INFORMAL HELPING AND SUBSEQUENT HEALTH AND WELL-BEING IN OLDER U.S. ADULTS

by

Julia Sharon Nakamura

B.S., University of California, Los Angeles, 2020

A THESIS SUBMITTED IN PARTIAL FULFILLMENT OF
THE REQUIREMENTS FOR THE DEGREE OF

MASTER OF ARTS

in

The Faculty of Graduate and Postdoctoral Studies

(Psychology)

THE UNIVERSITY OF BRITISH COLUMBIA

(Vancouver)

August 2022

The following individuals certify that they have read, and recommend to the Faculty of Graduate and Postdoctoral Studies for acceptance, the thesis entitled:

Informal Helping and Subsequent Health and Well-Being in Older U.S. Adults

submitted by Julia Sharon Nakamura in partial fulfillment of the requirements for

the degree of Master of Arts

in Psychology

Examining Committee:

Anita DeLongis, Professor, Psychology, UBC
Supervisor

Frances Chen, Associate Professor, Psychology, UBC
Supervisory Committee Member

Christiane Hoppmann, Professor, Psychology, UBC
Supervisory Committee Member

Abstract

While there is a large and growing body of evidence around the health and well-being benefits of formal volunteering, less is known about the downstream benefits of informal helping behaviors. A small literature has evaluated associations between informal helping and health and well-being outcomes. However, epidemiological studies have not evaluated if *changes* in informal helping are associated with subsequent health and well-being. Using data from 12,998 participants in the Health and Retirement Study, a national cohort of US adults aged >50, we evaluated if changes in informal helping (between t_0;2006/2008 and t_1;2010/2012) were associated with 35 indicators of physical, behavioral, and psychosocial health and well-being (at t_2;2014/2016). Over the four-year follow-up period, informal helping $\geq$100 (versus 0) hours/year was associated with a 32% lower mortality risk (95% CI [0.54, 0.86]), and improved physical health (e.g., 21% reduced risk of stroke (95% CI [0.64, 0.96])), health behaviors (e.g., 11% increased physical activity (95% CI [1.04, 1.19])), and psychosocial outcomes (e.g., higher purpose in life (β=0.13, 95% CI [0.08, 0.18])). However, there was also evidence that informal helping was associated with higher negative affect (β=0.06, 95% CI [0.003, 0.12]) and little evidence of associations with various other outcomes. In secondary analyses, we adjusted for formal volunteering and caregiving and results were largely unchanged. Encouraging informal helping may improve various aspects of health and well-being and also promote societal well-being.

Lay Summary

Growing evidence suggests that informal helping (unpaid volunteering not coordinated by an organization or institution) is associated with improved health and well-being outcomes. However, studies have not investigated whether *changes* in informal helping are associated with subsequent health and well-being. We evaluated if changes in informal helping were associated with 35 indicators of physical, behavioral, and psychosocial health and well-being using data from 12,998 participants in the Health and Retirement study — a national cohort of US adults aged >50. Over the four-year follow-up period, informal helping ≥100 (versus 0) hours/year was associated with a decreased risk of mortality and improved physical health, health behaviors, and psychosocial outcomes. However, there was also evidence that informal helping was associated with increased negative affect and there was little evidence of associations with other outcomes. Encouraging informal helping may improve various aspects of individuals' health and well-being and also promote societal well-being.

Preface

I am the primary author of the work presented in this thesis. I identified the research question and created plans for data analysis. I was responsible for study conception, data analysis, data interpretation, and manuscript writing. Dr. Eric S. Kim acted as supervisory author by providing guidance and feedback on concept formation, data analysis and interpretation, and manuscript revisions. This study used publicly-available data from the Health and Retirement Study (HRS; http://hrsonline.isr.umich.edu/). The HRS is sponsored by the National Institute on Aging (grant number NIA U01AG009740) and is conducted by the University of Michigan. This study was exempt from additional review by the Institutional Review Board at the University of British Columbia because we used publicly available, de-identified data.

A shorter version of this manuscript is currently under review. Eric S. Kim and Tyler J. VanderWeele were involved with the formation of the project. Eric S. Kim, Tyler J. VanderWeele, and Matthew T. Lee contributed to manuscript drafting.

Table of Contents

Abstract iii

Lay Summary iv

Preface v

Table of Contents vi

List of Tables viii

Acknowledgements ix

Introduction 1

Informal Helping in Older Adults 1

Informal Helping and Health and Well-Being 2

Informal Helping and Physical Health 3

Informal Helping and Health Behaviors 4

Informal Helping and Psychosocial Well-Being 6

Limitations in Prior Research 8

The Present Study 9

Methods 10

Design and Sample 10

Measures 10

Statistical Analysis 12

Results 15

Additional Analyses 16

Discussion 18

Results in the Context of Prior Research 18

Informal Helping vs. Formal Volunteering 21

Formal Volunteering and Caregiving 22

Limitations and Future Directions 22

Conclusion 25

Tables 26

References 34

Appendices 46

Appendix A. Assessment of Outcomes 46

Appendix B. Proof Illustrating How Adjusting for Pre-Baseline Informal Helping Can Help Us Evaluate How "Changes" in Informal Helping are Associated with Subsequent Health and Well-Being Outcomes Over Time 57

Appendix C. Considering Causes of Death 58

Appendix D. Tables 59

List of Tables

Table 1: Characteristics of Health and Retirement Study Participants at Pre-Baseline by Baseline Categories of Informal Helping (N=10,956).....26

Table 2: Changes in Informal Helping and Subsequent Health and Well-being (Health and Retirement Study [HRS]: N=12,998) 29

Table 3: Robustness to Unmeasured Confounding (E-Values) for the Associations Between Informal Helping (≥100 hours vs. 0 hours) and Subsequent Health and Well-Being (N=12,998).....32

Table A1: Changes in Informal Helping from the Pre-Baseline Wave (t_0) to the Baseline Wave (t_1) 59

Table A2: Informal Helping and Subsequent Health and Well-being (After Adjustment for Conventional Covariates or All Covariates; Health and Retirement Study [HRS]: N =12,998).....60

Table A3: Complete-Case Analyses: Informal Helping and Subsequent Health and Well-being (Health and Retirement Study [HRS]: N ranged from: 5,576 to 8,482).....63

Table A4: Informal Helping and Subsequent Health and Well-being (Health and Retirement Study [HRS]: N=12,998) – Baseline Adjustment for Formal Volunteering and Caregiving 65

Acknowledgements

Firstly, I would like to thank my advisor, Dr. Eric S. Kim, for his support and mentoring throughout my Master's degree. He has challenged me and helped me grow in both my research and professional development in meaningful ways.

I would also like to give thanks to my supervisory committee members. A special thank you to Dr. Anita DeLongis, Dr. Frances Chen, and Dr. Christiane Hoppmann for their insightful feedback on my thesis project, their personal support, and encouragement throughout my time in the Health Psychology Program.

I would also like to thank Dr. Tyler VanderWeele and the Harvard Human Flourishing Program for their time and research support during my time working with them over the past few years. Further, I'd like to express many thanks to Dr. Kate Kuhlman, my undergraduate advisor, for encouraging me to pursue graduate studies, for helping me develop the fundamental skills needed for graduate school, and for her continued mentoring.

Thank you to all of our co-op students, lab managers, and research assistants at the Psychosocial Flourishing and Health Lab for their encouragement and support of the lab's research.

Additionally, I am grateful for the generous funding provided by the UBC Faculty of Arts, UBC Department of Psychology, and the Social Sciences and Humanities Research Council of Canada, which helped support me during the course of my Master's program.

Finally, I would like to thank my mom and dad, my two sisters (Jane and Jenn), and my friends, who have supported me in my pursuit of a graduate degree. I am honored and immensely grateful to have such an incredible support system.

Introduction

Prosocial behaviors are a key building block of society at the micro- (e.g., individual origins of prosociality), meso- (e.g., helper-recipient dyads), and macro- (e.g., volunteering in the context of groups and large organizations) levels (Penner et al., 2005). While there is a large and growing body of evidence around the health and well-being benefits of formal volunteering (Burr et al., 2021; Kim et al., 2020), less work has evaluated the downstream benefits of the most common type of prosocial behavior: informal helping. Informal helping (i.e., informal volunteering) is defined as, "unpaid volunteering not coordinated by an organization or institution" (e.g., babysitting, cooking meals, providing transportation, etc.) directed toward helping people outside of one's household (Einolf et al., 2016; Gottlieb, 1978). Three partially overlapping perspectives offer insight into why people participate in informal helping: 1) social capital (helping establish networks, norms, and trust that facilitates cooperation among people), 2) social network theory (reciprocity for helping others within one's social network), and 3) evolutionary biology (informal helping is present in all human societies, and evolutionary biologists consider it an important aspect of our evolution as social animals; Einolf et al., 2016). Further, there may be altruistic, spiritual, and/or moral reasons for engaging in informal helping behaviors. Yet, little is known about how informal helping is associated with downstream health and well-being.

Informal Helping in Older Adults

Four factors converge to underscore the importance of studying informal helping in older adults. First, populations are rapidly aging in many countries throughout the world (United Nations, Department of Economic and Social Affairs, Population Division, 2020). With rapid population aging, researchers and policymakers are increasingly interested in health assets that

enhance our society's health and well-being (Kubzansky et al., 2018; VanderWeele, 2017), and informal helping may be a promising health asset to consider. Second, informal helping is often more accessible than formal volunteering for various subgroups of people (Einolf et al., 2016; Martinez et al., 2011). For example, it is difficult for many in lower socioeconomic situations to formally volunteer because for people who work labor-intensive jobs and/or multiple jobs, there is little energy and time to formally volunteer (Martinez et al., 2011). Further, for older adults constrained by physical health conditions (e.g., people in nursing homes with limited mobility) who want to act on their intrinsic prosocial desire, informal helping is an accessible outlet. Third, there are also economic benefits to informal helping. Though work assigning monetary value to informal volunteering is rare, prior research has found that informal helping contributes more to the economy than formal volunteering (Einolf et al., 2016). Fourth, informal helping is more common than formal volunteering, and people in many countries (e.g., the United States, United Kingdom, and France) spend more time informally helping than formally volunteering (Einolf et al., 2016). Because many people are already engaging in informal helping, from a public health perspective, it is important to understand the downstream health and well-being consequences of these common informal helping behaviors (e.g., if increased informal helping is associated with increased depressive symptoms because it is perceived as stressful or overwhelming, it would be important to know this).

Informal Helping and Health and Well-Being

Hui et al. (2020), in the first meta-analysis to compare formal volunteering and informal helping, found that informal helping was more strongly associated with overall well-being. The authors propose that the difference between formal volunteering and informal helping might be understood through self-determination theory (Deci & Ryan, 2000), with informal helping

(which is potentially more casual, freely chosen, and spontaneous) fulfilling more basic psychological needs for autonomy and relatedness (Hui et al., 2020) than formal volunteering. In the sections below, we evaluate associations between informal helping and physical-, behavioral-, and psychosocial- health and well-being outcomes.

Informal Helping and Physical Health

Physical health conditions are a leading cause of death worldwide. For example, in 2019, cardiovascular diseases, cancers, and respiratory diseases were the leading causes of death worldwide for older adults aged 50-69 and ≥70 (Ritchie & Roser, 2018). Informal helping has been associated with improved physical health outcomes, but with mixed findings. While some studies find that informal helping is associated with improved subsequent physical health (e.g., a decreased risk of mortality), others do not find associations between informal helping and physical health outcomes (e.g., cardiovascular disease (CVD)), and some studies find that associations are conditional on key moderating variables (e.g., gender).

Informal helping has been associated with a reduced risk of mortality (Han et al., 2017; Qu et al., 2020). For example, in a longitudinal study of 10,841 older adults in the Health and Retirement Study (HRS), informal helping at a moderate amount (1 to 99 hours or ≥100 hours/year) was associated with a lower risk of non-CVD mortality (Han et al., 2017). Likewise, in a longitudinal study of older adults in the Wisconsin Longitudinal Study (N = 3,544), providing support to friends, neighbors, or coworkers (e.g., via transportation, errands or shopping, etc.) was associated with a lower risk of mortality (Qu et al., 2020).

When assessing associations between informal helping and other physical health outcomes (e.g., cardiovascular disease and self-rated health), findings are mixed. Han et al. (2017) (described above) observed that informal helping was not associated with incident CVD.

Conversely, in a longitudinal study of 11,418 adults aged ≥51 in the HRS, informal helping was associated with a lower risk of CVD, but only for men (Burr et al., 2018). Men who engaged in informal helping had a lower risk of incident CVD as compared to men who did not engage in informal helping. Informal helping has also been associated with higher self-reported health in some groups. In a longitudinal study of 681 older adults, informal volunteering directed toward fellow church members (e.g., tangible support such as household chores or yardwork, transportation, etc.) was associated with higher self-rated health, but only for participants who were more deeply committed to their faith (Krause, 2009).

While it is plausible that informal helping is associated with improved physical health outcomes, there is limited evidence to support this claim due to a lack of prior research on this topic. Most studies look at one (or a limited number of) health outcome(s) and many relevant physical health outcomes in older adults have not been assessed. For example, informal helping may decrease the risk of physical functioning limitations as a result of increased physical engagement (e.g., through running errands, buying groceries, etc.). Likewise, informal helping may decrease the risk of subsequent cognitive impairment (due to increased cognitive activity and engagement with the individuals being helped). We now shift to assess associations between informal helping and health behaviors.

Informal Helping and Health Behaviors

Health behaviors include a range of behaviors that explain about 40% of premature mortality as well as substantial morbidity in the United States (McGinnis et al., 2002). Some of these health behaviors (e.g., physical activity) promote health, while others (e.g., smoking, alcohol consumption) undermine health. Prior work shows that social relationships may influence health behaviors (Umberson & Montez, 2010) in various ways. To name a few,

religious ties might influence health behaviors through social control, social ties can instill a sense of responsibility for others that might lead individuals to protect the health of others as well as their own, and social ties might provide norms and information that influence behaviors (Umberson & Montez, 2010). Applying this framework (in which social ties influence health behaviors) to informal helping, informal helpers might engage in healthier behaviors and reduce behaviors that undermine health (e.g., due to a sense of responsibility for others or the nature of informal helping activities themselves).

While some prior work has established associations between prosocial behaviors (e.g., volunteering) and improved health behaviors (Burr et al., 2021), little research has assessed associations between informal helping and health behaviors. Han et al. (2017), mentioned previously, assessed three health behaviors (smoking, physical activity, and alcohol consumption) as mediators of the informal helping – non-CVD mortality and informal helping – incident CVD associations. Physical activity mediated the relationship between informal helping and incident CVD and non-CVD mortality, smoking mediated the association between informal helping and non-CVD mortality, and alcohol consumption did not mediate the associations between informal helping and either non-CVD mortality or incident CVD risk (Han et al., 2017). Building on this prior research investigating the mechanisms through which informal helping is associated with improved physical health, we recognize that health behaviors may be outcomes of interest in and of themselves (i.e., not through mediation), due to their associations with mortality and morbidity in later life. To our knowledge, associations between informal helping and some plausible health behaviors (e.g., sleep problems) have not been studied at all. While little work has assessed associations between informal helping and health behaviors, more research has assessed associations between informal helping and psychosocial outcomes.

Informal Helping and Psychosocial Well-Being

Psychological (Kubzansky et al., 2018; VanderWeele, 2017) and social (Holt-Lunstad et al., 2010; House et al., 1988) well-being are increasingly being understood as important contributors to later-life health outcomes, and are also desired by many for their own sake. As compared to physical- and behavioral- health and well-being outcomes, a larger literature has assessed associations between informal helping and psychosocial factors (e.g., psychological well-being- (e.g., life satisfaction), psychological distress- (e.g., depressive symptoms), and social-factors (e.g., loneliness)).

