ORIGINAL PAPER

Social isolation, health dynamics, and mortality: evidence across 21 European countries

Yarine Fawaz[1] · Pedro Mira[1]

Abstract

We provide a comprehensive picture of the health effects of social isolation using longitudinal data from 21 European countries. First, using Cox regressions, we find a significant, strong, and robust association between our social isolation index and mortality. The association is much stronger in Eastern European countries. While all of our pooled countries estimates ranged between a 20 and 30% increase in the mortality hazard for the socially isolated that number jumps to 45% for Eastern European countries. We then estimate linear regressions to study the dynamic "value-added" effects of social isolation on health and other mediator outcomes. We find that social isolation at baseline leads to worsening health in subsequent waves along all of the dimensions observed. Up to 13% of the effect of baseline social isolation on mortality can be attributed to the combined one-wave-ahead impact of social isolation on increased frailty, reduced cognitive function, and increased smoking.

1 Introduction

The current evidence regarding the interplay between social isolation and health indicates that such isolation, sometimes the subjective side of it in the form of loneliness, can be devastating to a person's health. According to recent studies, feeling lonely

✉ Yarine Fawaz
 fawaz@cemfi.es

 Pedro Mira
 mira@cemfi.es

1 CEMFI, Casado del Alisal, 5, 28014 Madrid, Spain

and being socially isolated (i.e., lacking social connections) are at least as bad for one's health as being obese or a heavy smoker (Holt-Lunstad et al. 2015). Ever since social sciences (Berkman and Syme 1979) revealed the importance of social ties as predictors of survival for persons aged under 70 at baseline, the impact of loneliness and social isolation on health and mortality has been increasingly investigated in public health, medicine, epidemiology, gerontology, and other health-related disciplines. The bulk of evidence has pointed at social isolation and loneliness being linked to a variety of physical and mental conditions such as high blood pressure, cardiovascular diseases, obesity, a weakened immune system, anxiety, depression, fatigue, pain, cognitive decline, Alzheimer's disease, and even death (Steptoe et al. 2013; Cohen et al. 1997; Shankar et al. 2013; Heffner et al. 2011; Yu et al. 2020; Teguo et al. 2016; Powell et al. 2021). In contrast, the economics literature on the topic has remained relatively scarce.

According to one meta-analysis of scientific literature on the subject (Holt-Lunstad et al. 2015), social isolation, that is, having few network ties, increases the risk of death over 7 years by approximately 30%, while the effect of loneliness is estimated at approximately 26%, and living alone seems to be the utmost risk factor with a weighted average effect of 32%.[1] That meta-analysis also reviews a number of previous studies that showed that individuals with less social connection have disrupted sleep patterns, altered immune systems, more inflammation, and higher levels of stress hormones. Another recent meta-analysis Valtorta et al. (2016) found that social isolation increases the risk of heart disease by 29% and stroke by 32%.

These meta-analyses report results from a variety of articles that do not share a common level of rigor. For example, out of the 70 studies analyzed in Holt-Lunstad et al. (2015), 31 are fully "unadjusted," meaning that they include no control of any sort, and 20% of the remaining studies fail to control for baseline health, which, according to the meta-analyses, radically changes the findings. The remaining multivariate analyses that do control for baseline health and other factors rarely have background data on individuals and are seldom based on random samples since participants are often recruited from a medical setting. Even when studies recruit participants from the general community, they rarely collect as much information as gathered in multi-disciplinary surveys such as the data used in our study and cannot claim to be fully representative.

We rely on longitudinal data on a large representative population across 21 European countries to investigate the association between social isolation and mortality and health.[2] Data from the Survey of Health, Ageing and Retirement in Europe (SHARE) enables us to follow individuals across time and mitigate part of the endogeneity concerns; it provides us with a comprehensive set of health indicators and social isolation and loneliness variables, which are observed every 2 years over 18 years, from 2004 to 2021. A few studies have exploited similar data, such as the American Health and Retirement Study (HRS) and its UK equivalent, the English Longitudinal

---

[1] Seven years is the average of the follow-up periods across the studies analyzed in the meta-analysis.

[2] Loneliness is also considered, but rather as a mediator in the association under study. Our focus is on objective social isolation, which we define according to objective criteria, such as living alone, participation in social activities, and frequency of contact with family.

Study of Ageing (ELSA), to look at correlations between social isolation, loneliness, and mortality or a specific health outcome. One noteworthy study, by Steptoe et al. (2013), uses ELSA to investigate how social isolation and loneliness at baseline are associated with mortality over a 7-year follow-up period.

Relative to those studies that use a representative longitudinal dataset, we make several contributions. First, we present a rigorous analysis of the effects of social isolation (SI) at baseline on mortality over a long follow-up period, in a harmonized multi-country framework, and produce novel findings on the heterogeneous effects of SI across country groups.

Second, in addition to the baseline health controls included in the main mortality analysis, we explore health behaviors, healthcare utilization, loneliness, and a poor social network as additional potential mediators of the relationship between SI and mortality.

Third, we run linear regressions to study the dynamic "value-added" effects of SI on health and other mediator outcomes, that is, the effect of baseline SI on all observable dimensions of health, health behavior, and healthcare utilization in each future wave when controlling for their baseline levels. The Cox analysis is then combined with the dynamic regressions to compute a simple metric of how much of the SI effect on mortality can be attributed to the one-wave-ahead detrimental effect of baseline SI on each dimension of health, health behavior, and healthcare utilization. This metric can serve as a guide toward understanding where it is most necessary to intervene in order to curb the detrimental effects of social isolation.

We find a significant and robust association between our social isolation index and mortality. Heterogeneous effects of social isolation on mortality across countries are revealed as a major finding. The impact of social isolation on the elderly may possess a cultural and/or institutional dimension, which should be examined in a cross-national framework. A much stronger association is found between social isolation and mortality in Eastern countries. While all of our pooled countries estimates range from a 20 to 30% increase in the mortality hazard for the socially isolated, this number jumps to 45% for Eastern countries. This objective measure of social isolation does not produce the same health consequences across countries, albeit using harmonized data, which suggests that public health policies have a role to play in moderating the health risks posed by social isolation.

Remarkably, controlling for loneliness barely weakens the relationship between our social isolation index and mortality, which is also true in Steptoe et al. (2013). This suggests that loneliness cannot be the only mechanism through which social isolation affects health. While we find that socially isolated individuals are more likely to adopt a worse lifestyle (particularly in terms of physical inactivity), the inclusion of unhealthy behavior measures at baseline in our regressions does not affect the coefficient on social isolation. Likewise, healthcare utilization does not appear as a major channel for the effect of social isolation on future health. On the one hand, we find that socially isolated individuals do not use more healthcare services than their non-socially-isolated counterparts, with the exception of prescription drug consumption. This holds in spite of the fact that the health of the socially isolated continues to worsen, which suggests that social isolation might inhibit the use of some healthcare

services. However, on the other hand, the inclusion of measures of current healthcare use in our regressions incurs no effect on the coefficient on social isolation.

We also find that social isolation at baseline correlates with worsening health in the subsequent waves along all the dimensions under observation, and these effects are persistent. Up to 13% of the effect of baseline social isolation on mortality can be attributed to the combined one-wave-ahead impact of social isolation on increased frailty, reduced cognitive function, and increased smoking.

Given our careful empirical strategy and multiple robustness checks, we believe that a causal interpretation of our findings is plausible. However, we acknowledge the need to be cautious regarding causality because our study is observational and does not make use of any so-called natural experiment.[3] As an additional check to provide support for a causal relationship, we use Oster's test for selection on unobservables (Oster 2019). We also use education as a benchmark for the health effects of social isolation. We find the education gradient in mortality is smaller than the social isolation gradient, but the association of education with future health is stronger than the association we find for social isolation in dynamic value-added regressions.

The rest of the paper is structured as follows: Section 2 establishes a link between social isolation at baseline and future mortality while controlling for a rich set of socioeconomic and health indicators that cover physical, functional, mental, and cognitive health at baseline. In Section 3, additional potential mediators of the social isolation-mortality association are explored by adding them as baseline controls in the Cox regressions. Section 4 presents the dynamic regressions of health and other mediating variables on the SI index, with the same baseline controls as in the Cox regression and the baseline value of the mediating variable. We present a metric to measure how much of the SI effects on mortality can be attributed to the effect of SI at baseline on each dimension of health, health behaviors, and healthcare utilization, one survey wave ahead. Section 5 discusses the causality challenge, Oster's test of selection on unobservables, and education as a benchmark of the SI effects on mortality and health outcomes. Lastly, Section 6 provides the conclusions.

2 Social isolation and mortality

2.1 Data

We use longitudinal survey data from the Survey of Health, Ageing and Retirement in Europe (SHARE) that covers 8 waves from 2004 to 2019, plus the two "Corona" waves of Spring 2020 and Summer 2021, across 20 European countries plus Israel. This multidisciplinary and cross-national panel database holds micro data on health, socio-economic status, and social and family networks on approximately 140,000 individuals aged 50 and over (around 530,000 interviews). Our sample is made of 67,676 non-institutionalized individuals (i.e., those not living in a nursing home, at least at baseline), corresponding to 243,515 observations, whom we observe at least

---

[3] Henceforth, we indistinctively report an "effect" of social isolation on health or an "association" between them.

twice over the 10 waves (the second observation might be an "exit" interview, i.e., a post-mortem interview), and whose information on the set of variables used herein is available (see Table 2 for more details).[4] Individuals enter the study in any wave between wave 1 (in 2004) and wave 6 (in 2015) and are followed for a minimum period of 24 months, due to the minimum follow-up restriction imposed.[5] The maximum follow-up time is 207 months (i.e., 17 years and 3 months). The median follow-up time is 79 months (i.e., 6 years and 7 months). Over the course of our study, 9802 deaths, are observed, which corresponds to 14% of our sample.

In order to carry out our empirical strategy, we create a set of health indicators that covers physical, functional, mental, and cognitive health. Physical health is investigated along several lines: objective (i.e., number of diagnosed chronic diseases) and subjective (self-assessed health status); focusing on functional health (Activities of Daily Living (ADLs) and Instrumental Activities of Daily Living (IADLs)); and constructing an index of frailty (Fried et al. 2001), which aggregates unintentional weight loss, self-reported exhaustion, weakness (grip strength), difficulties in walking, and low physical activity. Mental health is summarized by the EURO-D score, which is the sum of 12 items that can be relied on to diagnose depression in older adults, such as suicidal thoughts, sadness, no hopes for the future, excessive guilt, sleep issues, fatigue, irritability, loss of appetite, tearfulness, concentration issues, lack of enjoyment, and difficulties in maintaining interest in things. Cognitive functioning is an average of immediate and delayed word recall (i.e., the number of words an individual is able to remember out of a list of ten words). We also make use of the so-called "exit interviews," which allow us to keep track of respondents' death.

