of societal religiosity on population health implies that promoting secularization can improve the health of the population as a whole—if (as our results suggest) secularization is accompanied by an increase in public health expenditures and/or if secularization accelerates the development and use of new medical technologies (which is indirectly supported by our results). How public policy can promote secularization, however, is a separate question that is beyond the scope of this paper. Nevertheless, we can say that education is a potentially important policy variable that can negatively affect religiosity by providing knowledge that contradicts many religious claims and promoting critical thinking (e.g., Becker at al., 2017). This is supported by the studies of Arias-Vazquez (2012) and Becker at al. (2017) who found that education contributes to secularization. Since education may also impart knowledge about health and health behaviors, and since several studies have found positive effects of education on health (e.g., Brunello et al., 2016; Lleras-Muney, 2005; Ricci & Zachariadis, 2013), education policies might be powerful tools for improving population health.

ACKNOWLEDGMENTS

Thanks to a Co−Editor and an Associate Editor as well as two referees for their helpful comments and suggestions.

Open access funding enabled and organized by Projekt DEAL.

CONFLICT OF INTEREST

Dierk Herzer has nothing to disclose.

ETHICS STATEMENT

No ethical approval was required for this study.

DATA AVAILABILITY STATEMENT

The data that support the findings of this study are available from the corresponding author upon reasonable request.

ORCID

Dierk Herzer https://orcid.org/0000-0002-5310-0385

ENDNOTES

[1] To explain this phenomenon, Herzer & Strulik, 2020, developed a theoretical model in which it is assumed that reflective/analytical thinking improves research productivity, and that while religious individuals tend to use intuitive thinking, non-religious individuals tend to use reflective/analytical thinking. The authors showed that, under these assumptions, a decline in the number of religious people relative to the number of non-religious people leads to growth in TFP.

[2] While Azzi & Ehrenberg, 1975, did not explicitly consider life expectancy in their rational choice model, Papyrakis & Selvaretnam, 2011, extended their model to study the impact of life expectancy on religiosity explicitly.

[3] Table A12 in the Supplementary Appendix reports the results of regressions that consider the relationship between religiosity and health in level-level form (rather than in log-log form). The main results are qualitatively unchanged when using models with the unlogged level of life expectancy as the dependent variable and the unlogged level of religiosity as the independent variable.

[4] Several studies use a log-linear time trend as a proxy for medical technological change. However, direct measures of medical progress, such as (health) R&D expenditures or the number of approved new drugs, do not generally follow a log-linear trend (e.g., Okunade & Murthy, 2002). It is therefore reasonable to assume that the estimate of β captures the total effect of religiosity on health, including the effect that operates through medical progress, despite the inclusion of country-specific time trends. Moreover, the trends are statistically significant for 12 of the 17 countries in this sample, and the evidence in favor of cointegration is relatively weak when we exclude the trends, implying that country-specific time trends should be included.

[5] The implicit assumption behind the use of demeaned data is that the loadings of the common factors are homogeneous across the cross-sectional units and, consequently, that the responses to the common factors do not differ across the units. However, if this assumption does not hold, then the demeaning procedure may be ineffective in eliminating (error) cross-sectional dependence. Therefore, to check the robustness of our β estimate, we also use the so-called common correlated effects (CCE) approach of Pesaran, 2006; this approach allows for heterogeneous responses to unobserved common factors by including weighted cross-sectional averages of the dependent and independent variables. The CCE results are reported in the Supplementary Appendix (Section A3). It should be noted that the CCE method is designed for large N and large T. Although Westerlund et al., 2019, show that the pooled CCE estimator is consistent and asymptotically (mixed) normal when T is fixed and $N \to \infty$, it is unclear how large T should be to ensure reliable estimates when N is relatively small (as in the present application). Therefore, the CCE method is not preferred here, although the CCE results are qualitatively and quantitatively similar to those reported in Column 1 of Table 1.

[6] As shown by Kao, 1999, the tendency for spuriously indicating a relationship may even be stronger in panel data regressions than in pure time-series regressions. Interestingly, the standard fixed-effects OLS estimator of the regression coefficient is consistent for its true value (so that the point estimate of the regression coefficient converges to zero), but the t-statistic diverges as the cross-sectional dimension increases.