Several studies have found associations between informal helping and improved psychological well-being (Appau & Awaworyi Churchill, 2019; Kahana et al., 2013; Matthews & Nazroo, 2021). In a study of 23 countries from the European Social Survey, after adjustment for sociodemographic variables, informal helping was associated with improved psychological well-being across hedonic indicators (e.g., happiness and positive affect) and eudaimonic indicators (e.g., accomplishment and worthwhileness; Plagnol & Huppert, 2010). These findings have been observed in other populations. In a cross-sectional study from the Community Life Survey in the UK, informal helping was associated with higher subjective well-being (measured via: life satisfaction, happiness, worthwhileness in this study; Appau & Awaworyi Churchill, 2019). Likewise, Kahana et al. (2013), in a study of 585 older adults in a two-wave longitudinal study using data from the Elderly Care Research Center (based on a large retirement community in West Florida), observed that informal helping was associated with higher life satisfaction and positive affect. Matthews & Nazroo (2021), in a study of 3,740 participants in the English Longitudinal Study of Ageing, observed that engagement in two or more informal helping activities was associated with improved life satisfaction and quality of life (assessed as control,

autonomy, self-realization, and pleasure). Conversely, Plagnol & Huppert (2010) did not observe associations between informal helping and life satisfaction, and other findings are mixed. For example, in one cross-sectional study of adults aged >60 from the Americans' Changing Lives Panel Study, informal volunteering was associated with increased positive affect among Black men and White women only (McIntosh & Danigelis, 1995).

Prior work has assessed associations between informal helping and some psychological distress outcomes (e.g., depressive symptoms and negative affect). Associations between informal helping and depressive symptoms are mixed. One cross-sectional study of 14,477 respondents (aged 52-66) from the French GAZEL cohort study found that informal helping at all amounts (less often, almost every week, almost daily) vs. not at all was associated with lower depressive symptoms (Wahrendorf et al., 2008). Matthews & Nazroo (2021) also observed that informal helping was associated with lower depressive symptoms for participants who participated in one or more informal helping activities. Conversely, at least two other longitudinal studies have not found associations between informal helping and lower depression (Kahana et al., 2013; Li & Ferraro, 2005). Li & Ferraro (2005), in a three-wave longitudinal study (over 8 years) of older adults from the Americans' Changing Lives study, observed that informal helping was not associated with a reduced risk of depression. Likewise, Kahana et al. (2013) observed that informal helping was not associated with depressive symptoms. Associations between informal helping and depression may depend on the length of follow-up. In a longitudinal study of 2,688 older adults from the HRS, informal helping was initially associated with a small increase in depressive symptoms (for participants with dual sensory loss and the comparison group), but this effect was reduced over time. After the initial increase in depressive symptoms, informal helping was associated with decreased depressive symptoms

over time (McDonnall, 2011). With regards to negative affect, McIntosh & Danigelis (1995) observed that informal volunteering was associated with decreased negative affect among older Black men and White women only, and Kahana et al. (2013) and Plagnol & Huppert (2010) observed that informal helping was not associated with negative affect.

While social well-being is health promoting, few studies have assessed associations between informal helping and social outcomes. Matthews & Nazroo (2021) observed that engaging in two or more informal helping activities was associated with lower loneliness, but there is little prior literature assessing associations between informal helping and other social outcomes (e.g., frequency of contact with one's social network).

Through these key prior studies on informal helping and psychosocial well-being we see the importance of many future directions. Again, here we see that most studies look at one (or a limited number of) psychological outcome(s), some studies are still cross-sectional, and many relevant psychosocial outcomes in older adults have not been assessed. For example, informal helping may be associated with increased purpose in life (e.g., through purpose found in helping others), and purpose in life has been associated with improved subsequent health and well-being outcomes (Kim et al., 2021). Similarly, purpose in life, like many psychosocial outcomes, is desired by many for its own sake, giving one's activities and life goals a broader context (Hanson & VanderWeele, 2021; Lee et al., 2021; VanderWeele, 2017).

Limitations in Prior Research

These past observational studies helped break exciting new ground. However, there is relatively little prior research evaluating associations between informal helping and health and well-being outcomes and many of these studies have limitations that we aim to improve upon. Though these limitations were mentioned above, we summarize them here. First, while an

increasing number of studies are longitudinal, many are cross-sectional, making it difficult to assess the direction of causality (Einolf et al., 2016). Does informal helping protect against depression, or are depressed people less likely to informally help? Cross-sectional data cannot provide evidence concerning this question. Second, many studies have not yet evaluated a wide range of outcomes that holistically assess healthy aging (e.g., across physical-, behavioral-, and psychosocial domains). Third, most studies do not adjust for prior informal helping, or prior outcomes, in the pre-baseline wave. Pre-baseline adjustment for informal helping helps readers evaluate how *changes* in informal helping are associated with health and well-being.

The Present Study

Here we asked the question, if informal helping were increased, what improvements to health and well-being outcomes might we observe within a relatively short time frame (i.e., a 4-year follow up period)? To begin addressing this question, we used the new *outcome-wide* analytic approach which is described further in the *Analysis* section (VanderWeele et al., 2020). Using this approach, we examined if changes in informal helping were associated with better subsequent health and well-being across 35 separate outcomes, including indicators of physical health, health behaviors, and psychosocial factors. Outcome-wide analyses are a hypothesis-generating, data-driven approach aimed at discovering promising health and well-being outcomes associated with various exposures, such as increased informal helping, which may then undergo further investigation in future studies. These outcomes were chosen because they are frequently included in the conceptualization of seminal gerontological models that characterize the antecedents, processes, and outcomes that foster people's ability to age well (Aldwin & Igarashi, 2015; Depp & Jeste, 2006; Reich et al., 2010; Rowe & Kahn, 1987; Ryff & Singer, 2009).

Methods

Design and Sample

Data were from the HRS, a national sample of adults aged >50 in the United States. Approximately 50% of HRS respondents were randomly selected for an enhanced face-to-face (EFTF) interview in 2006 when most psychosocial factors were first assessed and the other half of respondents were assessed in 2008. After the interview, participants completed a psychosocial questionnaire which they mailed to the University of Michigan upon completion (88% response rate in 2006 and 84% response rate in 2008; Smith et al., 2017). These sub-cohorts alternate reporting of psychosocial factors: each participant reports psychosocial data every 4 years. To increase sample size and statistical power, data from 2006 and 2008 were combined. Participants were excluded if they did not report psychosocial data at baseline because over half of the study outcomes were included in this assessment, resulting in a final sample of 12,998 participants.

This study used data from three timepoints: 1) Covariates were assessed in the pre-baseline wave (t_0;2006/2008), 2) the exposure (informal helping) was assessed in the baseline wave (t_1;2010/2012), and 3) outcomes were assessed in the outcome wave (t_2;2014/2016). Further details about this study can be found on the HRS website (http://hrsonline.isr.umich.edu/). This study was exempt from additional review by the Institutional Review Board (blinded for peer review) because we used publicly available, de-identified data.

Measures

Informal Helping

Respondents were asked: “Have you spent any time in the past 12 months helping friends, neighbors, or relatives who did not live with you and did not pay you for the help?” If

they answered yes to this question, respondents were asked how many hours they volunteered: 1−49 hours, 50−99 hours, 100−199 hours, or ≥ 200 hours. Based on past research suggesting that approximately 100 hours/year of formal volunteering may be an inflection point for improved health and well-being, the top 2 informal helping groups were collapsed in the main analyses to increase statistical power (Johnson & Post, 2017).

Covariates

We adjusted for a wide range of covariates in the pre-baseline wave (t_0;2006/2008). Covariates included: sociodemographic factors (age (continuous), gender (male/female), race/ethnicity (White, African-American, Hispanic, Other), marital status (married/not married), income (<$50,000, $50,000-$74,999, $75,000-$99,999, ≥$100,000), total wealth (based on quintiles of the score distribution for total wealth in this sample), educational attainment (no degree, GED/high school diploma, ≥college degree), employment status (yes/no), health insurance (yes/no), geographic region (Northeast, Midwest, South, West), religious service attendance (none, <1x/week, ≥1x/week), personality (openness, conscientiousness, extraversion, agreeableness, neuroticism; continuous), and childhood abuse (yes/no)). We also adjusted for informal helping and all outcome variables in the pre-baseline wave.

Outcomes

We evaluated 35 outcomes in the outcome wave (t_2:2014/2016). These included measures of: physical health (all-cause mortality, number of chronic conditions, diabetes, hypertension, stroke, cancer, heart disease, lung disease, arthritis, overweight/obesity, physical functioning limitations, cognitive impairment, chronic pain, self-rated health), health behaviors (heavy drinking, smoking, physical activity, sleep problems), psychological well-being (positive affect, life satisfaction, optimism, purpose in life, mastery, health mastery, financial mastery),

psychological distress (depression, depressive symptoms, hopelessness, negative affect, perceived constraints), and social factors (loneliness, living with a spouse/partner, frequency of contact with 1) children, 2) other family, and 3) friends). The appendix (Appendix Text 1) and HRS materials provide further details about each of these variables (Fisher et al., 2005; Jenkins et al., 2008; Smith et al., 2017).

Statistical Analysis

The outcome-wide analytic approach uses several analytic decisions not widely implemented outside of biostatistics and causal inference, thus we summarize these decisions here (VanderWeele et al., 2020). First, it is difficult to discern whether covariates are confounders or mediators if the covariate measurements used in analyses are assessed at the same timepoint as the exposure (t_1: 2010/2012; VanderWeele et al., 2020). Thus, to reduce this concern and allow for a comprehensive set of covariates to address confounding, we adjusted for covariates in the pre-baseline wave (t_0: 2006/2008). Second, we adjusted for all outcome variables in the pre-baseline wave (t_0) to reduce the likelihood of reverse causality. Third, to evaluate "*changes*" in informal helping, we adjusted for informal helping in the pre-baseline wave (t_0). This helps "hold constant" pre-baseline levels of informal helping. Participants who start in the $\geq$100-hour informal helping group in the pre-baseline wave (t_0) and remain there in the baseline wave (t_1) contribute to the final estimate. However, this estimate also corresponds to participants who start in the 0-hour informal helping group in the pre-baseline wave and move to the $\geq$100-hour group in the baseline wave. The model effectively assumes that there is no interaction between past and current informal helping (i.e., the $\geq$100-hour group coefficient is constant across past informal helping levels). Thus, we can evaluate how *change*s in informal helping (between t_0 and t_1) are associated with later health and well-being outcomes in the

outcome wave (at t_2: see Appendix Text 2). Adjusting for pre-baseline informal helping (t_0) has several other advantages. First, it reduces risk of reverse causality by "removing" the potential accumulating effects that informal helping might have had on health and well-being outcomes in the past (prevalent exposure). Second, it allows us to focus on how *changes* in informal helping (incident exposure) affect outcomes. Therefore, there is a focus on how changes in informal helping are associated with short-term changes in health and well-being outcomes.

We used an outcome-wide approach (VanderWeele et al., 2020) and ran separate models for each outcome. Depending on the nature of the outcome, we used 3 different models: 1) logistic regression for each binary outcome with a prevalence <10%, 2) generalized linear models (with a log link and Poisson distribution) for each binary outcome with a prevalence ≥10%, and 3) linear regression for each continuous outcome. We also standardized all continuous outcomes (mean=0, standard deviation=1) (so their effect sizes can be interpreted as a standard deviation change in the outcome variable) and marked multiple p-value cutoffs in our tables (including those making Bonferroni correction) because practices for multiple testing vary widely and are continuously evolving (Dunn, 1961; VanderWeele & Mathur, 2019).

Additional Analyses

We conducted several additional analyses. 1) We conducted E-value analyses. E-values allow us to evaluate the robustness of our results to unmeasured confounding by assessing the minimum strength that unmeasured confounder(s) would have to have on the risk ratio scale (with both informal helping and the outcome) to entirely explain away the association between informal helping and the outcome (VanderWeele & Ding, 2017). 2) We re-analyzed all models using a reduced list of conventional covariates (only sociodemographic factors) in the pre-baseline wave. This approach (i.e., not adjusting for prior informal helping) might crudely assess

the potential cumulative effects that the whole history of informal helping has on outcomes. 3) We re-analyzed models after removing people with a history of a given physical condition at baseline. 4) We re-analyzed all models using only complete-cases to assess the impact of multiple imputation on results. 5) We re-analyzed all models after additionally adjusting for both formal volunteering and caregiving at baseline, to evaluate the effects of informal helping above and beyond formal volunteering or caregiving (see Appendix Text 1 for formal volunteering and caregiving measures).

Multiple Imputation

All missing exposures, covariates, and outcomes were imputed using imputation by chained equations, and 5 datasets were created. This method generally provides a more flexible approach than other methods of handling missing data and addresses problems that arise from attrition (Harel et al., 2018).

Results

In the pre-baseline wave (t_0;2006/2008), participants were on average 65 years old (*SD*=10), predominantly women (59%) and married (67%). Table 1 provides the distribution of covariates by informal helping, and Table A1 shows the changes in informal helping from the pre-baseline wave (t_0) to the baseline wave (t_1).

Over the 4-year follow-up period, participants engaging in informal helping $\geq$100 hours/year (versus 0 hours/year), conditional on prior informal helping, had a 32% reduced risk of mortality (95% CI [0.54, 0.86]; Appendix Text 3), 21% reduced risk of stroke (95% CI [0.64, 0.96]), 20% reduced risk of physical functioning limitations (95% CI [0.70, 0.91]), 18% reduced risk of cognitive impairment (95% CI [0.69, 0.97]), and higher self-rated health (β=0.10, 95% CI [0.04, 0.15]; Table 2). There was less evidence of associations between informal helping and other physical health outcomes, including: number of chronic conditions and risk of diabetes, hypertension, cancer, heart disease, lung disease, arthritis, overweight/obesity, and chronic pain.

Among health behaviors, participants engaging in informal helping $\geq$100 hours/year (versus 0 hours/year), conditional on prior informal helping, had an 11% increased likelihood of engaging in frequent physical activity (95% CI [1.04, 1.19]) four years later. There was little evidence of associations between informal helping and heavy drinking, smoking, or sleep problems.

Amongst psychological factors, for psychological well-being, participants engaging in informal helping $\geq$100 hours/year (versus 0 hours/year), conditional on prior informal helping, had higher positive affect (β=0.09, 95% CI [0.02, 0.17]), optimism (β=0.09, 95% CI [0.03, 0.15]), purpose in life (β=0.13, 95% CI [0.08, 0.18]), mastery (β=0.10, 95% CI [0.03, 0.16]), and health mastery (β=0.07, 95% CI [0.0003, 0.15]). Among psychological distress factors,

participants engaging in informal helping ≥100 hours/year (versus 0 hours/year) had lower hopelessness (β=-0.06, 95% CI [-0.12, -0.004]) and higher negative affect (β=0.06, 95% CI [0.003, 0.12]). However, there was little evidence of associations between informal helping and life satisfaction, financial mastery, depression, depressive symptoms, and perceived constraints.

Finally, amongst social factors, participants engaging in informal helping ≥100 hours/year (versus 0 hours/year), conditional on prior informal helping, had a 13% increased likelihood of frequent contact with friends (95% CI [1.04, 1.24]). However, there was little evidence of associations between informal helping and loneliness, living with a spouse or partner, or contact with children or other family.