Regarding the key regressor, following Steptoe et al. (2013), we create a social isolation index including information on whether the individual lives alone, has infrequent/no social contact with their children (less than weekly contact, or does not have children), and does not participate in any social activities (including political, sports, educational, and voluntary work activities).[6] The resulting index lies between 0 and 3, with a 1.03 mean and 0.79 standard deviation for our sample at baseline. Higher values indicate increasing social isolation. Table 1 displays the distribution of our SI index. In our sample, 20% of individuals live alone, 61% participate in no social activities, and 20% have infrequent/no contact with their children (13% due to not having children). More than half our sample is massed at SI = 1, and very few individuals have

---

[4] Although SHARE now encompasses 29 countries and all waves are considered herein, only data on the 21 countries: that entered SHARE before the last wave can be included (since more than one observation per individual is required), and that appeared at least twice across the 8 first waves, excluding wave 3, which was dedicated to constructing life histories of SHARE respondents.

[5] Otherwise, the minimum follow-up time until death would be 1 month, which is too short for obvious reverse causality concerns. We argue in favor of a 24-month period when presenting the Cox model in Section 2.2.

[6] Note that this index is close to the original Berkman-Syme Social Network Index developed in (Berkman and Syme 1979) for a population aged under 70, which included (1) marital status; (2) contacts with close friends and relatives; (3) membership of a church group; and (4) memberships in other types of groups). We do not include contact with other family or friends because these items were absent from SHARE until wave 4, when a social networks module was introduced for the first time (it appears again in waves 6 and 8). We will use that module when creating an index of connectedness, but no item from that module is included in our main SI index so that we can follow respondents for a much longer time span.

Table 1 Distribution of the social isolation (SI) index and its components

	SI=0	SI=1	SI=2	SI=3	Total
Live Alone	0.00	0.09	0.58	1.00	0.20
No participation in any social activities	0.00	0.82	0.87	1.00	0.61
Less than weekly contact with children	0.00	0.09	0.56	1.00	0.19
Observations	18256	33992	12581	2847	67676
%	26.98	50.23	18.59	4.21	100

the maximum value of 3 (4%). Interestingly, the jump between a 0 and a 1 value of the social isolation index is largely driven by the non-participation in social activities, while the jump from 1 to 2 is due to a shift in both the other two components of living alone and having infrequent/no contact with children (or not having any children).

The short form of the R-UCLA loneliness scale is also utilized. This was created by aggregating 3 items (how much of the time they felt a sense of being left out, the lack of companionship, and isolation) into one single measure of loneliness.[7] The effects on mortality, both of objective social isolation and of the perception of social isolation, can therefore be studied.

Table 2 provides descriptive statistics at baseline (when individuals enter the data) on the SI-related variables, the aforementioned health outcomes, and the socio-demographic controls that are to be used in our analyses. It also puts forward the major differences between two populations, those who are not socially isolated at all according to our index, and those with at least 1 social isolation point. The more socially isolated exhibit worse health measurements along all dimensions (note that cognitive functioning is the only health measure where a higher value means better health), and are more likely to be female, less educated, and childless than the non-socially-isolated. Regarding gender and social isolation, the reality is more complex, since women become socially isolated at a higher rate than men as they grow older (past 60), which is likely to reflect gender imbalances in the probability of being widowed after the age of 60. We also add an employment indicator since whether one is working or retired may explain a major part of social isolation for the over-50s population.

Figure 1 and A2 (the latter in the Appendix) show that there is a great deal of variation in social isolation across countries, whereby Eastern and Southern Europe countries have the highest average levels of social isolation, and Western and Northern Europe the lowest. Since the correlation between loneliness and social isolation at the individual level is a mere 24%, it seems that objective and subjective measures of social isolation capture different aspects of social experience, as suggested in Hughes et al. (2004). As a consequence, several countries, such as Italy, Greece, and Israel, have very high levels of loneliness compared to other countries with a similar level of social isolation, while others, such as Switzerland, Austria, and Denmark, have very low rates of loneliness in comparison with other countries with a similar level of social isolation.

---

[7] See Hughes et al. (2004) for a validation of the short version of the RUCLA scale of loneliness.

Table 2 Descriptive statistics on observables, by social isolation status

	SI index=0		SI index>0		All	
	mean	sd	mean	sd	mean	sd
Social isolation						
SI> 0	0.00	(0.00)	1.00	(0.00)	0.73	(0.44)
Live Alone	0.00	(0.00)	0.27	(0.44)	0.20	(0.40)
No participation in any social activities	0.00	(0.00)	0.84	(0.37)	0.61	(0.49)
Less than weekly contact with children	0.00	(0.00)	0.26	(0.44)	0.19	(0.39)
Loneliness (RUCLA scale)	3.49	(0.96)	3.89	(1.36)	3.76	(1.26)
SN disconnectedness	1.88	(0.91)	2.17	(0.86)	2.09	(0.88)
Health behavior						
Currently Smokes	0.15	(0.36)	0.21	(0.40)	0.19	(0.39)
Sedentary	0.03	(0.16)	0.10	(0.30)	0.08	(0.28)
Drank alcohol past 3 months/7 days	0.82	(0.38)	0.64	(0.48)	0.69	(0.46)
Healthcare utilization						
Number of medicines	1.17	(1.28)	1.58	(1.54)	1.47	(1.49)
Number of doctor visits	5.41	(7.91)	6.65	(9.68)	6.32	(9.25)
Number of overnight stays in 12 mo	0.11	(0.32)	0.14	(0.35)	0.13	(0.34)
Health						
Self-perceived health	2.71	(1.03)	3.20	(1.07)	3.06	(1.08)
Frailty index	0.56	(0.81)	1.02	(1.16)	0.90	(1.09)
Number of chronic diseases	1.47	(1.38)	1.80	(1.58)	1.71	(1.54)
Number of limitations	0.17	(0.71)	0.47	(1.35)	0.39	(1.22)
Depression score (EuroD)	1.91	(1.86)	2.56	(2.30)	2.39	(2.21)
Cognitive recall test (the higher, the better)	5.09	(1.62)	4.24	(1.76)	4.47	(1.76)

Table 2 continued

	SI index=0		SI index>0		All	
	mean	sd	mean	sd	mean	sd
Socio-demographic controls						
Age	62.14	(8.33)	65.45	(9.73)	64.56	(9.49)
Female	0.49	(0.50)	0.57	(0.50)	0.55	(0.50)
Education: None	0.01	(0.12)	0.05	(0.22)	0.04	(0.20)
Education: Primary	0.09	(0.29)	0.23	(0.42)	0.20	(0.40)
Education: Secondary(i)	0.14	(0.35)	0.19	(0.39)	0.18	(0.38)
Education: Secondary(ii)	0.35	(0.48)	0.31	(0.46)	0.32	(0.47)
Education: Tertiary	0.39	(0.49)	0.20	(0.40)	0.25	(0.44)
Education: Other	0.01	(0.07)	0.01	(0.08)	0.01	(0.08)
Employed or Self-employed	0.42	(0.49)	0.26	(0.44)	0.30	(0.46)
Has at least one child	1.00	(0.00)	0.82	(0.39)	0.87	(0.34)
Lives in rural area	0.34	(0.47)	0.30	(0.46)	0.31	(0.46)
Lives in a house (not a building)	0.77	(0.42)	0.63	(0.48)	0.67	(0.47)
Total household income	65688	(2144930)	28498	(41142)	38530	(1114689)
Household net worth	376684	(664093)	618897	(64990243)	553558	(55537994)
Observations	18,256		49,420		67,676	

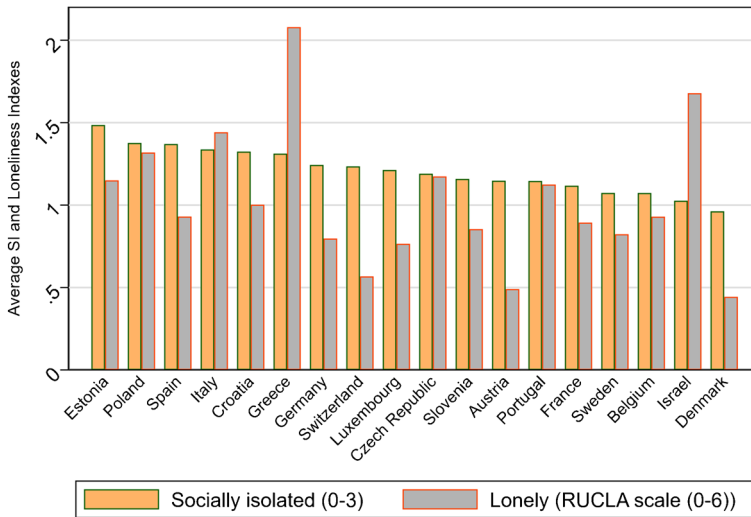Note: sample with a follow-up of 24-month minimum

Fig. 1 Social isolation and loneliness across Europe (18 countries)

2.2 Main results from Cox models

We first look at how social isolation at baseline is associated with future mortality, by estimating Cox proportional hazards regression models, from the date an individual enters the data (February 2004 at the earliest) until that individual dies or is followed up in subsequent surveys up to July 2021. Out of the 67,676 followed individuals, 9802 died over the period. Although our longitudinal data would allow us to let our explanatory variables vary across time, we keep them fixed at baseline in order to introduce some distance between the measurement of covariates, more particularly social isolation, and the outcome of mortality. In our preferred specification, a minimum of 24 months of follow-up is imposed between the moment social isolation is measured, and mortality, following the robustness checks performed in Steptoe et al. (2013): "we repeated the analysis excluding deaths within 24 months of baseline, and the results were very similar results to those for the full cohort, suggesting that existing terminal illness is not the primary explanation." This specification helps to alleviate concerns about reverse causality.

A potential concern may arise as to how stable our measure of social isolation is across time: if SI were to vary widely from one wave to another, then picking its first observation *ad hoc* might lead us to overestimate (or underestimate) the health effects of SI, if that observation was particularly low (or high) that precise year.

Evidence in favor of the stability over time of the SI index is presented in Fig. A1, where each line represents the average SI over time for individuals who were followed for two waves, three waves, and so on, up to seven waves. Out of the 67,676 individuals who enter our survival analysis, 60,454 are represented in this graph: those who are not represented have an SI value at baseline and (an)other value(s) at some future waves. Table A1 provides more detail on how many individuals have only a baseline

SI index, how many have an SI index at t_0 and $t_0 + 1$, how many at t_0, $t_0 + 1$, and $t_0 + 2$, and so on until t_0 to $t_0 + 6$. We can observe at most 7 values of the SI index, that is, Waves 1 to 8 inclusive, except for wave 3. The items that are part of the SI index were not present in the questionnaires of the two "Corona" waves.

This graph informs us about two things: (i) Those who "disappear" earlier have a (slightly) higher social isolation index than those who are followed over 4, 5, and 6 consecutive waves. Since higher social isolation leads to higher mortality, it is reasonable that the "survivors" exhibit lower social isolation; (ii) nevertheless, for both categories, the SI index seems quite stable over time.