[7] The (panel) FMOLS estimator uses a semi-parametric correction for endogeneity and serial correlation, based on the OLS residuals and the first differences of the regressors. An alternative (asymptotically efficient) estimation method for estimating cointegrating relationships, the dynamic OLS method, employs a parametric correction for endogeneity and serial correlation, based on lead, lag, and current values of the differenced regressors. A potential problem with this method is that the use of leads and lags of the first differences of the independent variables reduces the number of observations in each cross-sectional unit, and can thus lead to a substantial loss of observations when the number of time-series observations is relatively small, as in our sample. Therefore, to minimize the problem of low statistical power associated with the use of small sample sizes, we do not use the DOLS estimator.

[8] The concept of long-run (Granger) causality is to be distinguished from the more familiar notion of "Granger causality", which refers to short-run forecastability and does not account for long-run causality through the error correction term in a cointegrated error-correction model.

[9] Eberhardt & Teal, 2013, argue that cointegration and (long-run) Granger causality techniques are a useful alternative for causal inference that can be used when it is difficult to find variables that qualify as instruments.

[10] The Human Mortality Database is available at http://www.mortality.org.

[11] Unfortunately, the ISSP data do not allow estimation of church attendance rates for different religious groups for a sufficiently large number of countries and time periods, forcing us to use aggregate church attendance rates.

[12] We exclude two countries where, according to the World Religion Dataset (available at http://www.thearda.com/), the proportion of Christians in the population is smaller than 80% on average during our sample period, Israel and Japan.

[13] If there are more than two variables (as is the case in the regression reported in Column 2), there can be more than one cointegrating relationship and thus cointegration among the regressors.

[14] We use the demeaned data to construct the error-correction term and to estimate the error-correction equations. Given the small number of time-series observations, we set $k = 1$.

[15] Three things should be noted here. First, system panel cointegration approaches require long time series to provide reliable estimates. In addition, system panel cointegration tests are seriously biased toward overestimating the cointegrating rank in small samples (e.g., Wagner & Hlouskova, 2009). Thus, even if we had historical data on public health expenditures, it would not be appropriate to use these methods in our investigation because of the relatively short time dimension of our panel. Second, as discussed in footnote 9 of the Supporting Information, the finding of cointegration between $\log(LIFE_{it})$ and $\log(CHURCH_{it})$ is consistent with the idea that religiosity affects health through its effect on public health expenditures (and through its effect on medical progress). In other words, ignoring indirect effects of $\log(CHURCH_{it})$ on $\log(LIFE_{it})$ does not affect the cointegration between $\log(CHURCH_{it})$ and $\log(LIFE_{it})$. Third, of course, a single equation model such as (1), by its nature, ignores the potential indirect effect(s) of religiosity on health through public health expenditures (and medical progress). However, and in particular because of the super-consistency of cointegration parameter estimates, there is no reason to suggest that not modeling of indirect effects induces biased estimates of total effects (or results in underestimation of the number of cointegrating vectors). Finally, we note that in the Supplementary Appendix (Section A4), we provide at least rough support for the hypothesis that societal religiosity negatively affects population health through negative effects on public health expenditures and the level of medical technology by using total government expenditures, G_{it}, as a rough proxy for public health expenditures and the patent stock, P_{it}, as a rough proxy for the level of medical technology and showing that there is (1) a positive cointegrating relationship between $\log(G_{it})$, $\log(P_{it})$, and $\log(LIFE_{it})$, (2) a negative cointegrating relationship between $\log(CHURCH_{it})$ and $\log(G_{it})$, and (3) a negative cointegrating relationship between $\log(CHURCH_{it})$ and $\log(P_{it})$. Unfortunately, we cannot include the patent stock in the analysis of our cross-sectionally dominated panel dataset in Section 3 due to the limited availability of data on patents and our measure of religiosity.

[16] Due to the lack of sufficient observations on an individual country basis, we cannot include country-specific time trends as we did in the previous analysis.

[17] As noted in Section 2.1, the standard fixed-effects OLS estimator is consistent in large T panels even when the regressors are endogenous (but its t-ratio is not asymptotically standard normal).

[18] We do not use the log of the percentage of the population aged 0–14 because it is significantly correlated with life expectancy at birth (and thus not a good instrument), whereas the percentage of the population aged 0–14 is not significantly correlated with life expectancy at birth in our dataset.

[19] We do not use the log of the Internet variable because it has some zero values.

[20] Ram, 2009, argues that if one uses the log of population density (which is equal to the log of population size minus the log of country area) along with the log of population size, the coefficient on the log of population density captures the effect of area size rather than the effect of population density on public expenditures. Therefore, following Ram, 2009, we use population density rather than the log of population density.