Additional Analyses

Concerning the additional analyses, first, E-values suggested that many of the observed associations were moderately robust to unmeasured confounding (Table 3). For example, for mortality, an unmeasured confounder associated with both informal helping and mortality by risk ratios of 2.30 each (above and beyond the covariates already adjusted for) could explain away the association, but weaker joint confounder associations could not. Further, to shift the CI to include the null, an unmeasured confounder that was associated with both informal helping and mortality by risk ratios of 1.60 each could suffice, but weaker joint confounder associations could not. Second, adjustment for conventional covariates showed mostly larger estimates than found in the fully adjusted models (Table A2). Third, after removing anyone with a history of a given physical condition at baseline, estimates were generally similar (Table A2). Fourth, complete-case analyses showed similar results to those from the main imputed analyses (Table A3). Fifth, results from analyses that additionally adjusted for both formal volunteering and caregiving at baseline (Table A4) showed similar results to the main imputed analyses (with the

exception of stroke and some psychological well-being, psychological distress, and social factors), suggesting that informal helping is associated with improved subsequent health and well-being above and beyond the effects of formal volunteering and caregiving.

Discussion

In a large, longitudinal, and national sample of US adults aged >50, informal helping ≥100 hours/year (versus 0 hours/year) was associated with improved physical health (e.g., decreased risk of: mortality, stroke, physical functioning limitations, cognitive impairment; and increased self-rated health), health behaviors (e.g., increased physical activity), psychological well-being (e.g., increased: positive affect, optimism, purpose in life, mastery, and health mastery), psychological distress (e.g., decreased hopelessness) and social outcomes (e.g., more frequent contact with friends). However, informal helping was also associated with slightly increased psychological distress (e.g., higher negative affect), and showed little evidence of associations with other physical health factors (e.g., number of chronic conditions, diabetes, hypertension, cancer, heart disease, lung disease, arthritis, overweight/obesity, and chronic pain), health behaviors (e.g., heavy drinking, smoking, and sleep problems), psychological well-being (e.g., life satisfaction, financial mastery), psychological distress (e.g., depression, depressive symptoms, perceived constraints), and social factors (e.g., loneliness, living with a spouse or partner, contact with children and other family).

Results in the Context of Prior Research

Our findings converge with findings from some prior studies (e.g., associations between informal helping and mortality) and diverge from other studies (e.g., associations between informal helping and life satisfaction) as discussed more below. Findings may have diverged for many reasons, including: 1) different measurements to assess the exposure and outcomes, 2) differences in sample composition (e.g., some prior studies were conducted in specific samples, such as older adults with dual sensory loss; McDonnall, 2011), 3) we adjusted for an extensive set of covariates while most other studies used a more limited set of covariates (e.g.,

sociodemographic covariates only), 4) many previous studies were cross-sectional, and 5) most did not adjust for pre-baseline informal helping and outcomes (which allowed us to look at changes in informal helping (through pre-baseline adjustment for informal helping) and mitigate concerns of reverse causality (through pre-baseline adjustment for outcomes)).

Regarding physical health outcomes, our findings mostly converge with those in the prior literature. Our findings align with prior work in which informal helping was associated with a reduced risk of mortality (Han et al., 2017; Qu et al., 2020) and higher self-rated health (Krause, 2009), but not associated with CVD (Han et al., 2017). While Burr et al. (2018) found associations between informal helping and incident CVD, these associations were only observed for men. Most prior work on physical health outcomes, in line with our work, has consisted of longitudinal studies of aging in large samples (Burr et al., 2018; Han et al., 2017; Qu et al., 2020). It is plausible that informal helping is associated with improved physical health, as this is in line with work on other prosocial behaviors (e.g., volunteering) and improved physical health (Burr et al., 2021). When further adjusting for volunteering and caregiving, we observed that associations between informal helping and most physical health outcomes did not change substantially (though we no longer see an association between informal helping and stroke, the effect estimate is of similar magnitude and $p = 0.051$), suggesting that associations between informal helping and some physical health outcomes may be robust. Future work will need to investigate why other prosocial behaviors (e.g., volunteering) are associated with some physical health outcomes (e.g., cardiovascular disease, hypertension; Burr et al., 2021) while informal helping was not in our study.

Among health behaviors, we observed that informal helping was associated with an increased likelihood of engaging in frequent ($\geq$1x/week) physical activity. Prior work has found

that physical activity mediated the relationship between informal helping and incident CVD and non-CVD mortality (Han et al., 2017), but little prior research has assessed health behaviors as outcomes directly. In our study, informal helping was not associated with heavy drinking, smoking, or sleep problems.

Our findings converge with associations between informal helping and some psychosocial factors (e.g., associations between informal helping and higher positive affect (Kahana et al., 2013) and no associations between informal helping and depression (Kahana et al., 2013; Li & Ferraro, 2005)), but diverge from prior work which has observed associations between informal helping and higher life satisfaction (Appau & Awaworyi Churchill, 2019; Kahana et al., 2013; Matthews & Nazroo, 2021), lower depression (Matthews & Nazroo, 2021; Wahrendorf et al., 2008), and loneliness (Matthews & Nazroo, 2021). Our findings (in which informal helping was associated with increased negative affect) further diverge from prior work which found that informal helping was not associated with negative affect, or was associated with decreased negative affect (Kahana et al., 2013; McIntosh & Danigelis, 1995), though associations with increased negative affect may be attributable to caregiving (see Table A4). Some of these differences from prior studies may be explained by the reasons listed above (different measures of exposures and outcomes, different samples, robustness of covariate adjustment, and the longitudinal nature of our study with pre-baseline adjustment for informal helping and outcomes). In the case of depression and depressive symptoms, for example, the null results here using longitudinal data may indicate that prior cross-sectional studies found a protective association, not necessarily because informal helping protects against depression, but because people with depression are less likely to provide help to others.

Informal Helping vs. Formal Volunteering

While informal helping may appear to be a similar construct to formal volunteering, informal helping is differentially associated with (and conceptually distinct from) formal volunteering (Kim et al., 2020). For instance, informal helping likely has more reciprocity and obligation engineered into the exchanges. Here we comment on a few key differences that we observed between informal helping and formal volunteering. For instance, compared with findings that were also based on HRS data (Kim et al., 2020), informal helping was associated with improvements in a greater number of psychological well-being factors than formal volunteering, while formal volunteering was associated with improvements in more psychological distress outcomes. Perhaps when informally helping people that one knows, helpers receive more consistently positive feedback than that they might receive from formally volunteering to serve strangers (who may give mixed or no reactions). In regard to social factors, formal volunteering was associated with lower loneliness, while informal helping was not. Informal helping (e.g., picking up groceries for a neighbor) might provide less opportunity to form new relationships than formal volunteering, though we do observe associations between informal helping and increased objective contact with friends. If people are trying to decrease loneliness and increase social contact, formal volunteering might be a more intentional way of doing this, though there are avenues of informal helping that appear to increase the frequency of contact with others in one's existing social network. Though some differences between informal helping and formal volunteering are apparent, some key similarities (e.g., decreased risks of mortality, increased likelihood of engaging in physical activity, etc.) suggest that there may be common systems and mechanisms linking prosocial acts more generally with health and well-being. For example, the caregiving system model, which provides neurophysiological

mechanistic explanations for how helping others affects physical health and longevity, suggests that helping others triggers the release of hormones that downregulate the harmful effects of stress (e.g., inflammation and the immune response; Burr et al., 2021). Further, helping close others has been shown to buffer against the harmful effects of stress, and there are several plausible psychosocial mechanisms explaining this finding (e.g., increased: sense of meaning or mattering, social well-being, and opportunities for generativity; Poulin et al., 2013).

Formal Volunteering and Caregiving

In supplementary analyses, we further adjusted for formal volunteering and caregiving to isolate the effects of informal helping, and our results with health and well-being outcomes were largely maintained. However, for psychological factors, we no longer observed associations between informal helping and increased: positive affect, health mastery, and negative affect, or decreased: hopelessness. It is possible that caregiving for a relative could increase both positive and negative affect, and thus, caregiving may have accounted for the association between increased informal helping and increased negative affect. Adjusting for formal volunteering may have also accounted for associations between informal helping and formal volunteering and health mastery and hopelessness. Regarding social factors, after adjustment for formal volunteering, there was no longer notable evidence for associations with informal helping. Perhaps, caregiving and formal volunteering are more direct routes to increasing contact with friends, while informal helping behaviors (e.g., purchasing groceries for a neighbor), do not necessarily increase social contact.

Limitations and Future Directions

These findings should be considered in the context of their limitations, which inspire many important future directions. First, nearly all of the physical health outcomes and health

behaviors used in the current study were self-reported, and thus may be susceptible to self-report bias. Future studies should measure these outcomes objectively. However, study participants were blind to this study's hypotheses and reported informal helping before this study was conducted. Second, there is the potential for confounding by third variables. However, we reduced this concern by implementing a longitudinal study design, adjusting for a large number of covariates, and conducting E-value analyses. Third, we don't know the type or quality of informal helping being done (e.g., providing an instrumental service such as transportation or warm, caring emotional support; Erickson & Stacey, 2013; Gottlieb, 1978; Poulin et al., 2013), or the motivation (e.g., doing voluntary acts of kindness or involuntary chores for a sick relative), and these may have different implications for health and well-being. For example, doing chores in an obligatory capacity for a relative might increase negative affect, while voluntarily helping a grateful neighbor with groceries might decrease negative affect. Fourth, we don't know *how* informal helping is influencing health. Some of the associations we observed in our study may act as mechanistic pathways and should be evaluated using formal mediation methods. For example, perhaps with more informal helping, people experience fewer (subjective) physical functioning limitations because helping someone else helps oneself feel more capable (e.g., through an increased sense of mastery). Alternatively, informal helping may increase physical activity (perhaps one is acquiring increased physical activity through running errands/doing chores), and as a result might influence objective physical functioning. Fifth, we do not know for whom informal helping is the most influential for health outcomes. Future studies may benefit from assessing important candidate moderators (e.g., age, socioeconomic status, personality) that influence associations between informal helping and health and well-being outcomes. Informal helping, if associated with improved health and well-being in

traditionally marginalized populations, may be a route to helping others (and experiencing the health benefits of doing so) despite an inability to formally volunteer. In fact, informal helping in some studies is more common among lower SES populations (Einolf et al., 2016). The current study had several notable strengths. We used a large and national sample aged >50. Further, we simultaneously assessed a large number of outcomes using a longitudinal study design that allows for a direct comparison of effect sizes between outcomes.

Conclusion

Research that reveals how to encourage more acts of prosociality is valuable for its own sake and could be deployed to identify pathways for more timely and complete recovery from the stressors associated with COVID-19. Informal helping is a historically understudied, but potentially powerful driver of health and well-being in older adults. Informal helping behaviors, as opposed to formal volunteering, are open for wider segments of our population to engage in and may have positive effects on subsequent health and well-being that rival or exceed the more commonly studied formal volunteering. We also know that prosocial activities can be increased through intervention (e.g., community courses improving compassion and social trust; Krekel et al., 2021) and our findings highlight the effects that we might expect to observe if informal helping interventions and/or policies were implemented at scale. Interventions to promote compassion might also be understood as building the capacity to love others. A focus on the fundamental dignity of everyone as a human person, and promoting teachings and practices of love of neighbor, common to many religious traditions, might also foster informal helping, formal volunteering, and prosociality more generally. In line with calls to consider "the variety of ways in which prosocial behavior can be manifested" (Penner et al., 2005), with further evidence, we hope to encourage more micro-moments of kindness into the lives of many. Promoting informal helping may be a double-pronged approach to: 1) improving the health of the helpers on an individual level, and 2) improving society as a whole.

Table 1. Characteristics of Health and Retirement Study Participants at Pre-Baseline by Baseline Categories of Informal Helping (N=12,964)[a,b,c]

Participant Characteristics	0 hours (n=6,050)		1-49 hours (n=3,594)		50-99 hours (n= 1,649)		≥100 hours (n=1,671)	
	No. (%)	Mean (SD)	No. (%)	Mean (SD)	No. (%)	Mean (SD)	No. (%)	Mean (SD)
Sociodemographic factors								
Age (yr; range: 46-96)		67.5 (10.3)		63.3 (9.2)		63.2 (9.2)		62.6 (8.4)
Female (%)	3873 (64.0)		1840 (51.2)		895 (54.3)		1012 (60.6)	
Race/Ethnicity (%)								
White	4148 (68.6)		2721 (75.7)		1342 (81.4)		1348 (80.7)	
Black	953 (15.8)		523 (14.6)		173 (10.5)		195 (11.7)	
Hispanic	796 (13.2)		259 (7.2)		92 (5.6)		80 (4.8)	
Other	151 (2.5)		90 (2.5)		42 (2.6)		47 (2.8)	
Married (%)	3283 (62.0)		2100 (71.2)		971 (71.0)		987 (70.7)	
Annual Household Income (%)								
<$50,000	3443 (65.2)		1445 (49.2)		616 (45.2)		629 (45.4)	
$50,000-$74,999	762 (14.4)		526 (17.9)		239 (17.6)		256 (18.5)	
$75,000-$99,999	433 (8.2)		333 (11.3)		168 (12.3)		166 (12.0)	
≥$100,000	640 (12.1)		635 (21.6)		339 (24.9)		335 (24.2)	
Total Wealth (%)								
1st Quintile	1357 (25.7)		483 (16.4)		191 (14.0)		174 (12.6)	
2nd Quintile	1143 (21.7)		583 (19.8)		235 (17.3)		225 (16.2)	
3rd Quintile	987 (18.7)		617 (21.0)		269 (19.8)		317 (22.9)	
4th Quintile	955 (18.1)		608 (20.7)		315 (23.1)		315 (22.7)	
5th Quintile	836 (15.8)		648 (22.1)		352 (25.8)		355 (25.6)	
Education (%)								
<High School	1517 (25.1)		381 (10.7)		140 (8.5)		128 (7.7)	
High School	3220 (53.3)		1999 (55.9)		912 (55.6)		923 (55.4)	
≥College	1300 (21.5)		1198 (33.5)		589 (35.9)		614 (36.9)	
Employed (%)	1748 (33.2)		1470 (50.0)		685 (50.3)		686 (49.5)	
Health Insurance (%)	5010 (95.0)		2783 (94.7)		1306 (95.9)		1313 (94.7)	
Geographic Region (%)								
Northeast	827 (15.3)		422 (14.0)		202 (14.4)		230 (16.0)	
Midwest	1336 (24.7)		838 (27.8)		435 (31.0)		411 (28.6)	
South	2262 (41.8)		1191 (39.5)		490 (35.0)		505 (35.2)	
West	993 (18.3)		564 (18.7)		275 (19.6)		289 (20.1)	
Childhood Abuse (%)	303 (6.3)		181 (6.7)		94 (7.6)		106 (8.4)	
Physical Health								