Another way to ascertain the stability of the (binary) SI index over time is by looking at transitions into and out of being socially isolated between t and $t + 1$, as shown in Table A2. Those who are socially isolated at t remain socially isolated at $t + 1$ with an almost 90% probability, while those who are not socially isolated at t remain in that state with an almost 70% probability. The transition from not being socially isolated to being socially isolated stands at approximately 30%, while the transition from being socially isolated to not being socially isolated is lower, at around 11%.

We estimate the following Cox model:

$$h(i, t|SI_{i,t_0}, X_{i,t_0}) = h_0(t)exp(SI_{i,t_0}\beta + X_{i,t_0}\gamma) \tag{1}$$

The hazard of dying, h, at time t is a function of a fully flexible baseline hazard h_0 common to all individuals in our sample, which is shifted proportionally upwards or downwards by social isolation SI at baseline and by the individual characteristics X_i introduced in the model at baseline t_0. A total of 5 models are fitted, each adding further constraints to the relationship between social isolation and mortality.

Our results in Table 3 fall closely in line with those of Steptoe et al. (2013), who find a hazard ratio for a comparable discrete social isolation index that varies between 1.50 and 1.26. Our first two models suffer from obvious omitted-variable bias since no background information on the individual is included in Model (1), and no health information is included at baseline in Model (2), while gender, age, and health at baseline are potentially both correlated with social isolation at baseline (it is hard to be socially connected to people when in bad health, for instance) and future mortality. The results are displayed in columns (1) and (2). The point estimates of the hazard ratio for the social isolation index, in its continuous version from 0 to 3, decrease from 1.21 to 1.13 when controlling for all health indicators at baseline, since part of the association between social isolation and future health passes through initial health conditions, although the coefficient remains both large and strongly significant. Models (4) and (5) restrict follow-up to individuals who are still alive after the first 24 months, as a way to mitigate reverse causality, thereby ensuring that no life-threatening conditions of our respondents had been overlooked by our observable covariates at baseline. This restriction poses no challenge to our estimate. The comparison of individuals with at least 1 point of social isolation to those with zero point (column (5): the social isolation index is a binary variable) leads to a higher hazard ratio than looking at the effect of one extra point of social isolation (column (4)), meaning that a change from 0 to 1, 0 to 2, or 0 to 3, has more impact than the average increase of one point of SI over all the

Table 3 Cox models: effect of social isolation at baseline on mortality up to Summer 2021 (SHARE second Covid-19 survey)

	(1)	(2)	(3)	(4)	(5)
Social isolation index	1.188***	1.208***	1.127***	1.118***	1.254***
	(0.013)	(0.016)	(0.017)	(0.018)	(0.038)
Basic	Yes	Yes	Yes	Yes	Yes
Socio-Demo	No	Yes	Yes	Yes	Yes
Health	No	No	Yes	Yes	Yes
Follow-up	No	No	No	>24 mo	>24 mo
Binary SI index	No	No	No	No	Yes
Observations	341806	325406	273201	243515	243515
Individuals	97751	92558	72659	67676	67676

Exponentiated coefficients; standard errors in parentheses

Note: Basic controls: age, age squared, gender. Socio-demographic controls: education categories, whether employed, whether has at least one child, income and wealth quartiles, living in a house (vs. building), rural (vs. urban). Health controls: frailty, self-assessed health, number of chronic diseases, number of limitations, depression score, cognitive recall test. All regressions include wave and country-specific fixed effects

$^{*}(p < 0.10)$, $^{**}(p < 0.05)$, $^{***}(p < 0.01)$

possibilities of SI increasing by 1. In all the following analyses, we adopt the latter, with a binary SI index and the exclusion of follow-up times shorter than 24 months.

2.3 Sensitivity checks

A number of sensitivity checks are performed: (1) to open the black box of the SI index and check whether a specific component is driving most of the SI effects on mortality; (2) to make sure that our results are not driven by a specific subsample; (3) and to verify whether marital status is a confounder in our analysis. (This variable has remained excluded due to its high correlation with living alone).

Since the three components that make the SI index may reflect different dimensions of social interaction and support, evidence regarding the relative importance of each component is provided. In Table A3 (in the Appendix), the results of Cox regressions are displayed that include each of the eight possible combinations of the three items. The only item that does not have any significant impact on mortality by itself is the factor "living alone." Even considered jointly with one or two other SI items, living alone does not seem to add (at least significantly) to the effect of the other items. For instance, living alone and not participating in any associations jointly lead to a 24% increase in the mortality hazard compared with not being socially isolated at all (SI=0), although the single factor of not participating in any associations leads to a 28% greater hazard. When associated with infrequent contact with children, although living alone seems to have a greater impact than when considering "few contacts" alone, there is no statistical difference between the two estimated hazard ratios. Both non-participation in associational activities and infrequent contact with children (or no children), considered separately, exert a major impact on mortality, with non-participation in associational activities possibly carrying even greater weight and the

most widespread presence across the sample, with 41% of our sample possessing only that item). When considered jointly, whether this means a value of 2 or of 3 for the SI, the impact of these factors on mortality is even stronger (with the hazard being shifted upwards by 42%). This exercise is informative as to the contribution to the index of each component and hence is considered useful in terms of policy implications since it enables a better understanding of which variables should be targeted by policymakers: a policy aiming at incentivizing older individuals to participate in associational activities (on which local authorities, for instance, have agency, in contrast to the frequency of contact with one's children), might lead to curbing at least one of the two most harmful dimensions of SI for individuals' health.

Second, in order to ensure our results are not driven by a specific subsample, we estimate our specification (column (5) of Table 3) on several subgroups. The main concern is related to the construction of the social isolation index: if a particular population is more likely to have less contact with their children (say, males), to live alone (unmarried individuals), or to participate in social activities (working versus retired individuals), then the results found across the whole representative sample could be misleading. Since these characteristics are already controlled for in our regressions, this is less of a concern, but the results of this sensitivity analysis are still displayed in Table A4 Panel A (in the Appendix). Apart from a few exceptions, such as the employed who have a higher mortality risk associated with social isolation than the non-employed (hazard ratios of 1.49 against 1.22), the hazard ratio remains remarkably stable at approximately $1.22-1.29$ across almost all subsamples. Individuals with no children are mechanically assigned to the "socially isolated" group when the SI index is binary and defined as "$SI > 0$," since the item "infrequent contact with children" is set to 1 for individuals with no children. Hence, the coefficient of the (binary) SI index cannot be identified for the childless.

One solution towards ascertaining whether SI affects mortality differentially for the childless and those with children is to re-run these regressions using the continuous index of social isolation: we learn how a one-unit increase of the SI index affects the two groups, albeit that the estimate of the impact of the SI index cannot incorporate the effect of going from 0 to 1 for the childless group. Table A4 Panel B shows the impact of the (continuous) SI index is remarkably stable across the childless and "with children" subsamples. One difference with the binary case is the difference between the married and non-married individuals; when the continuous measure is used, those married seem to be more at risk of mortality when more socially isolated than those who are not married. Again, this could be due to the married being much less socially isolated than the unmarried (an SI average of 0.79 vs. 1.71), partly due to the very high correlation between one of the SI components, that of living alone, and the SI index.

Hitherto we have not controlled for marital status due to its very high correlation with "living alone" (between 75 and 83% depending on the definition of marital status). This might still constitute a major confounder, and hence, it is checked by re-running the main Cox estimation: (1) using marital status instead of "living alone" in the definition of the SI index (columns (2) and (3) of Table A5) 2) controlling for marital status in the main regression (columns (4) and (5)); and while 3) adding marital status together with the other two items in a regression (columns (6) and (7)). We define being married as "married and living with spouse" or "in a registered partnership," versus

"married, not living with spouse," "never married," "divorced," and "widowed"; we define "in a couple" as being in a couple and living with one's partner regardless of the official marital status.

As shown in Table A5 in the Appendix, our SI index (column (1)) seems to have a greater impact on mortality than when replacing "living alone" with marital status (columns (2) and (3)), which was not as straightforward as it seemed since the "living alone" item had no particular relevance in the SI index. Second, controlling for either version of marital status does not change anything regarding the estimated impact of our SI index. Third, hardly any evidence can be found of the well-known protective effect of being married on individuals' health, which we interpret as evidence that our baseline controls do a good job at capturing health at baseline. When dropping these (columns (8) and (9)), married individuals face a 7% lower hazard of death over the follow-up period, which is close to the lower bound of the effects of marital status found in the literature (12% according to a meta-analysis consisting of more than 250,000 elderly subjects (Manzoli et al. 2007)), as one could expect given the richness of the set of baseline controls used herein. All in all, marital status does not seem to drive the effect of social isolation on mortality.

2.4 Country heterogeneity

One of the most unique features of the SHARE datasets is that it is harmonized across all of Europe, and therefore, we can study how social isolation affects health and mortality differentially across countries. We group countries into four culturally and geographically consistent subgroups: Western (Austria, Germany, the Netherlands, France, Switzerland, Belgium, and Luxembourg); Northern (Sweden and Denmark); Southern (Spain, Italy, Greece, and Portugal); and Eastern countries (Czech Republic, Poland, Hungary, Slovenia, Estonia, and Croatia). Ireland and Israel are excluded from this part of the analysis.

A first look at heterogeneity across these four groups of countries (see Table 4, columns (1) and (2)) suggests that the hazard ratio found for the socially isolated against the non-socially-isolated conceals major differences across countries. While in Western, Northern, and Southern countries, social isolation (defined as having at least one social isolation point) is associated with a 1.19 hazard ratio (whereby Northern countries exhibit a higher HR but do not differ statistically regarding traditional thresholds), the social isolation HR is much higher in Eastern countries (1.45=1.19x1.22). In other words, social isolation has a similar impact in Western, Southern, and, to a certain extent, in Northern countries, but there is a very strong and significant difference between these countries and Eastern countries. Columns (3) to (5) introduce each country group against the other three in order to ascertain whether any other pattern would appear when pooling more countries together in the reference group. Again, only in Eastern countries are the socially isolated at a greater risk of mortality compared with other countries, and strikingly so: in all non-Eastern countries, being socially isolated is associated with a 23% extra likelihood of dying over the follow-up period; in Eastern countries that likelihood increases by a further 18%.