[21] In a previous version of this paper, we also estimated Equations (5) and (6) as a system using three-stage least squares (3SLS) and found qualitatively similar results. However, 3SLS has the drawback that if one equation in the system is misspecified, it can cause bias in estimates of other (correctly specified) equations. Therefore, it was decided not to report the results of the 3SLS analysis.

[22] WDI data are available at https://databank.worldbank.org/source/world-development-indicators.

[23] The WVS was implemented in 1981 and currently has six waves (with a seventh in the works). The WVS data are available at http://www.worldvaluessurvey.org/WVSContents.jsp?CMSID=Home.

[24] These statistics are available at https://data.imf.org/?sk=a0867067-d23c-4ebc-ad23-d3b015045405.

[25] The subsample does not include Albania, Croatia, Dominican Republic, El Salvador, Libya, New Zealand, and Puerto Rico.

[26] This subgroup does not include Azerbaijan, Bosnia and Herzegovina, Haiti, Japan, Lebanon, Montenegro, Puerto Rico, Venezuela, Vietnam, Yemen, and Zambia (due to the limited availability of data).

[27] In the Supplementary Appendix (Section A9), we also examine whether the effect of religiosity on health varies substantially across different religious traditions, but find no evidence for this.

[28] Since $INTERNET_{it}$ is the only instrument, the model is exactly identified. Thus, tests of overidentifying restrictions cannot be applied because the test statistic will have zero degrees of freedom. Therefore, we follow Evans & Schwab, 1995, who suggest that including the instrument as a regressor in the OLS model provides suggestive evidence on its validity.

[29] We also find (not reported) that the coefficient on $POP0\_14_{it}$ is statistically insignificant when $POP0\_14_{it}$ is included without $INTERNET_{it}$.

[30] To control for possible endogeneity, we also estimate specifications using future values, or leads, of $\log(LIFE_{it})$ as dependent variables (which is equivalent to using lagged values of the explanatory variables). The results are reported in the Supplementary Appendix (Section A6); they are qualitatively and quantitatively similar to those reported in Table 4.

[31] Some studies, such as Maung et al., 2020, McCleary & Barro, 2006b, and Scheve & Stasavage, 2006, use the instrument set suggested by Barro & McCleary, 2003, which consists of a religious pluralism measure and dummy variables for state religion and state regulation of

religion. In the Supplementary Appendix (Section A11), we experiment with these instruments and find that they are both weak and invalid, a finding which is also supported by Young, 2009.

[32] To strengthen our argument that societal religiosity has a negative effect on public health spending (because religious individuals are healthier and therefore demand less public health services overall than non-religious individuals)—and to ensure that the negative effect of religiosity on public health spending is not due to omitted factors that influence both religiosity and public spending, including spending on health—, we regress other (non-health) government expenditure categories on religiosity in the Supplementary Appendix (Section A8). These regressions show that there is no general negative effect of religiosity on categories of government spending.

[33] We treat all regressors (except the time dummies) as endogenous. Since GMM can exhibit the problem of too many instruments when the number of instruments is greater than the number of number of cross-sectional units, we use only the third lag of the regressors as instruments, and we collapse the instrument matrix. Since the two-step estimator is more efficient than the one-step estimator, we use the former. However, a well-known property of the two-step estimator is that the standard errors may be severely biased downwards in small samples. To address this problem, we adopt the Windmeijer, 2005, finite sample correction to the standard errors.

[34] We also estimated Equation (6) using 2SLS with country fixed-effects and found qualitatively similar results. To ensure consistency with the results in Tables 4, 6 and 7 (which are not based on fixed-effects regressions), we report in Table 8 only the OLS and 2SLS results without country fixed-effects. It is perhaps needless to say that the GMM estimator (implicitly) accounts for country fixed-effects.

[35] We also tested whether a decline in societal religiosity positively affects population health via an increase in public expenditures for social protection, but found no evidence for this. We also found no evidence that income inequality is a channel through which health is affected by religiosity.

REFERENCES

Alesina, A., & Wacziarg, R. (1998). Openness, country size and government. *Journal of Public Economics*, *69*, 305–321.

Arias-Vazquez, F. J. (2012). A note on the effect of education on religiosity. *Economics Letters*, *117*, 895–897.

Azzi, C., & Ehrenberg, R. (1975). Household allocation of time and church attendance. *Journal of Political Economy*, *83*, 27–56.

Barro, R. J., & McCleary, R. M. (2003). Religion and economic growth across countries. *American Sociological Review*, *68*, 760–781.