Participant Characteristics	0 hours (n=6,050)		1-49 hours (n=3,594)		50-99 hours (n= 1,649)		≥100 hours (n=1,671)	
	No. (%)	Mean (SD)	No. (%)	Mean (SD)	No. (%)	Mean (SD)	No. (%)	Mean (SD)
Number of chronic conditions (range: 0-8)		2.6 (1.4)		2.3 (1.4)		2.3 (1.3)		2.3 (1.3)
Diabetes (%)	1111 (21.1)		458 (15.6)		191 (14.0)		165 (11.9)	
Hypertension (%)	3048 (57.8)		1486 (50.6)		661 (48.6)		663 (47.9)	
Stroke (%)	393 (7.5)		115 (3.9)		67 (4.9)		57 (4.1)	
Cancer (%)	744 (14.1)		390 (13.3)		164 (12.1)		163 (11.8)	
Heart disease (%)	1196 (22.7)		559 (19.0)		228 (16.8)		233 (16.8)	
Lung disease (%)	496 (9.4)		184 (6.3)		81 (6.0)		76 (5.5)	
Arthritis (%)	3247 (61.6)		1563 (53.2)		719 (52.9)		756 (54.6)	
Overweight/obesity (%)	3720 (71.6)		2109 (72.6)		966 (71.5)		1010 (73.6)	
Physical functioning limitations (%)	1383 (26.2)		416 (14.2)		168 (12.3)		149 (10.8)	
Cognitive impairment (%)	1050 (20.2)		283 (9.8)		96 (7.2)		90 (6.6)	
Chronic pain (%)	1913 (36.3)		919 (31.3)		395 (29.0)		422 (30.5)	
Self-rated health (range: 1-5)		3.1 (1.1)		3.5 (1.0)		3.5 (0.9)		3.6 (1.0)
Health Behaviors								
Heavy drinking (%)	307 (7.1)		206 (8.6)		76 (6.9)		79 (7.2)	
Smoking (%)	710 (13.6)		365 (12.5)		140 (10.4)		158 (11.5)	
Frequent physical activity (%)	3577 (67.8)		2424 (82.6)		1182 (86.9)		1195 (86.3)	
Sleep problems (%)	1185 (42.7)		667 (40.6)		282 (37.1)		321 (38.3)	
Religious Service Attendance (%)								
Never	1402 (26.6)		678 (23.1)		273 (20.0)		268 (19.3)	
<1x/week	1638 (31.1)		958 (32.6)		450 (33.0)		511 (36.9)	
≥1x/week	2230 (42.3)		1302 (44.3)		639 (46.9)		607 (43.8)	
Psychological Well-Being								
Positive affect (range: 1-5)		3.5 (0.8)		3.7 (0.7)		3.7 (0.7)		3.7 (0.7)
Life satisfaction (range: 1-7)		5.0 (1.5)		5.2 (1.4)		5.3 (1.3)		5.3 (1.4)
Optimism (range: 1-6)		4.4 (1.0)		4.6 (0.9)		4.7 (0.9)		4.7 (0.9)
Purpose in life (range: 1-6)		4.5 (0.9)		4.8 (0.9)		4.8 (0.9)		4.9 (0.8)
Mastery (range: 1-6)		4.7 (1.1)		4.9 (1.0)		5.0 (1.0)		4.9 (1.0)
Health mastery (range: 0-10)		7.2 (2.4)		7.5 (2.1)		7.6 (2.0)		7.6 (2.1)
Financial mastery (range: 0-10)		7.4 (2.7)		7.4 (2.4)		7.6 (2.3)		7.4 (2.4)
Psychological Distress								
Depression (%)	819 (15.8)		275 (9.5)		109 (8.2)		119 (8.7)	
Depressive symptoms (range: 0-8)		1.5 (2.0)		1.1 (1.7)		1.0 (1.6)		1.0 (1.7)

Participant Characteristics	0 hours (n=6,050)		1-49 hours (n=3,594)		50-99 hours (n= 1,649)		≥100 hours (n=1,671)	
	No. (%)	Mean (SD)	No. (%)	Mean (SD)	No. (%)	Mean (SD)	No. (%)	Mean (SD)
Hopelessness (range: 1-6)		2.5 (1.3)		2.1 (1.1)		2.0 (1.1)		2.0 (1.1)
Negative affect (range: 1-5)		1.7 (0.7)		1.6 (0.6)		1.6 (0.5)		1.6 (0.6)
Perceived constraints (range: 1-6)		2.3 (1.2)		2.0 (1.0)		1.9 (1.0)		1.9 (1.0)
Social Factors								
Loneliness (range: 1-3)		1.5 (0.6)		1.4 (0.5)		1.4 (0.5)		1.4 (0.5)
Living with a spouse/partner (%)	3082 (65.3)		2026 (75.2)		906 (74.1)		937 (75.1)	
Contact children ≥1x/week (%)	3519 (74.3)		1997 (74.6)		942 (77.3)		985 (78.5)	
Contact other family ≥1x/week (%)	2420 (51.0)		1380 (51.0)		668 (53.8)		718 (56.8)	
Contact friends ≥1x/week (%)	2913 (61.0)		1863 (68.8)		892 (71.5)		907 (71.6)	
Personality								
Openness (range: 1-4)		2.9 (0.6)		3.0 (0.5)		3.1 (0.5)		3.1 (0.5)
Conscientiousness (range: 1-4)		3.3 (0.5)		3.4 (0.4)		3.5 (0.4)		3.5 (0.4)
Extraversion (range: 1-4)		3.1 (0.6)		3.3 (0.5)		3.3 (0.5)		3.3 (0.5)
Agreeableness (range: 1-4)		3.5 (0.5)		3.5 (0.4)		3.6 (0.4)		3.6 (0.4)
Neuroticism (range: 1-4)		2.1 (0.6)		2.0 (0.6)		2.0 (0.6)		2.0 (0.6)

[a]This table was created based on non-imputed data.

[b]All variables in Table 1 were used as covariates, and assessed in the pre-baseline wave (t_0;2006/2008).

[c]The percentages in some sections may not add up to 100% due to rounding.

Table 2. Changes in Informal Helping and Subsequent Health and Well-being (Health and Retirement Study [HRS]: N=12,998)[a,b,c,d]

	Informal Helping			
Outcomes	**0 hours (n=6,070) (Reference)**	**1-49 hours (n=3,603) RR/OR/β (95% CI)**	**50-99 hours (n=1,651) RR/OR/β (95% CI)**	**≥100 hours (n=1,674) RR/OR/β (95% CI)**
Physical Health				
All-cause mortality	1.00	0.70 (0.60, 0.81)***	0.73 (0.59, 0.90)**	0.68 (0.54, 0.86)**
Number of chronic conditions	0.00	-0.04 (-0.06, -0.01)**	-0.04 (-0.07, 0.00)	-0.01 (-0.05, 0.03)
Diabetes	1.00	0.97 (0.89, 1.06)	0.98 (0.87, 1.11)	1.00 (0.88, 1.14)
Hypertension	1.00	0.99 (0.93, 1.04)	0.98 (0.91, 1.06)	1.02 (0.94, 1.09)
Stroke	1.00	0.89 (0.77, 1.02)	0.86 (0.71, 1.04)	0.79 (0.64, 0.96)*
Cancer	1.00	0.94 (0.84, 1.04)	0.89 (0.77, 1.03)	0.92 (0.79, 1.05)
Heart disease	1.00	0.97 (0.89, 1.05)	0.98 (0.87, 1.09)	0.96 (0.85, 1.07)
Lung disease	1.00	0.96 (0.85, 1.10)	1.02 (0.86, 1.22)	0.93 (0.78, 1.11)
Arthritis	1.00	1.00 (0.95, 1.06)	1.01 (0.94, 1.09)	1.03 (0.96, 1.11)
Overweight/obesity	1.00	1.01 (0.96, 1.07)	1.01 (0.94, 1.09)	1.04 (0.96, 1.12)
Physical functioning limitations	1.00	0.85 (0.78, 0.94)***	0.75 (0.65, 0.86)***	0.80 (0.70, 0.91)***
Cognitive impairment	1.00	0.95 (0.86, 1.05)	0.78 (0.67, 0.92)**	0.82 (0.69, 0.97)*
Chronic pain	1.00	1.00 (0.92, 1.08)	0.99 (0.90, 1.10)	0.99 (0.89, 1.09)
Self-rated health	0.00	0.08 (0.04, 0.11)***	0.10 (0.04, 0.15)***	0.10 (0.04, 0.15)***
Health Behaviors				
Heavy drinking	1.00	1.05 (0.80, 1.37)	1.10 (0.79, 1.53)	1.20 (0.88, 1.64)
Smoking	1.00	1.04 (0.90, 1.21)	1.07 (0.87, 1.32)	1.16 (0.95, 1.42)
Frequent physical activity	1.00	1.08 (1.02, 1.15)*	1.12 (1.04, 1.21)**	1.11 (1.04, 1.19)**
Sleep problems	1.00	0.99 (0.92, 1.07)	0.95 (0.86, 1.06)	1.02 (0.91, 1.14)
Psychological Well-Being				
Positive affect	0.00	0.07 (0.02, 0.12)**	0.11 (0.05, 0.17)***	0.09 (0.02, 0.17)*
Life satisfaction	0.00	0.02 (-0.02, 0.07)	0.03 (-0.02, 0.09)	0.01 (-0.05, 0.06)
Optimism	0.00	0.03 (-0.01, 0.08)	0.04 (-0.02, 0.10)	0.09 (0.03, 0.15)**
Purpose in life	0.00	0.09 (0.05, 0.12)***	0.11 (0.06, 0.16)***	0.13 (0.08, 0.18)***
Mastery	0.00	0.05 (0.01, 0.09)*	0.08 (0.01, 0.16)*	0.10 (0.03, 0.16)**
Health mastery	0.00	0.04 (-0.01, 0.09)	0.06 (0.00, 0.12)*	0.07 (0.00, 0.15)*
Financial mastery	0.00	0.04 (-0.01, 0.09)	0.04 (-0.03, 0.11)	0.05 (-0.01, 0.12)

Outcomes	Informal Helping			
	0 hours (n=6,070) (Reference)	1-49 hours (n=3,603) RR/OR/β (95% CI)	50-99 hours (n=1,651) RR/OR/β (95% CI)	≥100 hours (n=1,674) RR/OR/β (95% CI)
Psychological Distress				
Depression	1.00	0.95 (0.82, 1.09)	1.04 (0.87, 1.25)	0.95 (0.78, 1.17)
Depressive symptoms	0.00	-0.04 (-0.08, 0.00)*	-0.04 (-0.09, 0.01)	-0.01 (-0.07, 0.04)
Hopelessness	0.00	-0.06 (-0.10, -0.01)**	-0.06 (-0.12, 0.00)*	-0.06 (-0.12, 0.00)*
Negative affect	0.00	0.02 (-0.03, 0.07)	0.03 (-0.03, 0.09)	0.06 (0.00, 0.12)*
Perceived constraints	0.00	-0.05 (-0.10, 0.00)*	-0.06 (-0.11, 0.00)	-0.05 (-0.12, 0.01)
Social Factors				
Loneliness	0.00	-0.02 (-0.06, 0.02)	-0.08 (-0.14, -0.03)**	-0.03 (-0.10, 0.03)
Living with a spouse/partner	1.00	1.00 (0.94, 1.07)	1.02 (0.94, 1.10)	0.99 (0.92, 1.07)
Contact children ≥1x/week	1.00	1.02 (0.96, 1.08)	1.02 (0.94, 1.10)	1.05 (0.98, 1.13)
Contact other family ≥1x/week	1.00	1.03 (0.97, 1.11)	1.01 (0.93, 1.11)	1.04 (0.95, 1.14)
Contact friends ≥1x/week	1.00	1.10 (1.02, 1.18)*	1.15 (1.06, 1.25)**	1.13 (1.04, 1.24)**

Note. Abbreviations: CI, confidence interval; OR, odds ratio; RR, risk ratio.

[a]If the reference value is "1," the effect estimate is OR or RR; if the reference value is "0," the effect estimate is β.

[b]The analytic sample was restricted to those who had participated in the baseline wave (t_1;2010 or 2012). Multiple imputation was performed to impute missing data on the exposure, covariates, and outcomes. All models adjusted for pre-baseline sociodemographic characteristics (age, sex, race/ethnicity, marital status, annual household income, total wealth, level of education, employment status, health insurance, geographic region), pre-baseline childhood abuse, pre-baseline religious service attendance, pre-baseline values of the outcome variables (diabetes, hypertension, stroke, cancer, heart disease, lung disease, arthritis, overweight/obesity, physical functioning limitations, cognitive impairment, chronic pain, self-rated health, heavy drinking, current smoking status, physical activity, sleep problems, positive affect, life satisfaction, optimism, purpose in life, mastery, health mastery, financial mastery, depressive symptoms, hopelessness, negative affect, perceived constraints, loneliness, living with a spouse/partner, contact children ≥1x/week, contact other family ≥1x/week, contact friends ≥1x/week), personality factors (openness, conscientiousness, extraversion, agreeableness, neuroticism), and the pre-baseline value of the exposure (informal helping). These variables were adjusted for in the wave pre-baseline to the exposure assessment (t_0;2006 or 2008).

[c]An outcome-wide analytic approach was used, and a separate model for each outcome was run. A different type of model was run depending on the nature of the outcome: 1) for each binary outcome with a prevalence of ≥10%, a generalized linear model (with a log link and Poisson distribution) was used to estimate a RR; 2) for each binary outcome with a prevalence of <10%, a logistic regression model was used to estimate an OR; and 3) for each continuous outcome, a linear regression model was used to estimate a β.

[d]All continuous outcomes were standardized (mean=0; standard deviation=1), and β was the standardized effect size.

*p<0.05 before Bonferroni correction; **p<0.01 before Bonferroni correction; ***p<0.05 after Bonferroni correction (the p-value cutoff for Bonferroni correction is p=0.05/35 outcomes=p<0.001).

Table 3. Robustness to Unmeasured Confounding (E-Values) for the Associations Between Informal Helping (≥100 hours vs. 0 hours) and Subsequent Health and Well-Being (N=12,998)[a]

	Effect Estimate[b]	Confidence Interval Limit[c]
Physical Health		
All-cause mortality	2.30	1.60
Number of chronic conditions	1.10	1.00
Diabetes	1.07	1.00
Hypertension	1.14	1.00
Stroke	1.86	1.23
Cancer	1.41	1.00
Heart disease	1.26	1.00
Lung disease	1.37	1.00
Arthritis	1.20	1.00
Overweight/obesity	1.23	1.00
Physical functioning limitations	1.82	1.43
Cognitive impairment	1.74	1.22
Chronic pain	1.14	1.00
Self-rated health	1.41	1.24
Health Behaviors		
Heavy drinking	1.69	1.00
Smoking	1.59	1.00
Frequent physical activity	1.47	1.23
Sleep problems	1.16	1.00
Psychological Well-being		
Positive affect	1.40	1.18
Life satisfaction	1.08	1.00
Optimism	1.40	1.22
Purpose in life	1.49	1.36
Mastery	1.41	1.22
Health mastery	1.34	1.07
Financial mastery	1.28	1.00
Psychological Distress		
Depression	1.27	1.00
Depressive symptoms	1.12	1.00
Hopelessness	1.31	1.08
Negative Affect	1.31	1.07
Constraints	1.27	1.00
Social Factors		
Loneliness	1.21	1.00
Living with a spouse/partner	1.08	1.00
Contact children ≥1x/week	1.27	1.00
Contact other family ≥1x/week	1.23	1.00
Contact friends ≥1x/week	1.52	1.23

[a]See VanderWeele and Ding (2017) for the formula for calculating E-values.

[b]The E-values for effect estimates are the minimum strength of association on the risk ratio scale that an unmeasured confounder would need to have with both the exposure and the outcome to fully explain away the observed association between the exposure and outcome, conditional on the measured covariates.
[c]The E-values for the limit of the 95% confidence interval (CI) closest to the null denote the minimum strength of association on the risk ratio scale that an unmeasured confounder would need to have with both the exposure and the outcome to shift the confidence interval to include the null value, conditional on the measured covariates.