Table 4 Cox models: Country heterogeneity of the impact of social isolation at baseline on mortality up to Summer 2021 (SHARE second Covid-19 survey)

	(1)	(2)	(3)	(4)	(5)
SI: SI index>0	1.27***	1.19***	1.29***	1.26***	1.23***
	(0.04)	(0.06)	(0.04)	(0.04)	(0.04)
Western countries	ref	ref	ref	ref	ref
Southern countries	1.07**	1.06	1.13	1.07**	1.08**
	(0.03)	(0.09)	(0.09)	(0.03)	(0.03)
Northern countries	1.58***	1.46***	1.58***	1.54***	1.58***
	(0.06)	(0.10)	(0.06)	(0.10)	(0.06)
Eastern countries	1.49***	1.25***	1.48***	1.49***	1.29***
	(0.05)	(0.09)	(0.05)	(0.05)	(0.09)
SI × Western		ref	ref	ref	ref
SI × Southern		1.02	0.94		
		(0.09)	(0.08)		
SI × Northern		1.10		1.04	
		(0.09)		(0.08)	
SI × Eastern		1.22**			1.18**
		(0.10)			(0.09)
Observations			235154		
Individuals			65210		

Exponentiated coefficients; Standard errors in parentheses
Controls included: all health and socio-demographic variables displayed in Table 2, and wave FE
Follow-up of 24-month minimum
*($p < 0.10$), **($p < 0.05$), ***($p < 0.01$)

This finding is important to the extent that if the same level of social isolation is associated with different mortality hazards across countries, then there may be room for public policies to weaken that association. Our results provide no indication of what makes people in Eastern countries more vulnerable when they are socially isolated, but our model enables us to rule out several hypotheses: at the individual level, we control for income and wealth quartiles within country and also for educational attainment. The focus is therefore on the effect of social isolation for individuals with a similar socio-economic status. Furthermore, indicator variables are included for country groups (or even for each country when not looking specifically at the effect of a certain group of countries), and hence, whatever may make individuals more or less healthy in a country (such as aggregate economic conditions, and generosity of the healthcare system) is already captured by these indicators. If these country-specific characteristics are still reflected in the interaction term, then it would mean that socially isolated individuals are more vulnerable to aggregate economic or healthcare conditions than non-socially-isolated individuals, even when controlling for their income and health.

We attempt to shed light on the specific aspects of Eastern countries that could explain this result by first re-estimating the same model separately for each country:

despite the lack of power to find significant results, this re-estimation enables us to ascertain whether a particular country might be driving the results of its group. As shown in Table A6 (in the Appendix), Poland, for instance, where the socially isolated are 83% more likely to die over the period than the non-socially-isolated, could simply be an outlier. With a significant hazard ratio of 0.59, Portugal could also be the reason why the socially isolated in Eastern countries suffer a higher mortality hazard over the period than do the socially isolated in Southern countries. The Cox regressions are then re-estimated with interactions presented in Table 4 (more specifically Column (2)) by excluding one country at a time, in order to check whether the hazard ratio for the interaction is stable or whether an outlier country might be driving our result for Eastern countries. Table A7 (in the Appendix) confirms that regardless of which country is excluded from the analysis, the mortality pattern of the socially isolated in Western, Southern, and Northern countries does not differ, while Eastern countries face an additional mortality hazard of between 17 and 27% for the socially isolated.

In addition to studying differences in the observables between the four groups of countries in Table A8, we also look into potential cultural and policy factors, using data from the European Social Survey (ESS). Regarding social isolation, Eastern and Southern countries are much more socially isolated than Northern and Western countries, mostly due to their high rate of non-participation in associational activities; they are also lonelier and less connected in terms of social networks (using the scale defined in Section 3).

Both Eastern and Southern countries also perform worse than Western and Northern countries across all health dimensions, with Eastern countries in a worse situation, especially in terms of self-assessed health, number of chronic diseases, and number of limitations, but they do much better than Southern countries in terms of cognitive functioning. Notably, our samples of Eastern and Southern countries differ greatly along the education dimension, with older Eastern Europeans being much more educated.

Eastern countries seem to differ from the rest of the countries due to a combination of high social isolation and bad health, although they remain similar to Southern countries in many features. One potential explanation for these heterogeneous mortality effects of SI is that, conditional on older individuals suffering from poorer health in both Southern and Eastern countries, Eastern healthcare systems might be worse than those of Southern countries.

Table A9 shows suggestive evidence that this could be the case: while the proportion of people who declare they have suffered from symptoms of depression that lasted at least 2 weeks remains remarkably stable across the four groups of countries (approximately 26%), the share of those who have ever been treated for depression by a doctor or psychiatrist (amongst those who have ever been depressed) is much lower in Eastern countries (40% against 54% in Southern countries). The European Social Survey (ESS) also points in the same direction, with individuals from Eastern countries rating the "state of health services in [their] countries nowadays" as worse than in the rest of the countries.

Another possibility is that our social isolation index lacks the friendship dimension that is present in Steptoe et al. (2013) and that the frequency of meeting friends is positively correlated with other items of our index. The proportion of individuals who never socialize with their friends, relatives, or colleagues is much higher in Eastern

Europe (24%) than in Southern Europe (19%), Western (10%), and Northern countries (5%). If this constitutes an important dimension of social isolation and if it correlates with, for instance, participation in associational activities or frequency of contact with children, then what we capture in the interaction with Eastern Europe could actually be due to this precise dimension.

3 Pathways from social isolation to death: loneliness, social connectedness, healthcare utilization, and health behavior

The association we uncovered between social isolation and mortality was found to be robust not only to several definitions of the social isolation index, but also to the inclusion of all the confounders suspected of being correlated both with social isolation and mortality (such as health at baseline, education, income and wealth, and country of residence), to the exclusion of the first 24 months after baseline and to restriction to several subgroups.

Once this relationship is established, the main question becomes that of the underlying mechanisms. Loneliness, social connectedness, health behavior, and healthcare utilization are all studied in turn as potential mediators in the relationship between social isolation and health.

How might social isolation lead to adverse health outcomes? Apart from the biological pathways, whereby the inflammatory and antiviral processes are suspected to be at the very core of this association (see Leschak and Eisenberger (2019)), health behavior (such as smoking, drinking, and a sedentary lifestyle), and healthcare utilization (in the form of doctor visits, preventive screenings, etc.) might explain why social isolation is so monotonically associated with worse health. These two mechanisms are in turn very much linked with one's social network, as "individuals who are socially engaged and connected are exposed to stronger normative pressures from and control by friends and loved ones to perform healthy behaviors and to access healthcare when needed" (Cacioppo and Hawkley 2003). We therefore create an index of "connectedness," which takes into account both the quantity and quality of one's social network, and check whether either of these three pathways mediates the association between social isolation and mortality.[8] It is also often argued that perceived social isolation, which is also referred to as the feeling of loneliness, could be the channel through which objective social isolation influences mortality. Another question is: what contributes the most to an individual's (bad) health: the objective or the subjective dimension of social isolation? We hence test whether loneliness mediates or mitigates this relationship.

Our preferred specification is first estimated in which the RUCLA scale of loneliness is added as a control. Regardless of whether we include it as a binary or continuous variable (columns (1) to (3) of Table 5), the hazard ratio of the social isolation index remains unchanged, even though loneliness by itself is positively and significantly

---

[8] This index of connectedness strives to summarize the richness of the social networks modules of SHARE waves 4, 6, and 8, which use name generators to construct respondents' networks of confidants, into one measure.

Table 5 Cox models: does loneliness mediate the association between social isolation and mortality?

	(1)	(2)	(3)	(4)
Social isolation index	1.27***	1.27***	1.27***	1.20***
	(0.07)	(0.07)	(0.07)	(0.08)
RUCLA loneliness at baseline		1.05***		
		(0.02)		
RUCLA loneliness at baseline (d)			1.26***	
			(0.09)	
RUCLA loneliness:time-varying				1.07***
				(0.02)
Observations	101894	101894	101894	61677
Individuals	34544	34544	34544	33347

Exponentiated coefficients; Standard errors in parentheses

Controls included: all health and socio-demographic variables displayed in Table 2, and wave and country FE

Waves 1 to 3 are excluded from the analysis (no RUCLA information)

Follow-up of 24-month minimum

$^{*}(p < 0.10)$, $^{**}(p < 0.05)$, $^{***}(p < 0.01)$

associated with mortality. If we allow the loneliness measure to vary over time (column (4)), instead of being fixed at baseline, then the impact of the SI index drops from 27 to 20%, but reverse causality is also more of a concern as there is less time between loneliness measured and death observed. It therefore seems that loneliness assumes little of the explanatory power of social isolation.

A social connectedness scale is subsequently constructed to test whether a poor social network could mediate the relationship between social isolation and mortality. Rich information from the social network modules introduced in waves 4 and 6 is utilized, in which respondents are asked to name up to 7 confidants, or people with whom [they] most often discussed important things, and to provide information about their relations to each one thereof. Following Malter and Börsch-Supan (2017), our measure of social connectedness uses (1) the number of people cited (network size); (2) the number of cited social network members living within a proximity of 25 km; (3) the number of cited persons with weekly or more contact (contact frequency); (4) the number of cited persons with very or extremely close emotional ties (support); and (5) the number of different types of relationships present within the network (diversity).[9] The connectedness scale is then reversed into a "disconnectedness" scale. The resulting scale lies between 0 and 4 (with a mean of 2), with higher values indicating a poorer social network. As shown in Table 6, similarly to the loneliness scale, social network disconnectedness is associated with higher mortality, but does not seem to mediate the relationship between social isolation and mortality, since it barely takes away anything from the impact of SI on mortality.

Health behavior is then explored as a potential pathway between social isolation and higher mortality: socially isolated individuals may have worse lifestyles in terms

---

[9] See Malter and Börsch-Supan (2017) for details on the construction of the connectedness scale.

Table 6 Cox models: does social disconnectedness mediate the association between social isolation and mortality?

	(1)	(2)	(3)	(4)
Social isolation index	1.30***	1.28***	1.29***	1.26**
	(0.07)	(0.07)	(0.07)	(0.11)
SN disconnectedness at baseline		1.09***		
		(0.02)		
SN disconnectedness at baseline (d)			1.16***	
			(0.04)	
SN disconnectedness: time-varying				1.16***
				(0.04)
Observations	84818	84818	84818	38707
Individuals	25234	25234	25234	22949

Exponentiated coefficients; Standard errors in parentheses

Controls included: all health and socio-demographic variables displayed in Table 2, and wave and country FE

Follow-up of 24-month minimum

*($p < 0.10$), **($p < 0.05$), ***($p < 0.01$)

of smoking and drinking more and moving less, which could put them at a higher risk of mortality. When adding health behavior information into the model (column (2) of Table 7), the hazard ratio of the social isolation index drops slightly from 1.23 to 1.19. All three variables come up as highly significant: smoking at baseline increases the mortality hazard by 83%, while being sedentary at baseline (i.e., never engaging in vigorous or moderate activity) raises this rate by 14%. Our measure of alcohol consumption, on the other hand, fails to capture the harmful effects of alcohol, since it is associated with lower mortality. This has to be either due to the definition used, that is, having drunk any alcohol over the last 3 months or the last 7 days depending on the wave, or due to the fact that by controlling for health and socio-economic status, alcohol can be associated with positive outcomes. Had we been able to observe more extreme forms of alcohol consumption, such as binge drinking and alcoholism (which is included from Wave 4), then the result might have been different.

Regarding healthcare utilization as another potential pathway, the addition of healthcare utilization information (column (3)) appears not to alter the SI coefficient whatsoever. Neither the number of medicines an individual takes at baseline nor the number of doctor visits one has had over the last 12 months, seem to matter once health is taken into account (when health controls are not included in the equation, then the number of medicines does). In contrast, having stayed overnight at a hospital over the last 12 months is associated with a 13% higher hazard, even though all observable dimensions of health are controlled for at baseline. This latter finding highlights one dimension that is not well captured by our health controls, that is, the severity of one's condition: although self-assessed health and the number of chronic diseases are controlled for, the severity of the illness of an individual might be better proxied by adding the number of overnight stays in hospital in the past 12 months.