Becker, S. O., Nagler, M., & Woessmann, L. (2017). Education and religious participation: City-level evidence from Germany's secularization period 1890–1930. *Journal of Economic Growth*, *22*, 273–311.

Bénabou, R., Ticchi, D., & Vindigni, A. (2015). Religion and innovation. *American Economic Review*, *105*, 346–351.

Bernet, P., Gumus, G., & Vishwasrao, S. (2020). Maternal mortality and public health programs: Evidence from Florida. *Milbank Quarterly*, *98*, 150–171.

Blundell, R., & Bond, S. (1998). Initial conditions and moment restrictions in dynamic panel data models. *Journal of Econometrics*, *87*, 115–143.

Bokhari, F. A. S., Gai, Y., & Gottret, P. (2007). Government health expenditures and health outcomes. *Health Economics*, *16*, 257–273.

Brossard, D., Scheufele, D. A., Kim, E., & Lewenstein, B. V. (2009). Religiosity as a perceptual filter: Examining processes of opinion formation about nanotechnology. *Public Understanding of Science*, *8*, 546–558.

Brunello, G., Fort, M., Schneeweis, N., & Winter-Ebmer, R. (2016). The causal effect of education on health: What is the role of health behaviors? *Health Economics*, *25*, 314–336.

Deaton, A. (2011). Aging, religion, and health. In D. Wise (Ed.), *Explorations in the Economics of Aging* (pp. 237–262). University of Chicago Press.

Dills, A. K., & Hernández-Julián, R. (2014). Religiosity and state welfare. *Journal of Economic Behavior & Organization*, *104*, 37–51.

Downey, A. B. (2014). *Religious affiliation, education and Internet use*. arXiv:1403.5534. Retrieved from http://arxiv.org/abs/1403.5534

Eberhardt, M., & Teal, F. (2013). No mangoes in the tundra: Spatial heterogeneity in agricultural productivity analysis. *Oxford Bulletin of Economics & Statistics*, *75*, 914–939.

Ecklund, E. H. (2010). *Science vs. religion. What scientists really think*. Oxford University Press.

Engle, R. E., & Granger, C. W. J. (1987). Cointegration and error-correction: Representation, estimation, and testing. *Econometrica*, *55*, 251–276.

Evans, W. N., & Schwab, R. M. (1995). Finishing high school and starting college: Do Catholic schools make a difference? *Quarterly Journal of Economics*, *110*, 941–974.

Filmer, D., & Pritchett, L. (1999). The impact of public spending on health: Does money matter? *Social Science & Medicine*, *49*(3), 1309–1323.

Fonseca, R., Michaud, P.-C., Galama, T., & Kapteyn, A. (2021). Accounting for the rise of health spending and longevity. *Journal of the European Economic Association*, *19*, 536–579.

Granger, C. W. J. (1988). Some recent developments in a concept of causality. *Journal of Econometrics*, *39*, 199–211.

Gruber, J., & Hungerman, D. M. (2007). Faith-based charity and crowd-out during the great depression. *Journal of Public Economics*, *91*, 1043–1069.

Herzer, D., & Strulik, H. (2020). Religiosity and long-run productivity growth. *Journal of Economics, Management and Religion*, *1*, 1–40.

Hungerman, D. (2005). Are state and church substitutes? Evidence from the 1996 welfare reform. *Journal of Public Economics*, *89*, 2245–2267.

Iannaccone, L. (2003). *Looking backward: A cross-national study of religious trends*. George Mason University.

Kao, C. (1999). Spurious regression and residual-based tests for cointegration in panel data. *Journal of Econometrics*, *90*, 1–44.

Koenig, H. G. (2012). Religion, spirituality, and health: The Research and clinical implications. Religion, spirituality, and health: The research and clinical implications. *ISRN Psychiatry*, *2012*, 1–33.