References

Abbott, R., Ploubidis, G., Huppert, F., Kuh, D., Wadsworth, M., & Croudace, T. (2006). Psychometric evaluation and predictive validity of Ryff's psychological well-being items in a UK birth cohort sample of women. *Health and Quality of Life Outcomes*, *4*, 76. https://doi.org/10.1186/1477-7525-4-76

Aldwin, C. M., & Igarashi, H. (2015). Successful, optimal, and resilient aging: A psychosocial perspective. In P. A. Lichtenberg, B. T. Mast, B. D. Carpenter & W. J. Loebach (Eds.), *APA handbook of clinical geropsychology: Vol. 1. History and status of the field and perspectives on aging* (pp. 331–359). American Psychological Association. https://doi.org/10.1037/14458-014

Appau, S., & Awaworyi Churchill, S. (2019). Charity, volunteering type and subjective wellbeing. *VOLUNTAS: International Journal of Voluntary and Nonprofit Organizations*, *30*(5), 1118–1132. https://doi.org/10.1007/s11266-018-0009-8

Beck, A. T., Weissman, A., Lester, D., & Trexler, L. (1974). The measurement of pessimism: The Hopelessness Scale. *Journal of Consulting and Clinical Psychology*, *42*, 861–865. https://doi.org/10.1037/h0037562

Brim, O. G., & Featherman, D. L. (1998). Surveying midlife development in the United States. *Unpublished Manuscript*.

Burr, J. A., Han, S., Lee, H. J., Tavares, J. L., & Mutchler, J. E. (2018). Health benefits associated with three helping behaviors: Evidence for incident cardiovascular disease. *The Journals of Gerontology: Series B, 73*(3), 492–500. https://doi.org/10.1093/geronb/gbx082

Burr, J. A., Mutchler, J. E., & Han, S. H. (2021). Chapter 19 - Volunteering and health in later life. In K. F. Ferraro & D. Carr (Eds.), *Handbook of aging and the social sciences (9th ed.,* pp. 303–319). Academic Press. https://doi.org/10.1016/B978-0-12-815970-5.00019-X

Crimmins, E. M., Kim, J. K., Langa, K. M., & Weir, D. R. (2011). Assessment of cognition using surveys and neuropsychological assessment: The Health and Retirement Study and the Aging, Demographics, and Memory Study. *The Journals of Gerontology. Series B, 66* (Suppl 1), 162-171. https://doi.org/10.1093/geronb/gbr048

Deci, E. L., & Ryan, R. M. (2000). The "what" and "why" of goal pursuits: Human needs and the self-determination of behavior. *Psychological Inquiry, 11*(4), 227–268. https://doi.org/10.1207/S15327965PLI1104_01

Department of Health and Human Services (n.d.). *Drinking levels defined.* National Institute on Alcohol Abuse and Alcoholism (NIAAA). Retrieved March 31, 2019, from https://www.niaaa.nih.gov/alcohol-health/overview-alcohol-consumption/moderate-binge-drinking

Depp, C. A., & Jeste, D. V. (2006). Definitions and predictors of successful aging: A comprehensive review of larger quantitative studies. *The American Journal of Geriatric Psychiatry, 14*(1), 6–20. https://doi.org/10.1097/01.JGP.0000192501.03069.bc

Diener, E., Emmons, R. A., Larsen, R. J., & Griffin, S. (1985). The satisfaction with life scale. *Journal of Personality Assessment*, *49*(1), 71–75. https://doi.org/10.1207/s15327752jpa4901_13

Dunn, O. J. (1961). Multiple comparisons among means. *Journal of the American Statistical Association, 56*(293), 52–64. https://doi.org/10.1080/01621459.1961.10482090

Einolf, C. J., Prouteau, L., Nezhina, T., & Ibrayeva, A. R. (2016). Informal, unorganized volunteering. In D. H. Smith, R. A. Stebbins, & J. Grotz (Eds.), *The Palgrave handbook of volunteering, civic participation, and nonprofit associations* (pp. 223–241). Palgrave Macmillan UK. https://doi.org/10.1007/978-1-137-26317-9_10

Erickson, R. J., & Stacey, C. L. (2013). Attending to mind and body: Engaging the complexity of emotion practice among caring professionals. In A. A. Grandey, J. M. Diefendorff & D. E. Rupp (Eds.), Emotional *labor in the 21st century: Diverse perspectives on emotional regulation at work (pp. 175-196)*. Routledge. https://doi.org/10.4324/9780203100851

Everson, S. A., Kaplan, G. A., Goldberg, D. E., Salonen, R., & Salonen, J. T. (1997). Hopelessness and 4-year progression of carotid atherosclerosis: The Kuopio Ischemic Heart Disease Risk Factor Study. *Arteriosclerosis, Thrombosis, and Vascular Biology*, *17* (8), 1490–1495. https://doi.org/10.1161/01.atv.17.8.1490

Fisher, G. G., Faul, J. D., Weir, D. R., & Wallace, R. B. (2005). *Documentation of chronic disease measures in the Health and Retirement Study (HRS/AHEAD)*. Institute for Social Research, University of Michigan. https://hrs.isr.umich.edu/sites/default/files/biblio/dr-009.pdf

Fisher, G. G., Halimah, H., Faul, J. D., Rogers, W. L., & Weir, D. R. (2017). *Health and Retirement Study imputation of cognitive functioning measures: 1992 – 2014*. Institute for Social Research, University of Michigan. https://hrs.isr.umich.edu/sites/default/files/biblio/COGIMPdd.pdf

Gottlieb, B. H. (1978). The development and application of a classification scheme of informal helping behaviours. *Canadian Journal of Behavioural Science/Revue Canadienne Des Sciences Du Comportement*, *10*(2), 105–115. https://doi.org/10.1037/h0081539

Han, S. H., Tavares, J. L., Evans, M., Saczynski, J., & Burr, J. A. (2017). Social activities, incident cardiovascular disease, and mortality: Health behaviors edition. *Journal of Aging and Health*, *29*(2), 268–288. https://doi.org/10.1177/0898264316635565

Hanson, J. A., & VanderWeele, T. J. (2021). The comprehensive measure of meaning: Psychological and philosophical foundations. In M. Lee, L. D. Kubzansky & T. J. Vanderweele (Eds.). *Measuring well-being: Interdisciplinary perspectives from the social sciences and the humanities* (pp. 339-376). Oxford University Press. https://doi.org/10.1093/oso/9780197512531.003.0013

Harel, O., Mitchell, E. M., Perkins, N. J., Cole, S. R., Tchetgen Tchetgen, E. J., Sun, B., & Schisterman, E. F. (2018). Multiple imputation for incomplete data in epidemiologic studies. *American Journal of Epidemiology*, *187*(3), 576–584. https://doi.org/10.1093/aje/kwx349

Holt-Lunstad, J., Smith, T. B., & Layton, J. B. (2010). Social relationships and mortality risk: A meta-analytic review. *PLoS Medicine*, *7*(7), e1000316. https://doi.org/10.1371/journal.pmed.1000316

House, J. S., Umberson, D., & Landis, K. R. (1988). Structures and processes of social support. *Annual Review of Sociology*, *14*(1), 293–318. https://doi.org/10.1146/annurev.so.14.080188.001453

Hui, B. P. H., Ng, J. C. K., Berzaghi, E., Cunningham-Amos, L. A., & Kogan, A. (2020). Rewards of kindness? A meta-analysis of the link between prosociality and well-being. *Psychological Bulletin*, *146*(12), 1084–1116. https://doi.org/10.1037/bul0000298

Jenkins, C. D., Stanton, B. -A., Niemcryk, S. J., & Rose, R. M. (1988). A scale for the estimation of sleep problems in clinical research. *Journal of Clinical Epidemiology*, *41*(4), 313–321. https://doi.org/10.1016/0895-4356(88)90138-2

Jenkins, K. R., Ofstedal, M. B., & Weir, D. (2008). *Documentation of health behaviors and risk factors measured in the Health and Retirement Study (HRS/AHEAD)*. Institute for Social Research, University of Michigan. https://hrs.isr.umich.edu/publications/biblio/5736

Johnson, S. S., & Post, S. G. (2017). Rx it's good to be good (G2BG) 2017 commentary: Prescribing volunteerism for health, happiness, resilience, and longevity. *American Journal of Health Promotion*, *31*(2), 163–172. https://doi.org/10.1177/0890117117691705

Kahana, E., Bhatta, T., Lovegreen, L. D., Kahana, B., & Midlarsky, E. (2013). Altruism, helping, and volunteering: Pathways to well-being in late life. *Journal of Aging and Health*, *25*(1), 159–187. https://doi.org/10.1177/0898264312469665

Katz S., Ford, A. B., Moskowitz, R. W., Jackson, B. A., & Jaffe, M. W. (1963). Studies of illness in the aged: The index of ADL: A standardized measure of biological and psychosocial function. *JAMA*, *185*(12), 914–919. https://doi.org/10.1001/jama.1963.03060120024016

Kim, E. S., Chen, Y., Nakamura, J. S., Ryff, C. D., & VanderWeele, T. J. (2021). Sense of purpose in life and subsequent physical, behavioral, and psychosocial health: An outcome-wide approach. *American Journal of Health Promotion 36*(1), 137–147. https://doi.org/10.1177/08901171211038545

Kim, E. S., Whillans, A. V., Lee, M. T., Chen, Y., & VanderWeele, T. J. (2020). Volunteering and subsequent health and well-being in older adults: An outcome-wide longitudinal

approach. *American Journal of Preventive Medicine*, *29*(2), 176–186. https://doi.org/10.1016/j.amepre.2020.03.004

Krause, N. (2009). Church-based volunteering, providing informal support at church, and self-rated health in late life. *Journal of Aging and Health*, *21*(1), 63–84. https://doi.org/10.1177/0898264308328638

Krekel, C., De Neve, J. E., Fancourt, D., & Layard, R. (2021). A local community course that raises wellbeing and pro-sociality: Evidence from a randomised controlled trial. *Journal of Economic Behavior & Organization*, *188*, 322–336. https://doi.org/10.1016/j.jebo.2021.05.021

Kubzansky, L. D., Huffman, J. C., Boehm, J. K., Hernandez, R., Kim, E. S., Koga, H. K., Feig, E. H., Lloyd-Jones, D. M., Seligman, M. E. P., & Labarthe, D. R. (2018). Positive psychological well-being and cardiovascular disease. *Journal of the American College of Cardiology*, *72*(12), 1382–1396. https://doi.org/10.1016/j.jacc.2018.07.042

Lachman, M. E., & Weaver, S. L. (1997). *The Midlife Development Inventory (MIDI) personality scales: Scale construction and scoring [Technical Report]*. Brandeis University, Psychology Department. https://www.brandeis.edu/psychology/lachman/pdfs/midi-personality-scales.pdf

Lachman, M. E., & Weaver, S. L. (1998). The sense of control as a moderator of social class differences in health and well-being. *Journal of Personality and Social Psychology*, *74*(3), 763–773. https://doi.org/10.1037/0022-3514.74.3.763

Langa, K. M., Plassman, B. L., Wallace, R. B., Herzog, A. R., Heeringa, S. G., Ofstedal, M. B., Burke, J. R., Fisher, G. G., Fultz, N. H., & Hurd, M. D. (2005). The Aging,

Demographics, and Memory Study: Study design and methods. *Neuroepidemiology*, *25*(4), 181–191. https://doi.org/10.1159/000087448

Lawton, M.P., Brody E. M. (1969). Assessment of older people: Self-maintaining and instrumental activities of daily living. *Gerontologist*, *9*(3), 179–186. https://doi.org/10.1093/geront/9.3_Part_1.179

Lee, M. T., Bialowolski, P., Weziak-Bialowolska, D., Mooney, K. D., Lerner, P. J., McNeely, E., & VanderWeele, T. J. (2021). Self-assessed importance of domains of flourishing: Demographics and correlations with well-being. *The Journal of Positive Psychology*, *16*(1), 137–144. https://doi.org/10.1080/17439760.2020.1716050

Li, Y., & Ferraro, K. F. (2005). Volunteering and depression in later life: Social benefit or selection processes? *Journal of Health and Social Behavior*, *46*(1), 68–84. https://doi.org/10.1177/002214650504600106

Martinez, I. L., Crooks, D., Kim, K. S., & Tanner, E. (2011). Invisible civic engagement among older adults: Valuing the contributions of informal volunteering. *Journal of Cross-Cultural Gerontology*, *26*(1), 23–37. https://doi.org/10.1007/s10823-011-9137-y

Matthews, K., & Nazroo, J. (2021). The impact of volunteering and its characteristics on well-being after state pension age: Longitudinal evidence from the English Longitudinal Study of Ageing. *The Journals of Gerontology: Series B*, *76*(3), 632–641. https://doi.org/10.1093/geronb/gbaa146

McDonnall, M. C. (2011). The effect of productive activities on expressive symptoms among older adults with dual sensory loss. *Research on Aging*, *33*(3), 234–255. https://doi.org/10.1177/0164027511399106

McGinnis, J. M., Williams-Russo, P., & Knickman, J. R. (2002). The case for more active policy attention to health promotion. *Health Affairs*, *21*(2), 78–93. https://doi.org/10.1377/hlthaff.21.2.78

McIntosh, B. R., & Danigelis, N. L. (1995). Race, gender, and the relevance of productive activity for elders' affect. *The Journals of Gerontology: Series B*, *50B*(4), S229–S239. https://doi.org/10.1093/geronb/50B.4.S229

Mroczek, D. K., & Kolarz, C. M. (1998). The effect of age on positive and negative affect: A developmental perspective on happiness. *Journal of Personality and Social Psychology*, *75*(5), 1333–1349. https://doi.org/10.1037/0022-3514.75.5.1333

Nagi, S. Z. (1976). An epidemiology of disability among adults in the United States. *The Milbank Memorial Fund Quarterly. Health and Society*, *54*(4), 439–467. https://doi.org/10.2307/3349677

Nandi, A., Glymour, M. M., & Subramanian, S. V. (2014). Association among socioeconomic status, health behaviors, and all-cause mortality in the United States. *Epidemiology*, *25*(2), 170–177. https://doi.org/10.1097/EDE.0000000000000038

Ofstedal, M. B., Fisher, G. G., & Herzog, A. R. (2005). *Documentation of cognitive functioning measures in the Health and Retirement Study*. Institute for Social Research, University of Michigan. https://hrs.isr.umich.edu/sites/default/files/biblio/dr-006.pdf

Penner, L. A., Dovidio, J. F., Piliavin, J. A., & Schroeder, D. A. (2005). Prosocial behavior: Multilevel perspectives. *Annual Review of Psychology*, *56*(1), 365–392. https://doi.org/10.1146/annurev.psych.56.091103.070141

Plagnol, A. C., & Huppert, F. A. (2010). Happy to help? Exploring the factors associated with variations in rates of volunteering across Europe. *Social Indicators Research*, *97*(2), 157–176. https://doi.org/10.1007/s11205-009-9494-x

Poulin, M. J., Brown, S. L., Dillard, A. J., & Smith, D. M. (2013). Giving to others and the association between stress and mortality. *American Journal of Public Health*, *103*(9), 1649–1655. https://doi.org/10.2105/AJPH.2012.300876

Qu, H., Konrath, S., & Poulin, M. (2020). Which types of giving are associated with reduced mortality risk among older adults? *Personality and Individual Differences*, *154*, 109668. https://doi.org/10.1016/j.paid.2019.109668

Radloff, L. S. (1977). The CES-D Scale: A self-report depression scale for research in the general population. *Applied Psychological Measurement*, *1*(3), 385–401. https://doi.org/10.1177/014662167700100306

Reich, J. W., Zautra, A. J., & Hall, J. S. (Eds.). (2010). *Handbook of adult resilience.* (pp. xix-540). The Guilford Press.