Table 7 Cox models: do healthcare and health behavior mediate the association between social isolation and mortality?

	(1)	(2)	(3)	(4)	(5)
Social isolation index	1.23***	1.19***	1.23***	1.19***	1.17***
	(0.04)	(0.04)	(0.04)	(0.04)	(0.05)
Currently smokes		1.83***		1.85***	1.71***
		(0.05)		(0.05)	(0.07)
Sedentary		1.14***		1.14***	1.92***
		(0.04)		(0.04)	(0.06)
Had alcohol in last 3mo/last week		0.92***		0.93***	0.79***
		(0.02)		(0.02)	(0.02)
Number of medicines			1.00	1.00	1.00
			(0.01)	(0.01)	(0.01)
Number of doctor visits			1.00***	1.00***	1.01***
			(0.00)	(0.00)	(0.00)
Stay overnight hospital in 12 mo			1.13***	1.13***	1.62***
			(0.03)	(0.03)	(0.05)
Observations	224691	224691	224691	224691	148512
Individuals	62553	62553	62553	62553	60458
Controls	at baseline	at baseline	at baseline	at baseline	time-varying

Exponentiated coefficients; standard errors in parentheses

Controls included: all health and socio-demographic variables displayed in Table 2, and wave and country FE

Follow-up of 24-month minimum

* ($p < 0.10$), ** ($p < 0.05$), *** ($p < 0.01$)

Lastly, when these potential "mediators" are allowed to vary wave by wave, then they tend to take away more of the SI impact on mortality, since sedentarism and hospital stays gain greater importance, which is consistent with them capturing some unobserved part of the health deterioration process. Although loneliness, social disconnectedness and health behavior show some correlation with both social isolation and mortality, none of these appear as major channels of the association between social isolation and mortality. The following section studies how social isolation affects the dynamics of health (in its observable dimensions), health behavior, and healthcare utilization, and assesses how much of its impact on mortality is directed through each of these dimensions.

4 The dynamic impact of social isolation on health, health behavior, and healthcare utilization outcomes

4.1 Health outcomes

As a second step in digging into the potential pathways from social isolation to health, the dynamics of the association between social isolation and all relevant dimensions of health are explored, some of which should show a significant decline (since social isolation leads to heightened mortality). As before, the focus is on major health indicators belonging to both the objective and subjective health spectrum and summarizing all relevant dimensions of health: physical (including frailty and functional health/limitations), mental, and cognitive health. Our sample is exactly the same as before but it should be borne in mind that the number of observations decreases over time as 41,821 individuals form part of our sample, observed through 2 consecutive waves, while only 7506 are observed 6 waves after their entry (which does not imply participating in all waves in between).

We estimate the following equations and plot the relevant coefficient, $alpha_1$, in the dynamic graphs displayed in Fig. 2:

$$Health_{i,t+j} = \alpha_0 + \alpha_1 SI_{i,t} + \alpha_2 Health_{i,t} + \alpha_3 X_{i,t} + \epsilon_{i,t} \tag{2}$$

where $j = 1, 2, ...6$. At $j = 0$ the coefficient of the SI index is mechanically 0, which is why it is not shown and the coefficients are plotted from $j = 1$. We regress each future health outcome $Health$ for individual i at time $t + j$ on social isolation SI at baseline t (the binary indicator that is equal to 1 whenever the index is non null) and for the exact same set of baseline characteristics used in the Cox model. These are summarized in Table 2, including the complete vector of health characteristics (amongst which is the outcome at baseline). Hence, the α_1 coefficient measures the correlation between SI at baseline and the deterioration (or the change), rather than the level, of health. Again, covariates are fixed at baseline in order to introduce some distance between the covariates and the outcomes.

For each health outcome, six regressions are run. There are more than 6 waves, but, in contrast with mortality, which is known at each wave and for which the date of death is known even if it occurs between two waves, here, several outcomes are not provided at all waves, such as at wave 3 for depression and frailty (SHARELIFE), and during the two "Corona" waves. "Time" indexes future waves, 1 for wave t + 1, and so on, up to 6, for wave t + 6 (individuals are observed at most from wave 1 to wave 7). In all our regressions, the outcome is measured at one of these future waves, while the rest of the variables are fixed at baseline. All health outcomes are standardized so that their mean is 0 and standard deviation is 1, which facilitates the comparison of the graphical representations of our regressions. They are coded so that higher values indicate worse health and are represented using the same scale on all graphs. As mentioned before, sample size shrinks over time, and hence confidence intervals become larger over time. A clear picture is still attained of how social isolation correlates with health over time.

Social isolation is undeniably associated with worsening health. Nevertheless, there is heterogeneity across outcomes: cognitive health (recall test) starts worsening in association with social isolation after one wave and the effect of social isolation remains at that same level after two waves, in line with Shankar et al. (2013), who find poorer cognitive functioning amongst the socially isolated four years after baseline using the ELSA data. Since we are controlling for baseline cognitive functioning, our results indicate a higher rate of decline for the socially isolated, in accordance with Ertel et al. (2008), who find a higher rate of memory loss using word recall for individuals with lower social integration (which is very close to our measure of social isolation) using HRS data. Other outcomes, such as frailty and self-assessed health, follow a similar trend, while certain outcomes return to their initial level, such as depression after four waves. It therefore appears that social isolation worsens both physical and cognitive health in the short and long run, but its association with mental health is only transitory. Functional health, when measured as "suffering at least one limitation," instead of the number of limitations as we had done so far, becomes increasingly correlated with social isolation over time, before possibly going back to the baseline level (the precision of the estimates does not allow us to derive any conclusion after 6 waves of follow-up). Nonetheless, the relationship between functional health and social isolation is sensitive to the definition of functional health: when it is defined as the sum of limitations with ADLs, our estimates are much closer to being non-significantly different from 0, in line with Shankar et al. (2017), who do not find a significant association between the number of ADLs and social isolation using two waves of the ELSA data.

This empirical exercise enables a relationship between social isolation and the deterioration of health to be revealed in almost all its facets. Even when considering mental health, for which the association does not persist in the long run, there is still a deterioration occurring two waves after baseline. Moreover, a high correlation at baseline between social isolation and poor mental health could be associated with higher mortality but would not be revealed in our estimates. In other words, if the socially isolated at baseline suffer from more symptoms of depression, even in the absence of further deterioration, poor mental health that remains poor over the follow-up period could also be a channel leading to higher mortality, which should be borne in mind when interpreting the results.
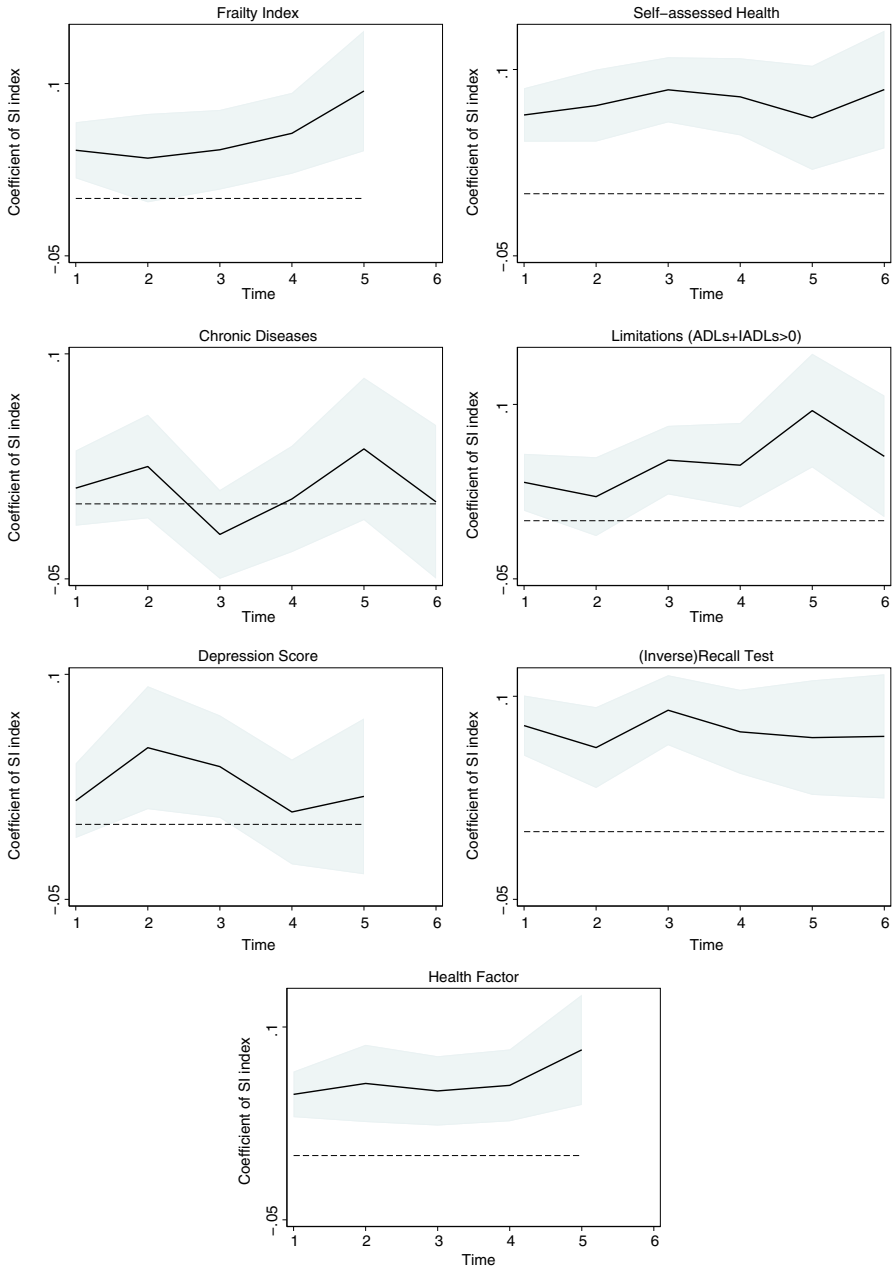
Fig. 2 Social isolation at baseline and health dynamics. Note: the figures display the coefficients and 95% confidence intervals for the effect of social isolation on all (standardized) health outcomes. "Time" indexes future waves: 1 is wave $t + 1$, 2 is wave $t + 2$, etc. Regressions are performed separately for each outcome and lag. All regressions include the health and socio-demographic variables displayed in Table 2 and wave (a dummy for which wave is baseline) and country FE, and control for the outcome at baseline

4.2 Health behavior and healthcare utilization outcomes

Apart from the "biological" channel, the literature puts forward health behavior and healthcare utilization as potential pathways from social isolation to worsened health and mortality. In the mortality section, health behavior has already been implied as playing a role in the SI-mortality relationship, whereas healthcare utilization has not. In the present section, we investigate whether there is a specific pattern of the socially isolated in terms of health behavior and healthcare. The same dynamic analysis is applied to the set of health behavior and healthcare utilization variables employed in Section 3. Again, in each regression, the outcome is controlled for at baseline, and hence looking at how social isolation is associated with smoking in later waves is equivalent to studying how it correlates with changes in smoking.