Lewbel, A. (2012). Using heteroscedasticity to identify and estimate mismeasured and endogenous regressor models. *Journal of Business & Economic Statistics*, *30*, 67–80.
Lleras-Muney, A. (2005). The relationship between education and adult mortality in the United States. *Review of Economic Studies*, *72*, 189–221.
Maung, M., Tang, Z., & Xu, X. (2020). Religion and venture investing: A cross-country analysis. *Financial Review*, *55*, 433–460.
McCleary, R. M., & Barro, R. J. (2006a). Religion and political economy in an international panel. *Journal for the Scientific Study of Religion*, *45*, 149–175.
McCleary, R. M., & Barro, R. J. (2006b). Religion and economy. *Journal of Economic Perspectives*, *20*, 49–72.
McClure, P. K. (2020). Tinkering with technology and religion in the digital age: The effects of internet use on religious belief, behavior, and belonging. *Journal for the Scientific Study of Religion*, *56*, 481–497.
McPhetres, J., & Zuckerman, M. (2018). Religiosity predicts negative attitudes towards science and lower levels of science literacy. *PLoS One*, *13*, 1–20, e0207125. https://doi.org/10.1371/journal.pone.0207125
Nisbet, M. C., & Goidel, R. K. (2007). Understanding citizen perceptions of science controversy: Bridging the ethnographic—survey research divide. *Public Understanding of Science*, *16*, 421–440.
Okunade, A. A., & Murthy, V. N. R. (2002). Technology as a 'major driver' of health care costs: A cointegration analysis of the newhouse conjecture. *Journal of Health Economics*, *21*, 147–159.
Papyrakis, E., & Selvaretnam, G. (2011). The greying church: The impact of life expectancy on religiosity. *International Journal of Social Economics*, *38*, 438–452.
Pedroni, P. (2001). Fully modified OLS for heterogeneous cointegrated panels. *Advances in Econometrics*, *15*, 93–130.
Perman, R., & Stern, D. I. (2003). Evidence from panel unit root and cointegration tests that the environmental Kuznets curve does not exist. *Australian Journal of Agricultural and Resource Economics*, *47*, 325–347.
Pesaran, M. H. (2006). Estimation and inference in large heterogeneous panels with a multifactor error structure. *Econometrica*, *74*, 967–1012.
Ram, R. (2009). Openness, country size, and government size: Additional evidence from a large cross-country panel. *Journal of Public Economics*, *93*, 213–218.
Ricci, F., & Zachariadis, M. (2013). Education externalities on longevity. *Economica*, *80*, 404–440.
Scheitle, C. P., & Ecklund, E. H. (2017). Recommending a child enter a STEM career: The role of religion. *Journal of Career Development*, *44*, 251–265.
Scheve, K., & Stasavage, D. (2006). Religion and preferences for social insurance. *Quarterly Journal of Political Science*, *1*, 255–286.
Skinner, J., & Staiger, D. (2015). Technology diffusion and productivity growth in health care. *Review of Economics and Statistics*, *97*, 951–964.
Smith, C., Denton, M. L., Faris, R., & Regnerus, M. (2002). Mapping American adolescent religious participation. *Journal for the Scientific Study of Religion*, *41*, 597–612.
Stock, J. H., & Watson, M. W. (2011). *Introduction to econometrics*. Addison-Wesley. Chapter 15.
VanderWeele, T. J. (2017). Religion and health: A synthesis. In J. R. Peteet & M. J. Balboni (Eds.), *Spirituality and Religion Within the Culture of Medicine: From Evidence to Practice* (pp. 357–401). Oxford University Press.
Wagner, M., & Hlouskova, J. (2009). The performance of panel cointegration methods: Results from a large scale simulation study. *Econometric Reviews*, *29*, 82–223.
Westerlund, J., Petrova, Y., & Norkute, M. (2019). CCE in fixed-*T* panels. *Journal of Applied Econometrics*, *34*, 746–761.
Windmeijer, F. (2005). A finite sample correction for the variance of linear efficient two-step GMM estimators. *Journal of Econometrics*, *126*, 25–51.
Wooldridge, J. M. (1995). Score diagnostics for linear models estimated by two stage least squares. In G. S. Maddala, P. C. B. Phillips, & T. N. Srinivasan (Eds.), *Advances in Econometrics and Quantitative Economics: Essays in Honor of Professor C. R. Rao* (pp. 66–87). Blackwell.
Young, C. (2009). Model uncertainty in sociological research: An application to religion and economic growth. *American Sociological Review*, *74*, 380–397.
Zimmer, Z., Jagger, C., Chiu, C.-T., Ofstedale, M. B., Rojoa, F., & Saito, Y. (2016). Spirituality, religiosity, aging and health in global perspective: A review. *SSM - Population Health*, *2*, 373–381.

SUPPORTING INFORMATION

Additional supporting information can be found online in the Supporting Information section at the end of this article.

How to cite this article: Herzer, D. (2022). The macro-level effect of religiosity on health. *Health Economics*, *31*(6), 993–1011. https://doi.org/10.1002/hec.4494