Ritchie, H., & Roser, M. (2018). *Causes of death*. Our world in data. https://ourworldindata.org/causes-of-death

Rosow, I., & Breslau, N. (1966). A Guttman health scale for the aged. *Journal of Gerontology*, *21*(4), 556–559. https://doi.org/10.1093/geronj/21.4.556

Rowe, J., & Kahn, R. (1987). Human aging: Usual and successful. *Science*, *237*(4811), 143. https://doi.org/10.1126/science.3299702

Russell, D. W. (1996). UCLA Loneliness Scale (Version 3): Reliability, validity, and factor structure. *Journal of Personality Assessment*, *66*(1), 20–40. https://doi.org/10.1207/s15327752jpa6601_2

Ryff, C. D., & Keyes, C. L. M. (1995). The structure of psychological well-being revisited. *Journal of Personality and Social Psychology*, *69*(4), 719–727. https://doi.org/10.1037/0022-3514.69.4.719

Ryff, C. D., & Singer, B. (2009). Understanding healthy aging: Key components and their integration. In M. Silverstein, V. L. Bengtson, N. Putney & D. Gans (Eds.), *Handbook of theories of aging, second edition.* (pp. 117–144). Springer Publishing Company.

Scheier, M. F., Carver, C. S., & Bridges, M. W. (1994). Distinguishing optimism from neuroticism (and trait anxiety, self-mastery, and self-esteem): A reevaluation of the Life Orientation Test. *Journal of Personality and Social Psychology*, *67*(6), 1063–1078. https://doi.org/10.1037/0022-3514.67.6.1063

Smith, J., Ryan, L., Fisher, G. G., Sonnega, A., & Weir, D. (2017). *Documentation report core section LB.* 72. Institute for Social Research, University of Michigan. https://hrs.isr.umich.edu/publications/biblio/9066

Steffeck, D. (2000). *Documentation of affective functioning measures in the Health and Retirement Study.* Institute for Social Research, University of Michigan. https://hrs.isr.umich.edu/sites/default/files/biblio/dr-005.pdf

Teo, A. R., Choi, H., Andrea, S. B., Valenstein, M., Newsom, J. T., Dobscha, S. K., & Zivin, K. (2015). Does mode of contact with different types of social relationships predict depression in older adults? Evidence from a nationally representative survey. *Journal of the American Geriatrics Society*, *63*(10), 2014–2022. https://doi.org/10.1111/jgs.13667

Umberson, D., & Montez, J. K. (2010). Social relationships and health: A flashpoint for health policy. *Journal of Health and Social Behavior*, *51*(Suppl), S54–S66. https://doi.org/10.1177/0022146510383501

United Nations, Department of Economic and Social Affairs, Population Division. (2020). In *World Population Ageing 2019*. United Nations.

VanderWeele, T. J. (2017). On the promotion of human flourishing. *Proceedings of the National Academy of Sciences, 114*(31), 8148–8156. https://doi.org/10.1073/pnas.1702996114

VanderWeele, T. J., & Ding, P. (2017). Sensitivity analysis in observational research: Introducing the E-value. *Annals of Internal Medicine*, *167*(4), 268–274. https://doi.org/10.7326/M16-2607

VanderWeele, T. J., & Mathur, M. B. (2019). Some desirable properties of the Bonferroni correction: Is the Bonferroni correction really so bad? *American Journal of Epidemiology*, *188*(3), 617–618. https://doi.org/10.1093/aje/kwy250

VanderWeele, T. J., Mathur, M. B., & Chen, Y. (2020). Outcome-wide longitudinal designs for causal inference: A new template for empirical studies. *Statistical Science*, *35*(3), 437–466. https://doi.org/10.1214/19-STS728

Wahrendorf, M., Ribet, C., Zins, M., & Siegrist, J. (2008). Social productivity and depressive symptoms in early old age–results from the GAZEL study. *Aging & Mental Health*, *12*(3), 310–316. https://doi.org/10.1080/13607860802120805

Watson, D., & Clark, L. A. (1994). *The PANAS-X: Manual for the positive and negative affect schedule-expanded form*. The University of Iowa. https://www2.psychology.uiowa.edu/faculty/clark/panas-x.pdf

Watson, D., Clark, L. A., & Tellegen, A. (1988). Development and validation of brief measures of positive and negative affect: The PANAS scales. *Journal of Personality and Social Psychology*, *54*(6), 1063–1070. https://doi.org/10.1037/0022-3514.54.6.1063

Weir, D. R. (2016). *Validating mortality ascertainment in the health and retirement study*. Institute for Social Research, University of Michigan. https://hrs.isr.umich.edu/sites/default/files/biblio/Weir_mortality_ascertainment.pdf

World Health Organization. (1995). *Physical status: The use and interpretation of anthropometry: Report of a WHO expert committee* [WHO Technical Report Series 854]. https://apps.who.int/iris/bitstream/handle/10665/37003/WHO_TRS_854.pdf

Appendix A.

Assessment of Outcomes

Reference Group

Unless otherwise noted, for all binary outcomes, the reference group was the healthiest group.

Physical Health

All-cause mortality. Information about death was obtained up to the 2016 questionnaire wave via two methods. First, an exit interview was conducted with next-of-kin. Then, after each wave of data collection, the National Death Index (NDI) was searched for death information. When comparing deaths reported by NDI versus exit interviews, there is a 95.5% match (Weir, 2016).

Chronic conditions. Participants self-reported whether they were ever told by a healthcare provider that they had (yes/no) the following conditions: 1) diabetes, 2) hypertension, 3) stroke, 4) cancer, 5) heart disease, 6) lung disease, or 7) arthritis. Validity and reliability of self-reported chronic conditions has previously been demonstrated in HRS (Fisher et al., 2005).

Overweight/obesity. Body mass index (BMI) was derived based on self-reported height and weight, and BMI was calculated as weight/height^2 (kg/m^2). A BMI of $\geq$25 kg/m^2 was considered as overweight/obese (*World Health Organization*, 1995).

Number of chronic conditions. To create a score for the number of chronic conditions, a summary score was calculated by summing the number of reported conditions (e.g., the 7 chronic conditions and also overweight/obesity; range=0 to 8).

Cognitive function problem. The HRS cognitive function assessment (Fisher et al., 2017; Ofstedal et al., 2005) was adapted from the modified Telephone Interview for Cognitive Status (TICS-M). The assessment is a 27-point scale that included an immediate and delayed 10-noun free recall test, a serial 7 subtraction test, and a backward count 20 test. This assessment tool has been shown to have high sensitivity and specificity for cognitive impairment in older adults; cut points were derived from previous research conducted on cognitive impairment in HRS (Crimmins et al., 2011; Langa et al., 2005). Respondents scoring 0-11 on the 27-point scale were classified as having "cognitive impairment," while those scoring ≥12 were classified as "normal" (the reference group). More detailed information about the cognitive assessments can be found in HRS reports (Fisher et al., 2017; Ofstedal et al., 2005).

Physical functioning limitations. Physical functioning limitations were assessed using items adapted from scales developed by Rosow and Breslau (1966), Nagi (1976), Katz, Ford, Moskowitz, Jackson, and Jaffe (1963), and Lawton and Brody (1969) (Katz et al., 1963; Lawton & Brody, 1969; Nagi, 1976; Rosow & Breslau, 1966). Participants were defined as having physical function limitations if they reported >4 limitations with physical functioning (i.e., walking several blocks, climbing one flight of stairs, pushing or pulling large objects, lifting or carrying 10 pounds, getting up from a chair, reaching or extending arms up, stooping, kneeling, or crouching, sitting for 2 hours) or activities of daily living (i.e., walking across a room,

dressing, eating, bathing, getting in/out bed, using the toilet, picking up a dime). Those reporting ≤4 limitations were considered "normal" in the physical function domain and served as the reference group. This criterion was determined by identifying the physical function score where 75% of participants could be considered as having healthy physical function at baseline.

Chronic pain. Chronic pain was assessed by asking respondents (yes/no): "Are you often troubled with pain?" No pain was the reference group.

Self-rated health. Participants were asked, "Would you say your health is excellent, very good, good, fair, or poor?" on a 5-point scale (reverse coded with higher scores indicating higher self-rated health).

Health Behaviors

Heavy drinking. Following the National Institute on Alcohol Abuse and Alcoholism guidelines (Department of Health and Human Services, n.d.), heavy drinking was defined as >14 for drinks/week for men and >7 drinks/week for women. Alcohol consumption was measured by multiplying the number of days/week that alcohol was consumed by the number of drinks/day, which resulted in the number of drinks/week. Participants not in this alcohol consumption range were classified as non-heavy drinkers (the reference group).

Smoking. Current smoking status was assessed by asking participants: "Do you smoke cigarettes now?" The response categories included "yes" or "no" (with "no" smoking as the reference group).

Frequent physical activity. Based on prior research, we created a binary physical activity variable where ≥1x/week of vigorous or moderate exercise was considered frequent physical activity and <1x/week of vigorous or moderate exercise was the reference group (Nandi et al., 2014). Physical activity was measured by asking participants their frequency of engaging in vigorous (e.g., running, swimming, aerobics), moderate (e.g., gardening, dancing, walking at a moderate pace), and light (e.g., vacuuming, laundry) activities over the past 12 months. Response categories included daily, >1x/week, 1x/week, 1-3x/month, and hardly ever or never.

Sleep problems. Participants completed the 4-item Jenkins Sleep Questionnaire, a validated and widely used screening instrument for sleep complaints, querying insomnia symptoms (Jenkins et al., 1988). Response categories included "most of the time," "sometimes," and "rarely or never." Having sleep problems was defined as reporting: "most of the time" for any of the three negatively worded items (e.g., "How often do you have trouble falling asleep?") and "rarely or never" to the one positively worded item (i.e., "feel really rested when you wake up in the morning"). Participants were considered unhealthy (i.e., having sleep problems) if they reported one or more sleep problems. The sleep questionnaire was only administered every other wave. Thus, sleep data was imputed for half of the sample. Imputed and complete-case analyses showed similar estimates.

Psychological Well-Being

Positive affect. Positive affect was measured (in 2006 only) with a 6-item scale originally developed for use in the Midlife in the United States Study (Brim & Featherman, 1998; Mroczek

& Kolarz, 1998; Watson et al., 1988). The scale assessed how often the participant felt "cheerful," "in good spirits," "extremely happy," "calm and peaceful," "satisfied," and "full of life" over the past 30 days. Response categories ranged from 1 (all of the time) to 5 (none of the time). Responses were reverse scored, so that a higher score indicated higher positive affect. An overall score was derived by averaging responses across all 6 items (α=0.91 in 2006, range=1 to 5). After the 2006 wave, the HRS switched to a more expansive measure of positive affect based on the Positive and Negative Affect Schedule (PANAS-X; Watson & Clark, 1994). It included the following 13 items: determined, enthusiastic, active, proud, interested, happy, attentive, content, inspired, hopeful, alert, calm, excited. An overall score was derived by averaging responses across all 13 items (α=0.92, range=1 to 5). A limitation of this study is that affect was measured in a different way during only the first wave of the study. However, scores were standardized and both the prior and current measures of affect operate very similarly (e.g., similar correlations with other variables, similar distributions, etc.).

Life satisfaction. Life satisfaction was assessed with the 5-item Satisfaction with Life Scale (e.g., "In most ways my life is close to ideal"; Diener et al., 1985). The scale has shown excellent psychometric properties in prior work. Response categories ranged from 1 (strongly disagree) to 7 (strongly agree). An overall score was derived by averaging responses across all 5 items, with a higher score indicating higher life satisfaction (α=0.88, range=1 to 7).

Optimism. Optimism was assessed using the Life Orientation Test-Revised (LOT-R). The measure has good discriminant and convergent validity, and good reliability (Scheier et al., 1994). Using a 6-point Likert scale (from 1 (strongly disagree) to 6 (strongly agree)), participants

were asked the degree to which they agreed with six statements such as, “In uncertain times, I usually expect the best.” After reverse coding negatively worded items, all items were averaged together to create a composite score, with higher scores indicating higher optimism (α=0.75, range=1 to 6).

Purpose in life. Purpose in life was assessed with a 7-item purpose in life subscale from the Ryff’s Psychological Well-Being Scale (Ryff & Keyes, 1995). The 7-item subscale has been validated in prior work and has shown good psychometric properties (Abbott et al., 2006). Using a 6-point Likert scale (from 1 (strongly disagree) to 6 (strongly agree)), participants were asked the degree to which they agreed with statements such as, “I have a sense of direction and purpose in my life.” Negatively worded items were reverse coded and all items were averaged to create a composite score, with a higher score indicating higher purpose (α=0.76, range=1 to 6).

Mastery. Mastery was measured with 5-items derived from Lachman and Weaver (1998) and rated on a scale from 1 (strongly disagree) to 6 (strongly agree). The measure has good discriminant and convergent validity, as well as good reliability (Lachman & Weaver, 1998). Participants were asked the degree to which they agreed with five statements such as, “I can do just about anything I really set my mind to.” All items were averaged together to create a composite score, with higher scores indicating higher mastery (α=0.90, range=1 to 6).

Health mastery. On a 0 to 10 scale where 0 means “no control at all” and 10 means “very much control,” participants were asked, “how would you rate the amount of control you have over your health these days?”

Financial mastery. On a 0 to 10 scale where 0 means "no control at all" and 10 means "very much control," participants were asked, "how would you rate the amount of control you have over your financial situation these days?"

Psychological Distress

Depressive symptoms and depression. Depressive symptoms over the past week were measured using the 8-item Center for Epidemiologic Studies Depression Scale (CESD; Radloff, 1977) (e.g., "Much of the time during the past week, I felt depressed"), and response options included "yes" or "no" for each item. Following HRS protocol, an overall score was derived ranging from 0 to 8, with a higher score indicating higher depressive symptoms. The scale has been previously validated in the Health and Retirement Study (Steffeck, 2000) and showed high reliability in this sample (α=0.80). Following prior work (Steffeck, 2000), participants with a score of ≥4 were considered as having significant depressive symptoms, or depression. Prior work suggested that the cutoff of 4 would produce comparable results as the 16 symptoms cutoff when using the full 20-item CESD scale (Steffeck, 2000). No depression was the reference group.

Hopelessness. Hopelessness was measured with 4 questionnaire items from two previously validated scales (Beck et al., 1974; Everson et al., 1997) (e.g., "I feel it is impossible for me to reach the goals that I would like to strive for", "The future seems hopeless to me and I can't believe that things are changing for the better"). Response categories ranged from 1

(strongly disagree) to 6 (strongly agree). An overall score was created by averaging the responses across all items (α=0.87, range=1 to 6).

Negative affect. Negative affect was measured (in 2006 only) with a 6-item scale originally developed for use in the Midlife in the United States Study (Brim & Featherman, 1998; Mroczek & Kolarz, 1998; Watson et al., 1988). The scale assessed how often the participant felt "so depressed that nothing could cheer you up," "hopeless," "restless or fidgety," "that everything was an effort," "worthless," and "nervous" over the past 30 days. Response categories ranged from 1 (all of the time) to 5 (none of the time). Responses were reverse scored, so that a higher score indicated higher negative affect. An overall score was derived by averaging responses across all 6 items (α=0.86 in 2006, range=1 to 5). After the 2006 wave, the HRS switched to a more expansive measure of negative affect based on the Positive and Negative Affect Schedule (PANAS-X; Watson & Clark, 1994). It included the following 12 items: afraid, upset, guilty, scared, frustrated, bored, hostile, jittery, ashamed, nervous, sad, distressed. An overall score was derived by averaging responses across all 12 items (α=0.89, range=1 to 5). A limitation of this study is that affect was measured in a different way during only the first wave of the study. However, scores were standardized and both the prior and current measures of affect operate very similarly (e.g., similar correlations with other variables, similar distributions, etc.).