As shown in Fig. 3, there is no significant relationship between social isolation and smoking (when controlling for smoking at baseline), except after one and three waves, but the dynamic pattern remains unclear. If anything, social isolation seems to be associated with less drinking. The one important behavior that is increasingly and significantly associated with social isolation over time is that of sedentarism, defined as "hardly ever or never" engaging in vigorous (e.g., sports) or moderate (e.g., gardening, going for a walk) physical activity. Sedentarism may then play a role in how socially isolated individuals become sicker, but it is also reasonable that as individuals get sicker they would engage less in physical activity. Shankar et al. (2011) and Kobayashi and Steptoe (2018) find similar results on inactivity, and a more clear-cut association with smoking and drinking, without controlling for these variables at baseline, and they conclude that loneliness and social isolation may affect health independently through their effects on health behavior. One way to ascertain whether sedentarism is a mechanism per se is to control for the health factor at future waves as well, in addition to at baseline. When doing so, the trend looks the same, but the coefficient is no longer significant, and hence the association between social isolation and sedentarism could also be spurious due to their common correlation with a worsened health status.

Regarding the relationship between social isolation and healthcare utilization, there are two (or more) possible directions in which social isolation might affect healthcare utilization: (1) since social isolation is associated with worse health (see previous subsection) and higher mortality (see Section 2.2), the socially isolated might use more healthcare; (2) social isolation might make individuals less inclined to use healthcare, as their loved ones "pressure" them less to undergo medical checkups, prompt them less often to seek medical help when needed, and even fail to accompany them to a visit to the doctor.[10] Our results help to shed light on this discussion: Fig. 4 indicates socially isolated people not using any more or less healthcare than non-socially-isolated individuals. Regardless of whether future health is included as a control or

---

[10] Socially isolated individuals might also resort less to healthcare due to lack of information, as put forward in Devillanova (2008), who documents a lower time to visit for immigrants with a strong social tie who know about healthcare opportunities.
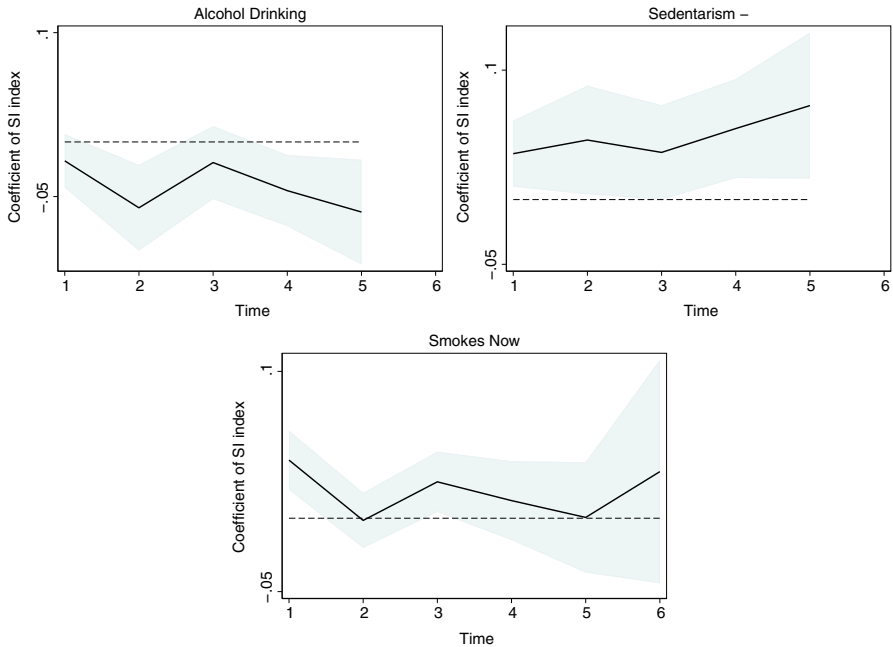
Fig. 3 Social isolation at baseline and health behavior dynamics. Note: the figures display the coefficients and 95% confidence intervals for the effect of social isolation on different health behavior outcomes. "Time" indexes future waves: 1 is wave $t + 1$, 2 is wave $t + 2$, etc. Regressions are performed separately for each outcome and lag. All regressions include the health and socio-demographic variables displayed in Table 2 and wave (a dummy for which wave is baseline) and country FE, and control for the outcome at baseline

not in the regressions, socially isolated individuals do not use more medicine than the non-socially-isolated, and neither do they visit their physician any more or less. If anything, socially isolated individuals spend more nights at the hospital, after a few waves, but this relationship is weak, both in terms of significance and magnitude.[11]

How is this "null" result compatible with the two aforementioned directions? First, as mentioned for the case of health behavior, any baseline correlation between social isolation and healthcare utilization levels is already factored in the baseline controls. Second, mechanisms (1) and (2) could theoretically cancel each other out. The fact that no positive effect of SI on healthcare utilization can be found due to health deterioration (i.e., direction (1)) is striking and consistent with mechanism (2) being a potential channel.

---

[11] This finding goes against some of the literature that points at lonely or socially isolated individuals using more healthcare than individuals who do not suffer from loneliness or social isolation. One example is Gerst-Emerso and Jayawardhana (2015), who find that the lonely are more likely to visit their doctor (but not to be hospitalized), even when controlling for their health, suggesting that individuals who suffer from chronic loneliness look for social support in their physician, but that the lack of healthcare use and barriers to healthcare access do not seem to drive the social isolation-health relationship.
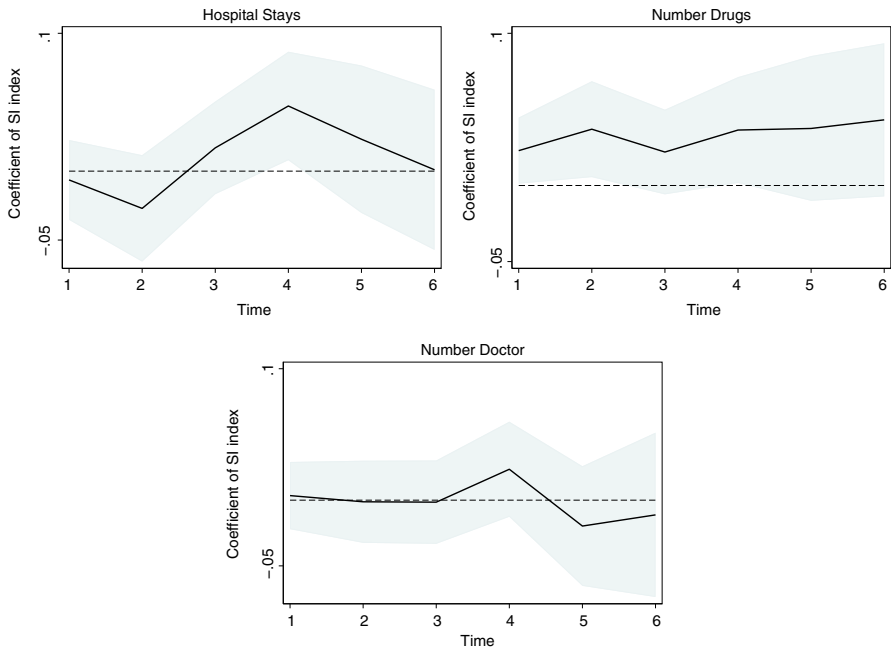
Fig. 4 Social isolation at baseline and healthcare utilization dynamics. Note: the figures display the coefficients and 95% confidence intervals for the effect of social isolation on different healthcare utilization outcomes. "Time" indexes future waves: 1 is wave $t + 1$, 2 is wave $t + 2$, etc. Regressions are performed separately for each outcome and lag. All regressions include the health and socio-demographic variables displayed in Table 2 and wave (a dummy for which wave is baseline) and country FE, and control for the outcome at baseline

4.3 Quantifying the contribution of each health dimension to the social isolation-mortality relationship

Having estimated the mortality effects of SI, we now dive into the dynamics of the social isolation effects in order to answer how social isolation might lead to higher mortality. In this section, two questions are asked: (1) Which mediator has the biggest impact on mortality hazards? (2) Which mediator is the most affected by social isolation? The answers to the two questions are then combined to compute how much of the SI effects on mortality can be attributed to the effect of SI at baseline on each dimension of health, in the subsequent period. In Table 8 we present the results of the same Cox regression as in the main table (Table 3 column(5)) but displaying the hazard ratios that correspond to the health controls. In column (1), the variables are introduced as in the descriptive statistics shown in Table 2, while in column (2) all the health variables are standardized to facilitate the comparison of their effects. In answer to the first question, we find that self-assessed health, frailty, and cognitive functioning exert a major impact on mortality and are of a similar magnitude to that of social isolation: 28, 22, and 16%, respectively. When controlling for all other health dimensions, depression is not associated with higher mortality. This is to be expected since it shares 57% correlation with the frailty dimension and 39% with self-assessed

Table 8 Cox models: How does health impact mortality?

	(1)	(2)
Social isolation index	1.254***	1.246***
	(0.038)	(0.040)
Self-perceived health	1.239***	1.278***
	(0.016)	(0.020)
Frailty index	1.203***	1.216***
	(0.013)	(0.015)
Number of chronic diseases	1.027***	1.041***
	(0.007)	(0.012)
Number of limitations	1.035***	1.051***
	(0.006)	(0.012)
Depression score (EuroD)	0.965***	0.923***
	(0.005)	(0.012)
Cognitive recall test (the higher, the better)	0.906***	0.839***
	(0.007)	(0.012)
Health variables	Non-standardized	Standardized
Observations	243515	225779
Individuals	67676	67127

Exponentiated coefficients; standard errors in parentheses

Controls included: all health and socio-demographic variables displayed in Table 2, and wave and country FE

Follow-up of 24-month minimum

*$(p < 0.10)$, **$(p < 0.05)$, ***$(p < 0.01)$

health. On removing these two dimensions from the equation, depression is significantly associated with higher mortality. Neither the number of chronic diseases nor the number of limitations in ADLs and IADLs constitute a sizeable effect. These coefficients are reported (as coefficients instead of hazard ratios this time), as $\beta(H)$ in Table 9 (first row), along with $\alpha(S)$, in the second row: the coefficient of social isolation in the health outcome regression (at $t + 1$, i.e. the first point plotted on each graph of Fig. 2). The second row provides an answer to the second question: self-assessed health, frailty, and cognitive health, which are the same as in question 1, are the health dimensions most affected by social isolation, at least at wave $t + 1$.