Perceived constraints. Perceived constraints were assessed with 5 items derived from Lachman and Weaver (1998), and this measure has good discriminant and convergent validity, as well as good reliability (Lachman & Weaver, 1998). Using a 6-point Likert scale (from 1

(strongly disagree) to 6 (strongly agree)), participants were asked the degree to which they agreed with statements such as, "What happens in my life is often beyond my control." All items were averaged to create an overall score, with higher scores indicating a higher sense of constraints on personal control (α=0.86, range=1 to 6).

Social Factors

Loneliness. Loneliness was measured with three items from the previously validated UCLA Loneliness Scale (i.e., How much of the time do you feel: 1) you lack companionship, 2) left out, and 3) isolated from others; Russell, 1996). Response categories ranged from 1 (often) to 3 (hardly ever or never). Responses were reverse scored, so that a higher score indicated higher loneliness. An overall score was derived by averaging the responses across the three items (α=0.80, range=1 to 3).

Living with partner/spouse. Participants were asked (yes/no), "Do you have a husband, wife, or partner with whom you live?" The reference group was not living with a spouse or partner.

Frequency of contact with: children, other family, and friends. Frequency of contact with children, other family, or friends was each queried separately, but in the same way. For example, participants were asked: "On average, how often do you do each of the following?" 1) "Meet up (include both arranged and chance meetings)," 2) "Speak on the phone," 3) "Write or email." For each of these 3 categories of questions, HRS respondents had the option of choosing 1 of the following 6 responses: 1) ≥3x/week, 2) 1x-2x/week, 3) 1-2x/month, 4) every few months, 5) 1-

2x/year, 6) <1x/year or never (Teo et al., 2015). Because contact of any kind (regardless of medium) was the main point of intertest, the highest value on any of the three modes of contact (e.g., meet up, phone, write/email) was taken. In other words, if the respondent did not meet in person very often with the other person but spoke on the phone very often with that person, contact was operationalized as fairly common, given that they speak on the phone very often. A binary frequency of contact variable was created where <1x/week of contact was considered infrequent contact (with this serving as the reference group) and ≥1x/week contact was considered frequent contact.

Other Factors

Personality. The "Big-5" personality traits (openness to experience, conscientiousness, extraversion, agreeableness, and neuroticism; Lachman & Weaver, 1997) were measured using 26 items derived from the Midlife Development Inventory Personality scales (MIDI) and International Personality Item Pool (IPIP). Using existing trait inventories, the goal of MIDI was to create the shortest possible collection of items that measured the Big-Five personality traits with high validity and reliability. In a pilot study conducted among a probability sample of 1,000 adults aged 30-70, items with the highest item-to-total correlations and factor loadings were selected for the MIDI. Forward regressions were then computed to determine the smallest number of items needed to account for more than 90 percent of the total scale variance. As an illustrative example, items on the conscientiousness scale included "organized," "responsible," "hardworking," and "careless." Response categories ranged from 1 (a lot) to 4 (not at all). Responses were reverse scored, so that a higher score indicated higher indication of a given

personality trait. An overall score for each personality trait was derived by averaging responses across all items of a given Big-5 Personality variable.

Formal volunteering. Volunteering hours was assessed by asking HRS participants: "Have you spent any time in the past 12 months doing volunteer work for religious, educational, health-related or other charitable organizations?" "If they responded yes, then HRS asked how many hours they volunteered: 1-49 hours, 50-99 hours, 100-199 hours, or ≥200 hours.

Caregiving. Caregiving was assessed by asking HRS participants: "Please tell us how often you do each activity. Care for a sick or disabled adult?" Response options included: daily, several times a week, once a week, several times a month, at least once a month, not in the last month, and never/not relevant.

Appendix B.

Proof Illustrating How Adjusting for Pre-Baseline Informal Helping Can Help Us Evaluate How "Changes" in Informal Helping are Associated with Subsequent Health and Well-Being Outcomes Over Time

Let Y be the outcome in 2014/2016, A_1 the informal helping exposure in 2010/2012, A_0 the informal helping exposure in 2006/2008, C the set of covariates in 2006/2008. For a continuous outcome, the regression model is: $E[Y|a_0, a_1, c] = v + b_0a_0 + b_1a_1 + b_2'c$

Let Y_a denote the potential outcome Y for an individual under an intervention to set A_1 to a. For an individual with baseline informal helping exposure $A_0=a_0$ and covariates c in 2006/2008, under the no-confounding (and positivity and consistency) and modeling assumptions, a change in informal helping of d points $A_0=a_0$ to $A_1=a_0+d$ in 2010/2012, rather than maintaining informal helping of $A_1=a_0$ in 2010/2012, will give rise to an effect (a difference in potential outcomes for Y) of:

$E[Y_{a0+d}| A_0=a_0, c] - E[Y_{a0}| A_0=a_0, c]$

$= E[Y_{a0+d}| A_1=a_0+d, A_0=a_0, c] - E[Y_{a0}| A_1=a_0, A_0=a_0, c]$

$= E[Y| A_1=a_0+d, A_0=a_0, c] - E[Y| A_1=a_0, A_0=a_0, c]$

$= [v + b_0a_0 + b_1(a_0+d) + b_2'c] - [v + b_0a_0 + b_1a_0 + b_2'c]$

$= b_1d$

where the first equality follows by the no-confounding assumption, the second by consistency, and the third by the statistical model.

Appendix C.

Considering Causes of Death

We considered the idea of creating aggregate measures that combined the incidence of a condition and death due to that condition. However, out of the 14 ways HRS categorizes causes of death, very few categories cleanly mapped onto health conditions we evaluated in this study without a large risk of misclassification error. Thus, we did not pursue this option. Causes of death included deaths due to: 1) musculoskeletal system and connective tissue; 2) heart, circulatory and blood conditions; 3) allergies; hay fever; sinusitis; tonsillitis; 4) endocrine, metabolic and nutritional conditions; 5) digestive system (stomach, liver, gallbladder, kidney, bladder); 6) neurological and sensory conditions; 7) reproductive system and prostate conditions; 8) emotional and psychological conditions; 9) miscellaneous; 10) other symptoms; 11) not a health condition; 12) none; 13) other health condition; 14) cancers and tumors; skin conditions.

Appendix D.

Appendix Tables

Table A1. Changes in Informal Helping from the Pre-Baseline Wave (t_0) to the Baseline Wave (t_1)[a,b]

	Baseline Wave (t_1)			
Pre-Baseline Wave (t_0)	**0 hours** **%**	**1-49 hours** **%**	**50-99 hours** **%**	**≥100 hours** **%**
0 hours	68.7	19.5	6.3	5.5
1-49 hours	36.6	40.2	13.6	9.6
50-99 hours	26.7	33.2	20.3	19.9
≥100 hours	25.6	24.1	20.4	29.8

[a]The percent of people in informal helping groups: 0 hours, 1-49 hours, 50-99 hours, ≥100 hours in the pre-baseline wave (t_0; 2006/2008) who (four years later) end up in informal helping groups: 0 hours, 1-49 hours, 50-99 hours, ≥100 hours (t_1; 2010/2012).
[b]The values in the third and fourth rows do not add up to 100% because of rounding.

Table A2. Informal Helping and Subsequent Health and Well-being (After Adjustment for Conventional Covariates or All Covariates; Health and Retirement Study [HRS]: N=12,998)[a,b,c]

Outcomes	Informal Helping		
	0 hours (Reference)	Conventionally-Adjusted Models[d] ≥100 hours RR/OR/β (95% CI)	Fully-Adjusted Models[e] ≥100 hours RR/OR/β (95% CI)
Physical Health			
All-cause mortality	1.00	0.55 (0.44, 0.69)***	0.68 (0.54, 0.86)**
Number of chronic conditions	0.00	-0.08 (-0.13, -0.03)**	-0.01 (-0.05, 0.03)
Diabetes[f]	1.00	1.05 (0.79, 1.40)	1.02 (0.76, 1.37)
Hypertension[g]	1.00	1.20 (0.93, 1.54)	1.25 (0.97, 1.61)
Stroke[h]	1.00	0.60 (0.42, 0.87)**	0.68 (0.45, 1.01)
Cancer[i]	1.00	1.09 (0.82, 1.45)	1.09 (0.78, 1.53)
Heart disease[j]	1.00	0.93 (0.74, 1.16)	0.99 (0.77, 1.28)
Lung disease[k]	1.00	0.93 (0.66, 1.33)	1.00 (0.67, 1.50)
Arthritis[l]	1.00	1.19 (0.96, 1.47)	1.19 (0.94, 1.50)
Overweight/obesity[m]	1.00	1.08 (0.80, 1.46)	1.15 (0.83, 1.59)
Physical functioning limitations[n]	1.00	0.80 (0.66, 0.95)*	0.87 (0.72, 1.05)
Cognitive impairment[o]	1.00	0.80 (0.66, 0.98)*	0.91 (0.73, 1.12)
Chronic pain[p]	1.00	0.95 (0.82, 1.12)	1.01 (0.86, 1.19)
Self-rated health	0.00	0.25 (0.19, 0.31)***	0.10 (0.04, 0.15)***
Health Behaviors			
Heavy drinking	1.00	0.98 (0.77, 1.23)	1.20 (0.88, 1.64)
Smoking	1.00	1.00 (0.83, 1.20)	1.16 (0.95, 1.42)
Frequent physical activity	1.00	1.25 (1.17, 1.34)***	1.11 (1.04, 1.19)**
Sleep problems	1.00	0.95 (0.87, 1.05)	1.02 (0.91, 1.14)
Psychological Well-Being			
Positive affect	0.00	0.33 (0.26, 0.40)***	0.09 (0.02, 0.17)*
Life satisfaction	0.00	0.15 (0.10, 0.21)***	0.01 (-0.05, 0.06)
Optimism	0.00	0.30 (0.24, 0.35)***	0.09 (0.03, 0.15)**
Purpose in life	0.00	0.35 (0.29, 0.40)***	0.13 (0.08, 0.18)***
Mastery	0.00	0.23 (0.17, 0.30)***	0.10 (0.03, 0.16)**
Health mastery	0.00	0.16 (0.08, 0.25)***	0.07 (0.00, 0.15)*
Financial mastery	0.00	0.11 (0.04, 0.17)**	0.05 (-0.01, 0.12)

Outcomes	Informal Helping		
	0 hours (Reference)	Conventionally-Adjusted Models[d] ≥100 hours RR/OR/β (95% CI)	Fully-Adjusted Models[e] ≥100 hours RR/OR/β (95% CI)
Psychological Distress			
Depression	1.00	0.74 (0.61, 0.89)**	0.95 (0.78, 1.17)
Depressive symptoms	0.00	-0.16 (-0.22, -0.11)***	-0.01 (-0.07, 0.04)
Hopelessness	0.00	-0.25 (-0.30, -0.19)***	-0.06 (-0.12, 0.00)*
Negative affect	0.00	-0.06 (-0.11, 0.00)	0.06 (0.00, 0.12)*
Perceived constraints	0.00	-0.22 (-0.28, -0.16)***	-0.05 (-0.12, 0.01)
Social Factors			
Loneliness	0.00	-0.17 (-0.24, -0.11)***	-0.03 (-0.10, 0.03)
Living with a spouse/partner	1.00	1.02 (0.95, 1.10)	0.99 (0.92, 1.07)
Contact children ≥1x/week	1.00	1.12 (1.04, 1.20)**	1.05 (0.98, 1.13)
Contact other family ≥1x/week	1.00	1.15 (1.06, 1.24)***	1.04 (0.95, 1.14)
Contact friends ≥1x/week	1.00	1.27 (1.18, 1.38)***	1.13 (1.04, 1.24)**

Note. Abbreviations: CI, confidence interval; OR, odds ratio; RR, risk ratio.

[a]If the reference value is "1," the effect estimate is OR or RR; if the reference value is "0," the effect estimate is β.

[b]An outcome-wide analytic approach was used, and a separate model for each outcome was run. A different type of model was run depending on the nature of the outcome: 1) for each binary outcome with a prevalence of ≥10%, a generalized linear model (with a log link and Poisson distribution) was used to estimate a RR; 2) for each binary outcome with a prevalence of <10%, a logistic regression model was used to estimate an OR; and 3) for each continuous outcome, a linear regression model was used to estimate a β.

[c]All continuous outcomes were standardized (mean=0; standard deviation=1), and β was the standardized effect size.

[d]The analytic sample was restricted to those who had participated in the baseline wave (t_1;2010 or 2012). Multiple imputation was performed to impute missing data on the exposure, covariates, and outcomes. All models adjusted for pre-baseline sociodemographic characteristics (age, sex, race/ethnicity, marital status, annual household income, total wealth, level of education). These variables were adjusted for in the pre-baseline wave (t_0;2006 or 2008).

[e]The analytic sample was restricted to those who had participated in the baseline wave (t_1;2010 or 2012). Multiple imputation was performed to impute missing data on the exposure, covariates, and outcomes. All models adjusted for pre-baseline sociodemographic characteristics (age, sex, race/ethnicity, marital status, annual household income, total wealth, level of education, employment status, health insurance, geographic region), pre-baseline childhood abuse, pre-baseline religious service attendance, pre-baseline values of the outcome variables (diabetes, hypertension, stroke, cancer, heart disease, lung disease, arthritis, overweight/obesity, physical functioning limitations, cognitive impairment, chronic pain, self-rated health, heavy drinking, current smoking status, physical activity, sleep problems, positive affect, life satisfaction, optimism, purpose in life, mastery, health mastery, financial mastery, depressive symptoms, hopelessness, negative affect, perceived constraints, loneliness, living with a spouse/partner, contact children ≥1x/week, contact other family ≥1x/week, contact friends ≥1x/week), personality factors (openness, conscientiousness, extraversion, agreeableness,

neuroticism), and the pre-baseline value of the exposure (informal helping). These variables were adjusted for in the wave pre-baseline to the exposure assessment (t_0;2006 or 2008).
[f]Includes only study participants with no history of diabetes (n=10,032).
[g]Includes only study participants with no history of hypertension (n=5,145).
[h]Includes only study participants with no history of stroke (n=11,915).
[i]Includes only study participants with no history of cancer (n=10,851).
[j]Includes only study participants with no history of heart disease (n=9,714).
[k]Includes only study participants with no history of lung disease (n=11,677).
[l]Includes only study participants with no history of arthritis (n=5,025). For this analysis, we did not adjust for arthritis in wave 1 because the cell size was too small and the analysis did not converge.
[m]Includes only study participants who were not overweight/obese (n=3,750).
[n]Includes only study participants who did not have physical functioning limitations (n=9,797).
[o]Includes only study participants who did not have cognitive impairment (n=10,407).
[p]Includes only study participants who did not have chronic pain (n=8,286).
*p<0.05 before Bonferroni correction; **p<0.01 before Bonferroni correction; ***p<0.05 after Bonferroni correction (the p-value cutoff for Bonferroni correction is p=0.05/35 outcomes=p<0.001).