These coefficients $\beta(H)$ and $\alpha(S)$ are then multiplied by each other in order to obtain a coefficient that is generated by SI at $t + 1$, through the health channels between t and $t + 1$ (see third row). Last, we compare this coefficient with $\beta(S)$, the coefficient of SI at baseline in the Cox model, which is equal to 0.227. In the last row, we show how much of the SI effect on mortality can be attributed to the dynamic effect of social isolation on health one wave ahead: while chronic diseases, functional health (limitations), and depression account for less than 1% of the SI effect on mortality, the impact of SI on mortality through self-assessed and cognitive health at the next wave, reported on the effect of these health variables on mortality, stands at approximately

Table 9 Chain-rule effect: How much of the SI impact on mortality hazards can be explained through dynamic health effects?

	SA Health	Frailty	Chronic diseases	Limitations	Depression	Cognitive (bad) health
Coeff of health variable in Cox model						
$\beta(H)$	0.245	0.195	0.040	0.050	−0.080	0.176
Coeff of SI in health outcome (at $t+1$) regression						
$\alpha(S)$	0.063	0.043	0.019	0.020	0.017	0.078
$(\beta(H)*\alpha(S))$	0.016	0.008	0.001	0.001	−0.001	0.014
Coeff of SI in Cox model						
$\beta(S)$				0.227		
Chain rule effect						
$(\beta(H)*\alpha(S))/\beta(S)*100$	6.866	3.710	0.331	0.446	−0.596	6.039

Note: 0.227 is the coefficient of SI in a Cox model with all the health and socio-demographic variables displayed in Table 2, and wave and country FE

6–7% (4% in the case of frailty). Although this does not seem like a very high figure, it is informative with respect to policy, that is, where and how to intervene in order to curb the SI effects on mortality. A policy designed to target social isolation directly, for instance promoting associational activities for older people, could therefore be viewed not only as a way to allow those who would otherwise be socially isolated to live longer, but also as a way to slow their cognitive decline. Reduction of social isolation could therefore become part of the recommendations, such as those emitted by the WHO in order to reduce the risk of dementia and cognitive decline (Chowdhary et al. 2021), and help " Understand the influence and interactions of non-modifiable (e.g., gender, genetics, age) and modifiable (e.g., physical activity, diet, and cognitive stimulation) risk and protective factors for dementia in population-based samples."

On reproducing the same exercise for health behaviors and healthcare utilization (see Table 10), we find that a small part of the impact of SI on mortality passes through the impact of SI on increased smoking (3%). On the other hand, the coefficient of smoking in the Cox regression is as big as that of self-assessed health, and hence, our "chain rule" accounting exercise still yields a non-null percentage for smoking. By comparison, none of the impact of SI on mortality passes through its impact on sedentarism at $t + 1$.

5 Discussion: the causality challenge and the education benchmark

5.1 Causality and other caveats

We uncovered a strong association between social isolation and mortality, in the form of a 25% increase in the mortality hazard rate for individuals who are socially isolated at baseline. The magnitude of this association is close to the estimates found in the literature for the causal impact of education on mortality. How causal can this association be proven to be? In order to make a stronger case for a causal association, all the people who die in the 24 months following the baseline (when social isolation is observed) are discarded from the sample. In this way, our sample is guaranteed not to suffer any life-threatening health condition that would fail to be captured in our health controls and that would still provide a reason for one to be socially isolated. This restriction, coupled with a very long follow-up period (up to 17.25 years, with a median follow-up of 79 months), makes reverse causality highly unlikely.

The main concern involves the potential existence of omitted variables that would affect both social isolation and mortality (or health outcomes). It is not straightforward to come up with potential confounders that are not controlled for in our regressions and would be correlated with both baseline social isolation and future health: not only is health already being controlled for extensively at baseline, but also socio-economic status, through income, wealth, and education, and other observable characteristics that could be related to both the main explanatory variable and the outcome, such as gender, whether individuals are working, and whether they have children. We also control for housing variables, as there could be a link between living in a rural area (vs. urban), and in a house (vs. in a building), and future health, whereby it may be harder for older individuals living in a rural area to seek medical attention when needed. At

Table 10 Chain-rule effect: how much of the SI impact on mortality hazards can be explained through health behavior and healthcare utilization?

	Smokes	Sedentarism	Drinking	Number of meds	Doctor visits	Hospital stays
Coeff of behavior/HCU variable in Cox model						
$\beta(H)$	0.208	0.005	-0.067	-0.087	0.069	0.111
Coeff of SI in behavior/HCU outcome (at $t+1$) regression						
$\alpha(S)$	0.040	0.035	-0.017	0.023	0.003	-0.006
$(\beta(H)*\alpha(S))$	0.008	0.000	0.001	-0.002	0.000	-0.001
Coeff of SI in Cox model						
$\beta(S)$	0.257	0.289	0.285	0.257	0.287	0.286
Chain-rule effect						
$(\beta(H)*\alpha(S))/\beta(S)*100$	3.195	0.065	0.406	-0.778	0.081	-0.252

Note: $\beta(S)$ is the coefficient of SI (binary) in a Cox model with all the health and socio-demographic variables, wave and country FE, plus the health behavior or healthcare utilization (HCU) variable that corresponds to each column

the same time, everything else being equal, living in a house with no neighbors, or in an isolated area, could also be correlated with social isolation.

An example of an unobservable factor that could determine social isolation is that of personality: Cacioppo et al. (2000) show that individuals from the lowest quintile on the UCLA loneliness scale were found to score lower on neuroticism and higher on surgency (extraversion), conscientiousness, and social agreeableness than individuals from the highest quintile, who in turn did not differ on any of these dimensions from individuals from the middle quintile. Introverted or neurotic individuals are probably more at risk of social isolation, since, for instance, they would be less prone to participate in social activities. Regarding personality and health, the latest evidence using SHARE's "Big 5" data that was collected for the first time at wave 7 indicates personality as being associated with an array of health indicators in older European adults. In particular, the strongest and most consistent personality-level correlates of good health are high conscientiousness and low neuroticism (Shemesh et al. 2019).[12] Nevertheless, these would be factored into our health indicators at baseline in our main Cox model. Since neuroticism, for instance, could also be correlated with a more acute deterioration of health, personality traits should become part of our controls. Their incorporation would not be without problems since these traits are measured at wave 7, while our baseline measures of health and social isolation might come from previous waves. Arguing for the stability of (at least a fraction of) the Big 5 over time in later life, we include these traits in our Cox regressions.[13] Table A10 shows that the same traits that correlated highly with several health indicators, that is, high conscientiousness (both as a continuous and binary variable) and low neuroticism (only as a binary variable) are associated with lower mortality, even when controlling for health at baseline. Individuals who have low conscientiousness (i.e., in the first quintile of that measure), which supposedly captures having a high propension to be self-controlled and to delay gratification, to be task- and goal-oriented, organized, efficient, precise, and deliberate (John et al. 1999), are exposed to a 25% higher mortality.[14] Even such a high correlation fails to remove any effect from our social isolation index.

One potentially omitted confounder which unfortunately cannot be included in our analysis is that of genetics. What if the same genes that are overexpressed in socially isolated individuals are also responsible for activating the immune system and the inflammation mechanism in the body? This has been found by Cole et al. (2007) amongst a sample of 230 Americans aged 50–67, which explains why lonely people suffer from chronic inflammation in spite of their high levels of cortisol and are

---

[12] Often referred to by the acronym OCEAN, these are openness to experience (vs. closedness), conscientiousness (vs. lack of direction), extraversion (vs. introversion), agreeableness (vs. antagonism), and neuroticism (vs. emotional stability).

[13] Cobb-Clark and Schurer (2012) shows that these traits are stable for at least a 4-year period, and Erlich and Litwin (2019) using the SHARE Big 5 data, establish that two personality attributes, conscientiousness and neuroticism, hardly vary across age.

[14] Conscientiousness is positively derived in SHARE from answers to the statement "I see myself as someone who does a thorough job" and negatively derived from answers to "I see myself as someone who tends to be lazy".

vulnerable to microbes, viruses, and other sources of tissue damage.[15] Genetics could therefore be a relevant source of omitted-variable bias in our study, if linked both to social isolation and mortality. Having established a list of potential confounders, what more could be done to assess how causal a relationship can be? In practice, several approaches can help mitigate omitted-variable concerns. The most straightforward way consists of including an appropriate set of observable controls (Angrist and Pischke 2010), such as when frailty and chronic diseases are included along with functional, self-assessed, mental, and cognitive health in order to capture the true health status of the individual. An additional approach that has been widely used in the empirical literature relies on demonstrating the stability of the key coefficient when faced with the inclusion of additional controls. Table A11 shows that the coefficient of the social isolation index remains stable over various specifications when adding a different subset of controls at each specification. For this table, we chose to regress self-assessed health six waves ahead on social isolation at baseline (at entry into the study) and other controls, such that the coefficient in column (4) corresponds to the point at time = 6 in the "self-assessed health" graph in Fig. 2, although the same stability could be shown for the other outcomes at other times.

As proposed in Oster (2019), despite being very intuitive, this idea relies on the selection on observables being informative regarding the selection on unobservables, which does not necessarily follow from the assumptions of the linear model. We therefore appeal to Oster's use of coefficient stability as a test for selection on unobservables. The test considers both the stability of the estimated social isolation treatment effect when adding key observables and the importance of these factors in explaining health outcomes. The estimate of the coefficient of proportionality proposed by Oster as a summary of the robustness of results is 1.37, which is higher than the lower bound of 1 which Oster proposes as a standard for robustness. This value implies that unobservables would need to be 37% more important than the observables in explaining the social isolation "treatment," in order for the treatment effect of social isolation to be zero, which seems unlikely given the richness of our data.[16] We are therefore more confident that our finding of a significant social isolation effect is not an artifact of omitted variable bias.

5.2 Education as a benchmark of social isolation effects on mortality

Education and mortality The association we find between social isolation and mortality in Table 3 appears both significant and robust to many checks. In order to gauge the magnitude of this association, the effects of social isolation on mortality are compared to those of education, which seems an ideal benchmark candidate since a compelling literature establishes significant positive associations between education and several dimensions of adult health and negative associations between education and mortal-

---

[15] This study looks at chronically lonely individuals, according to the RUCLA scale of loneliness, rather than at socially isolated individuals.

[16] In order to obtain this value of 1.37, we need to assume a multiplicative factor for the R-squared. This factor bounds the maximum R-squared that would be achieved if it were possible to include unobservable controls in the regression, relative to the R-squared we obtain with our observable controls. We choose a factor of 25%, which seems plausible based on Oster (2019).

ity. Whether these associations can be qualified as "causal" remains subject to debate, and diverging conclusions have been reached by several studies. Using changes in education legislation as a source of exogenous variation in educational attainment, Lleras-Muney (2005) in the US and Crespo et al. (2014) in Europe found support for a causal link. No such causal effect is found in Behrman et al. (2011), who uses comparisons of twins in Denmark. In contrast, Halpern-Manners et al. (2020), with a more recent assessment of this causal relationship based on representative US "twins data," support a causal interpretation of the education-mortality gradient.

We reproduce the same specification as before (see Eq. (1)) and display both the hazard ratio associated with our social isolation index and those associated with our education controls (see Panel A of Table A12 in the Appendix). Education is introduced as 4 categories, the higher educational attainment group being omitted as the reference group (the "other" category was dropped here, to make the interpretation more straightforward, hence the slight discrepancies in sample sizes with respect to Table 3). When controlling for socio-demographic information at baseline, lower education levels are associated with higher hazard ratios (ranging from 1.35 for the lowest education level to 1.22 for the upper secondary group), with greater magnitude than the hazard ratio corresponding to social isolation (1.21).

Note that when adding baseline health as a control, most of the effect of lower education disappears, and the magnitude of the effect becomes similar to that of social isolation. This effect is less for the upper secondary education group. This suggests that most of the effect of education on mortality has already been channeled into individuals' health status when they enter the study, except for those who are the closest to the higher education group. Therefore, the more controls and horizon constraints are added to the model, the less significant the association becomes between education and mortality. The only remaining significant difference lies between the upper secondary and the tertiary education group. In short, the education-mortality gradient has a similar magnitude as that between social isolation and mortality, but is less robust to the inclusion of the same additional controls and constraints. This result does not question the causal relationship between education and health obtained from the quasi-experimental studies cited above, which do not control for health when looking at education and mortality. When social isolation is removed from the equation and if years of education are considered instead of education categories (see Panel B of Table A5) in order to more closely resemble the study by Lleras-Muney (2005), we find that 1 year of education is associated with a 2.4% drop in the hazard of dying (over the period), which lies in the $1.3-3.6$ interval found in the aforementioned study, over a 10-year period. On adding baseline health into the equation, this result no longer holds.

Dynamic effects of education on health outcomes Again, in order to gauge the magnitude of the effects of social isolation on future health, we construct the same graphs as before, but display the coefficient of the lowest education category (with respect to the highest category) instead of that of the social isolation index (see Fig. A3). As before, controlling for baseline health implies we are looking at the effect of educa-

tion on changes in health (or on future health given health at entry), so one possibility would be that all of the impact of education is already factored in health at entry, and is not reflected in deviations of health from that point on. We find a very similar pattern as in Fig. 2, except that the magnitude of the coefficients is much bigger, e.g. coefficients are twice those of social isolation for self-assessed health, and 6 times those of social isolation for cognitive health (word recall). Looking at the health factor that summarizes all the observed dimensions of health, being socially isolated takes a toll on individuals' health, approximately half the toll of being in the lowest education group.

6 Conclusions

In this paper, a strong association between social isolation and future mortality is found, which is not solely mediated by concurrent loneliness, health behavior, and healthcare utilization, nor solely by social connectedness. We also explore the dynamics of the health impact of social isolation and find that social isolation correlates with a persistent worsening of all the facets of health considered herein (self-assessed, frailty, cognitive, mental, functional).

Although previous studies have identified similar associations, we contribute to the existing literature by looking at health across many dimensions instead of focusing on a single health outcome. Our study is carried out in a harmonized multi-country longitudinal framework that enables us to employ a long follow-up period and alleviates endogeneity concerns. We investigate heterogeneity in the social isolation-health relationship across numerous countries and find a much stronger association between social isolation and mortality in Eastern countries. The fact that this very same objective measure of social isolation does not produce the same health consequences across countries, albeit using harmonized data, suggests that public health policies might have a role to play in moderating the health risks posed by social isolation.

We test several potential underlying mechanisms and find that socially isolated individuals do not resort to more healthcare use in subsequent waves than non-socially-isolated individuals, although their health does worsen across all dimensions. This suggests that healthcare utilization might be a channel underlying the relationship between social isolation and health. When our mortality, health, and health behavior models are combined in an accounting exercise, we find that up to 13% of the effect of baseline social isolation on mortality can be attributed to the combined one-wave-ahead impact of social isolation on increased frailty, reduced cognitive functioning, and increased smoking.

Last but not least, we provide evidence in favor of a causal interpretation of our estimates using Oster's test for selection on unobservables. We also compare the social isolation-health gradient to the much-studied education-health gradient. Using the same models, the education gradient in mortality is less steep than the social isolation gradient, but the association of education with future health is stronger than that found for social isolation in dynamic value-added regressions.

In terms of public policy, our research shows that older individuals could benefit from public health initiatives that would focus on the strengthening of social connections through participating in clubs and associations. Furthermore, having established the existence of a pathway between social isolation and the deterioration of health for older individuals, it becomes relevant to investigate, using post-pandemic data, how the COVID-19 pandemic has impacted social isolation in the medium and longer run.

Declarations

Conflict of interest The authors declare no competing interests.

References

Angrist JD, Pischke J-S (2010) The credibility revolution in empirical economics: how better research design is taking the con out of econometrics. J Econ Perspect 24(2):3–30

Behrman JR, Kohler H-P, Jensen VM, Pedersen D, Petersen I, Bingley P, Christensen K (2011) Does more schooling reduce hospitalization and delay mortality? New evidence based on Danish twins. Demography 48(4):1347–1375

Berkman LF, Syme SL (1979) Social networks, host resistance, and mortality: a nine-year follow-up study of Alameda County residents. Am J Epidemiol 109(2):186–204

Cacioppo JT, Hawkley LC (2003) Social isolation and health, with an emphasis on underlying mechanisms. Perspect Biol Med 46(3):S39–S52

Cacioppo JT, Hawkley LC, Ernst JM, Burleson MH, McClintock MK, Malarkey WB, Hawkley LC, Kowalewski RB, Paulsen A, Hobson JA, Hugdahl K et al (2000) Lonely traits and concomitant physiological processes: the MacArthur social neuroscience studies. Int J Psychophysiol 35(2–3):143–154

Chowdhary N, Barbui C, Anstey KJ, Kivipelto M, Barbera M, Peters R, Zheng L, Kulmala J, Stephen R, Ferri CP et al (2021) Reducing the risk of cognitive decline and dementia: WHO recommendations. Front Neurol 12

Cobb-Clark DA, Schurer S (2012) The stability of big-five personality traits. Econ Lett 115(1):11–15

Cohen S, Doyle WJ, Skoner DP, Rabin BS, Gwaltney JM (1997) Social ties and susceptibility to the common cold. JAMA 277(24):1940–1944

Cole SW, Hawkley LC, Arevalo JM, Sung CY, Rose RM, Cacioppo JT (2007) Social regulation of gene expression in human leukocytes. Genome Biol 8(9):R189

Crespo L, López-Noval B, Mira P (2014) Compulsory schooling, education, depression and memory: new evidence from SHARELIFE. Econ Educ Rev 43:36–46

Devillanova C (2008) Social networks, information and health care utilization: evidence from undocumented immigrants in Milan. J Health Econ 27(2):265–286

Erlich B, Litwin H (2019) Personality, age and the well-being of older Europeans. Health and socioeconomic status over the life course: first results from SHARE waves 6:35–41

Ertel KA, Glymour MM, Berkman LF (2008) Effects of social integration on preserving memory function in a nationally representative US elderly population. Am J Public Health 98(7):1215–1220

Fried LP, Tangen CM, Walston J, Newman AB, Hirsch C, Gottdiener J, Seeman T, Tracy R, Kop WJ, Burke G, McBurnie MA (2001) Frailty in older adults: evidence for a phenotype. J Gerontol A 56(3): M146–M157

Gerst-Emerso K, Jayawardhana J (2015) Loneliness as a public health issue: the impact of loneliness on health care utilization among older adults. Am J Public Health 105(5):1013–1019

Halpern-Manners A, Helgertz J, Warren JR, Roberts E (2020) The effects of education on mortality: evidence from linked US Census and administrative mortality data. Demography 57(4):1513–1541

Heffner KL, Waring ME, Roberts MB, Eaton CB, Gramling R (2011) Social isolation, C-reactive protein, and coronary heart disease mortality among community-dwelling adults. Soc Sci Med 72(9):1482–1488

Holt-Lunstad J, Smith TB, Baker M, Harris T, Stephenson D (2015) Loneliness and social isolation as risk factors for mortality a meta-analytic review. Perspect Psychol Sci 10(2):227–237

Hughes ME, Waite LJ, Hawkley LC, Cacioppo JT (2004) A short scale for measuring loneliness in large surveys: results from two population-based studies. Res Aging 26(6):655–672

John OP, Srivastava S et al (1999) The big-five trait taxonomy: history, measurement, and theoretical perspectives

Kobayashi LC, Steptoe A (2018) Social isolation, loneliness, and health behaviors at older ages: longitudinal cohort study. Ann Behav Med 52(7):582–593

Leschak CJ, Eisenberger NI (2019) Two distinct immune pathways linking social relationships with health: inflammatory and antiviral processes. Psychosom Med 81(8):711

Lleras-Muney A (2005) The relationship between education and adult mortality in the United States. Rev Econ Stud 72(1):189–221

Malter F, Börsch-Supan A (2017) SHARE wave 6: panel innovations and collecting dried blood spots. Munich Center for the Economics of Aging (MEA), Munich

Manzoli L, Villari P, Pirone GM, Boccia A (2007) Marital status and mortality in the elderly: a systematic review and meta-analysis. Soc Sci Med 64(1):77–94

Oster E (2019) Unobservable selection and coefficient stability: theory and evidence. J Bus Econ Stat 37(2):187–204

Powell VD, Abedini NC, Galecki AT, Kabeto M, Kumar N, Silveira MJ (2021) Unwelcome companions: loneliness associates with the cluster of pain, fatigue, and depression in older adults. Gerontol Geriatr Med 7:2333721421997620

Shankar A, McMunn A, Banks J, Steptoe A (2011) Loneliness, social isolation, and behavioral and biological health indicators in older adults. Health Psychol 30(4):377

Shankar A, McMunn A, Demakakos P, Hamer M, Steptoe A (2017) Social isolation and loneliness: prospective associations with functional status in older adults. Health Psychol 36(2):179

Shankar A, Hamer M, McMunn A, Steptoe A (2013) Social isolation and loneliness: relationships with cognitive function during 4 years of follow-up in the English Longitudinal Study of Ageing. Psychosom Med 75(2):161–170

Shemesh J, Schwartz E, Litwin H (2019) Personality and physical health among older Europeans. Health and Socioeconomic Status Over the Life Course: First Results from SHARE Waves. p 43

Steptoe A, Aparna S, Panayotes D, Jane W (2013) Social isolation, loneliness, and all-cause mortality in older men and women. Proc Natl Acad Sci 110(15):5797–5801

Teguo MT, Simo-Tabue N, Stoykova R, Meillon C, Cogne M, Amiéva H, Dartigues J-F (2016) Feelings of loneliness and living alone as predictors of mortality in the elderly: the PAQUID study. Psychosom Med 78(8):904–909

Valtorta NK, Kanaan M, Gilbody S, Ronzi S, Hanratty B (2016) Loneliness and social isolation as risk factors for coronary heart disease and stroke: systematic review and meta-analysis of longitudinal observational studies. Heart. pp heartjnl–2015

Yu B, Steptoe A, Chen L-J, Chen Y-H, Lin C-H, Ku P-W (2020) Social isolation, loneliness, and all-cause mortality in patients with cardiovascular disease: a 10-year follow-up study. Psychosom Med 82(2):208–214