Table A3. Complete-Case Analyses: Informal Helping and Subsequent Health and Well-being (Health and Retirement Study [HRS]: N ranged from: 5,576 to 8,482)[a,b,c,d]

	Informal Helping			
Outcomes	**0 hours (Reference)**	**1-49 hours RR/OR/β (95% CI)**	**50-99 hours RR/OR/β (95% CI)**	**≥100 hours RR/OR/β (95% CI)**
Physical Health				
All-cause mortality	1.00	0.67 (0.56, 0.81)***	0.69 (0.53, 0.89)**	0.66 (0.50, 0.88)**
Number of chronic conditions	0.00	-0.04 (-0.08, -0.01)*	-0.04 (-0.08, 0.01)	0.01 (-0.04, 0.06)
Diabetes	1.00	0.99 (0.88, 1.11)	0.97 (0.82, 1.14)	1.06 (0.91, 1.25)
Hypertension	1.00	0.99 (0.91, 1.06)	0.99 (0.90, 1.09)	1.01 (0.92, 1.12)
Stroke	1.00	0.85 (0.70, 1.04)	0.85 (0.66, 1.09)	0.80 (0.61, 1.06)
Cancer	1.00	0.93 (0.82, 1.06)	0.95 (0.80, 1.13)	0.92 (0.77, 1.09)
Heart disease	1.00	0.97 (0.87, 1.08)	0.96 (0.84, 1.11)	0.96 (0.83, 1.11)
Lung disease	1.00	1.02 (0.85, 1.22)	1.02 (0.80, 1.29)	1.01 (0.79, 1.29)
Arthritis	1.00	0.98 (0.91, 1.06)	1.01 (0.92, 1.11)	1.02 (0.93, 1.12)
Overweight/obesity	1.00	1.00 (0.93, 1.08)	1.00 (0.91, 1.10)	1.04 (0.94, 1.14)
Physical functioning limitations	1.00	0.86 (0.76, 0.97)*	0.76 (0.64, 0.90)**	0.80 (0.67, 0.95)*
Cognitive impairment	1.00	0.98 (0.85, 1.12)	0.84 (0.69, 1.02)	0.78 (0.63, 0.96)*
Chronic pain	1.00	1.02 (0.92, 1.12)	1.02 (0.90, 1.16)	1.02 (0.90, 1.16)
Self-rated health	0.00	0.08 (0.03, 0.12)**	0.09 (0.03, 0.16)**	0.09 (0.03, 0.15)**
Health Behaviors				
Heavy drinking	1.00	1.02 (0.77, 1.35)	1.01 (0.70, 1.43)	1.10 (0.78, 1.55)
Smoking	1.00	1.07 (0.87, 1.32)	1.11 (0.84, 1.47)	1.19 (0.90, 1.57)
Frequent physical activity	1.00	1.09 (1.01, 1.18)*	1.13 (1.03, 1.25)**	1.13 (1.02, 1.24)*
Sleep problems	1.00	1.03 (0.94, 1.14)	0.99 (0.87, 1.13)	1.06 (0.93, 1.20)
Psychological Well-Being				
Positive affect	0.00	0.06 (0.01, 0.11)*	0.08 (0.01, 0.14)*	0.06 (-0.01, 0.12)
Life satisfaction	0.00	0.04 (-0.01, 0.09)	0.03 (-0.04, 0.09)	0.01 (-0.06, 0.07)
Optimism	0.00	0.03 (-0.02, 0.09)	0.01 (-0.06, 0.07)	0.11 (0.05, 0.18)***
Purpose in life	0.00	0.09 (0.04, 0.14)***	0.11 (0.04, 0.17)**	0.14 (0.07, 0.20)***
Mastery	0.00	0.06 (0.00, 0.12)*	0.08 (0.01, 0.15)*	0.09 (0.02, 0.17)*
Health mastery	0.00	0.03 (-0.03, 0.08)	0.00 (-0.08, 0.07)	0.03 (-0.04, 0.10)
Financial mastery	0.00	0.04 (-0.02, 0.10)	0.02 (-0.06, 0.09)	0.02 (-0.05, 0.10)

Outcomes	Informal Helping			
	0 hours (Reference)	1-49 hours RR/OR/β (95% CI)	50-99 hours RR/OR/β (95% CI)	≥100 hours RR/OR/β (95% CI)
Psychological Distress				
Depression	1.00	0.90 (0.74, 1.09)	1.17 (0.92, 1.50)	0.90 (0.69, 1.18)
Depressive symptoms	0.00	-0.04 (-0.09, 0.01)	0.00 (-0.06, 0.06)	0.01 (-0.05, 0.07)
Hopelessness	0.00	-0.02 (-0.07, 0.03)	-0.01 (-0.08, 0.06)	-0.02 (-0.09, 0.04)
Negative affect	0.00	0.02 (-0.03, 0.07)	0.08 (0.01, 0.15)*	0.07 (0.00, 0.14)*
Perceived constraints	0.00	-0.03 (-0.08, 0.03)	-0.02 (-0.09, 0.05)	-0.03 (-0.10, 0.03)
Social Factors				
Loneliness	0.00	-0.03 (-0.08, 0.02)	-0.08 (-0.15, -0.02)*	0.00 (-0.07, 0.06)
Living with a spouse/partner	1.00	1.03 (0.94, 1.12)	1.02 (0.91, 1.14)	1.00 (0.90, 1.12)
Contact children ≥1x/week	1.00	1.02 (0.95, 1.11)	1.02 (0.93, 1.13)	1.07 (0.97, 1.18)
Contact other family ≥1x/week	1.00	1.04 (0.95, 1.14)	1.03 (0.91, 1.16)	1.03 (0.92, 1.16)
Contact friends ≥1x/week	1.00	1.10 (1.02, 1.20)*	1.15 (1.03, 1.27)**	1.10 (0.99, 1.22)

Note. Abbreviations: CI, confidence interval; OR, odds ratio; RR, risk ratio.

[a]If the reference value is "1," the effect estimate is OR or RR; if the reference value is "0," the effect estimate is β.

[b]The analytic sample was restricted to those who had participated in the baseline wave (t_1;2010 or 2012). All models adjusted for pre-baseline sociodemographic characteristics (age, sex, race/ethnicity, marital status, annual household income, total wealth, level of education, employment status, health insurance, geographic region), pre-baseline childhood abuse, pre-baseline religious service attendance, pre-baseline values of the outcome variables (diabetes, hypertension, stroke, cancer, heart disease, lung disease, arthritis, overweight/obesity, physical functioning limitations, cognitive impairment, chronic pain, self-rated health, heavy drinking, current smoking status, physical activity, sleep problems, positive affect, life satisfaction, optimism, purpose in life, mastery, health mastery, financial mastery, depressive symptoms, hopelessness, negative affect, perceived constraints, loneliness, living with a spouse/partner, contact children ≥1x/week, contact other family ≥1x/week, contact friends ≥1x/week), personality factors (openness, conscientiousness, extraversion, agreeableness, neuroticism), and the pre-baseline value of the exposure (informal helping). These variables were adjusted for in the wave pre-baseline to the exposure assessment (t_0;2006 or 2008).

[c]An outcome-wide analytic approach was used, and a separate model for each outcome was run. A different type of model was run depending on the nature of the outcome: 1) for each binary outcome with a prevalence of ≥10%, a generalized linear model (with a log link and Poisson distribution) was used to estimate a RR; 2) for each binary outcome with a prevalence of <10%, a logistic regression model was used to estimate an OR; and 3) for each continuous outcome, a linear regression model was used to estimate a β.

[d]All continuous outcomes were standardized (mean=0; standard deviation=1), and β was the standardized effect size.

*p<0.05 before Bonferroni correction; **p<0.01 before Bonferroni correction; ***p<0.05 after Bonferroni correction (the p-value cutoff for Bonferroni correction is p=0.05/35 outcomes=p<0.001).

Table A4. Informal Helping and Subsequent Health and Well-being (Health and Retirement Study [HRS]: N=12,998) – Baseline Adjustment for Formal Volunteering and Caregiving[a,b,c,d]

	Informal Helping			
Outcomes	**0 hours (n=6,070) (Reference)**	**1-49 hours (n=3,603) RR/OR/β (95% CI)**	**50-99 hours (n=1,651) RR/OR/β (95% CI)**	**≥100 hours (n=1,674) RR/OR/β (95% CI)**
Physical Health				
All-cause mortality	1.00	0.71 (0.61, 0.83)***	0.79 (0.64, 0.98)*	0.74 (0.59, 0.94)*
Number of chronic conditions	0.00	-0.03 (-0.06, 0.00)*	-0.03 (-0.07, 0.01)	0.00 (-0.04, 0.04)
Diabetes	1.00	0.98 (0.90, 1.07)	0.99 (0.88, 1.12)	1.02 (0.89, 1.16)
Hypertension	1.00	0.99 (0.94, 1.05)	0.98 (0.91, 1.06)	1.02 (0.94, 1.10)
Stroke	1.00	0.88 (0.76, 1.01)	0.87 (0.72, 1.07)	0.81 (0.66, 1.00)
Cancer	1.00	0.94 (0.84, 1.04)	0.90 (0.78, 1.03)	0.93 (0.80, 1.07)
Heart disease	1.00	0.97 (0.90, 1.05)	0.99 (0.89, 1.11)	0.98 (0.87, 1.10)
Lung disease	1.00	0.97 (0.85, 1.11)	1.03 (0.86, 1.23)	0.94 (0.78, 1.13)
Arthritis	1.00	1.00 (0.95, 1.06)	1.01 (0.94, 1.09)	1.02 (0.95, 1.10)
Overweight/obesity	1.00	1.01 (0.96, 1.07)	1.01 (0.94, 1.09)	1.03 (0.96, 1.11)
Physical functioning limitations	1.00	0.86 (0.78, 0.94)**	0.77 (0.67, 0.89)***	0.82 (0.72, 0.94)**
Cognitive impairment	1.00	0.96 (0.87, 1.06)	0.80 (0.68, 0.94)**	0.84 (0.72, 0.99)*
Chronic pain	1.00	1.00 (0.92, 1.08)	1.00 (0.90, 1.11)	0.99 (0.90, 1.10)
Self-rated health	0.00	0.07 (0.04, 0.11)***	0.08 (0.03, 0.14)**	0.08 (0.03, 0.14)**
Health Behaviors				
Heavy drinking	1.00	1.07 (0.81, 1.41)	1.14 (0.81, 1.60)	1.24 (0.90, 1.70)
Smoking	1.00	1.05 (0.91, 1.22)	1.08 (0.87, 1.34)	1.17 (0.95, 1.44)
Frequent physical activity	1.00	1.07 (1.01, 1.14)*	1.11 (1.03, 1.19)**	1.10 (1.02, 1.18)*
Sleep problems	1.00	0.99 (0.92, 1.07)	0.95 (0.85, 1.05)	1.01 (0.91, 1.13)
Psychological Well-Being				
Positive affect	0.00	0.06 (0.02, 0.11)**	0.09 (0.03, 0.15)**	0.07 (-0.01, 0.15)
Life satisfaction	0.00	0.02 (-0.02, 0.07)	0.03 (-0.02, 0.08)	0.01 (-0.04, 0.06)
Optimism	0.00	0.03 (-0.02, 0.08)	0.03 (-0.03, 0.09)	0.10 (0.04, 0.16)**
Purpose in life	0.00	0.08 (0.05, 0.12)***	0.10 (0.05, 0.15)***	0.11 (0.06, 0.16)***
Mastery	0.00	0.05 (0.01, 0.10)*	0.09 (0.01, 0.16)*	0.11 (0.04, 0.17)**
Health mastery	0.00	0.04 (0.00, 0.09)	0.06 (0.00, 0.12)*	0.08 (0.00, 0.15)
Financial mastery	0.00	0.04 (-0.01, 0.08)	0.04 (-0.03, 0.10)	0.04 (-0.03, 0.12)

Outcomes	Informal Helping			
	0 hours (n=6,070) (Reference)	1-49 hours (n=3,603) RR/OR/β (95% CI)	50-99 hours (n=1,651) RR/OR/β (95% CI)	≥100 hours (n=1,674) RR/OR/β (95% CI)
Psychological Distress				
Depression	1.00	0.96 (0.83, 1.10)	1.05 (0.87, 1.26)	0.94 (0.77, 1.15)
Depressive symptoms	0.00	-0.03 (-0.07, 0.00)	-0.03 (-0.09, 0.02)	-0.02 (-0.07, 0.04)
Hopelessness	0.00	-0.05 (-0.09, -0.01)*	-0.05 (-0.11, 0.01)	-0.05 (-0.11, 0.00)
Negative affect	0.00	0.02 (-0.03, 0.07)	0.03 (-0.03, 0.09)	0.05 (-0.01, 0.11)
Perceived constraints	0.00	-0.05 (-0.10, 0.00)*	-0.05 (-0.11, 0.00)	-0.05 (-0.11, 0.01)
Social Factors				
Loneliness	0.00	-0.02 (-0.06, 0.02)	-0.08 (-0.14, -0.02)**	-0.04 (-0.10, 0.03)
Living with a spouse/partner	1.00	1.00 (0.94, 1.07)	1.02 (0.95, 1.11)	1.02 (0.94, 1.10)
Contact children ≥1x/week	1.00	1.01 (0.96, 1.07)	1.02 (0.94, 1.10)	1.04 (0.97, 1.12)
Contact other family ≥1x/week	1.00	1.03 (0.96, 1.10)	1.02 (0.93, 1.11)	1.04 (0.95, 1.14)
Contact friends ≥1x/week	1.00	1.09 (1.01, 1.17)*	1.12 (1.03, 1.22)**	1.09 (1.00, 1.20)

Note. Abbreviations: CI, confidence interval; OR, odds ratio; RR, risk ratio.

[a]If the reference value is "1," the effect estimate is OR or RR; if the reference value is "0," the effect estimate is β.

[b]The analytic sample was restricted to those who had participated in the baseline wave (t_1;2010 or 2012). Multiple imputation was performed to impute missing data on the exposure, covariates, and outcomes. All models adjusted for pre-baseline sociodemographic characteristics (age, sex, race/ethnicity, marital status, annual household income, total wealth, level of education, employment status, health insurance, geographic region), pre-baseline childhood abuse, pre-baseline religious service attendance, pre-baseline values of the outcome variables (diabetes, hypertension, stroke, cancer, heart disease, lung disease, arthritis), overweight/obesity, physical functioning limitations, cognitive impairment, chronic pain, self-rated health, heavy drinking, current smoking status, physical activity, sleep problems, positive affect, life satisfaction, optimism, purpose in life, mastery, health mastery, financial mastery, depressive symptoms, hopelessness, negative affect, perceived constraints, loneliness, living with a spouse/partner, contact children ≥1x/week, contact other family ≥1x/week, contact friends ≥1x/week), personality factors (openness, conscientiousness, extraversion, agreeableness, neuroticism), and the pre-baseline value of the exposure (informal helping). These variables were adjusted for in the wave pre-baseline to the exposure assessment (t_0;2006 or 2008). These models also adjusted for baseline formal volunteering and caregiving.

[c]An outcome-wide analytic approach was used, and a separate model for each outcome was run. A different type of model was run depending on the nature of the outcome: 1) for each binary outcome with a prevalence of ≥10%, a generalized linear model (with a log link and Poisson distribution) was used to estimate a RR; 2) for each binary outcome with a prevalence of <10%, a logistic regression model was used to estimate an OR; and 3) for each continuous outcome, a linear regression model was used to estimate a β.

[d]All continuous outcomes were standardized (mean=0; standard deviation=1), and β was the standardized effect size.

*p<0.05 before Bonferroni correction; **p<0.01 before Bonferroni correction; ***p<0.05 after Bonferroni correction (the p-value cutoff for Bonferroni correction is p=0.05/35 outcomes=p<0.001).