

A Service of

ZBW

Leibniz-Informationszentrum Wirtschaft Leibniz Information Centre for Economics

Eibich, Peter; Zai, Xianhua

Article — Published Version Are the grandparents alright? The health consequences of grandparental childcare provision

Journal of Population Economics

Provided in Cooperation with: Springer Nature

Suggested Citation: Eibich, Peter; Zai, Xianhua (2024) : Are the grandparents alright? The health consequences of grandparental childcare provision, Journal of Population Economics, ISSN 1432-1475, Springer, Berlin, Heidelberg, Vol. 37, Iss. 4, https://doi.org/10.1007/s00148-024-01044-5

This Version is available at: https://hdl.handle.net/10419/313833

Standard-Nutzungsbedingungen:

Die Dokumente auf EconStor dürfen zu eigenen wissenschaftlichen Zwecken und zum Privatgebrauch gespeichert und kopiert werden.

Sie dürfen die Dokumente nicht für öffentliche oder kommerzielle Zwecke vervielfältigen, öffentlich ausstellen, öffentlich zugänglich machen, vertreiben oder anderweitig nutzen.

Sofern die Verfasser die Dokumente unter Open-Content-Lizenzen (insbesondere CC-Lizenzen) zur Verfügung gestellt haben sollten, gelten abweichend von diesen Nutzungsbedingungen die in der dort genannten Lizenz gewährten Nutzungsrechte.



http://creativecommons.org/licenses/by/4.0/

Terms of use:

Documents in EconStor may be saved and copied for your personal and scholarly purposes.

You are not to copy documents for public or commercial purposes, to exhibit the documents publicly, to make them publicly available on the internet, or to distribute or otherwise use the documents in public.

If the documents have been made available under an Open Content Licence (especially Creative Commons Licences), you may exercise further usage rights as specified in the indicated licence.





ORIGINAL PAPER



Are the grandparents alright? The health consequences of grandparental childcare provision

Peter Eibich¹ · Xianhua Zai^{2,3,4}

Received: 28 September 2023 / Accepted: 9 September 2024 / Published online: 10 October 2024 © The Author(s) 2024

Abstract

This paper examines the causal effect of childcare provision on grandparents' health in the United States. We use the sex ratio among older adults' children as an instrument for grandparental childcare provision. Our instrument exploits that parents of daughters transition to grandparenthood earlier and invest more in their grandchildren than parents of sons. We estimate 2SLS regressions using data from the Health and Retirement Study. The results suggest that providing childcare is detrimental to grandparents' physical functioning and subjective health. We show that these effects increase with the intensity of grandchild care provision, and the effects are driven primarily by grandmothers.

Keywords Grandparents · Childcare provision · Instrumental variables · Health

JEL Classification $\, I10 \cdot J13 \cdot J14 \cdot C26$

1 Introduction

Grandparents across the globe play an important role in raising their grandchildren. For example, the US Census Bureau estimates that in 2011, 4.8 million children under age 5, or $\sim 24\%$ of young children, received care from their grandparents (Laughlin 2013).

Responsible editor: Xi Chen

Xianhua Zai zai@demogr.mpg.de

> Peter Eibich peter.eibich@dauphine.psl.eu

- ¹ Université Paris Dauphine-PSL, Place du Maréchal de Lattre de Tassigny, 75016 Paris, France
- ² Max Planck Institute for Demographic Research, 1 Konrad-Zuse-Str., 18057 Rostock, Germany
- ³ Max Planck University of Helsinki Center for Social Inequalities in Population Health, Rostock, Germany
- ⁴ Max Planck University of Helsinki Center for Social Inequalities in Population Health, Helsinki, Finland

In the UK, around 40% of grandparents provide regular care for their grandchildren, and 89% of these provide care at least once a week (Age UK 2017). In the EU, 21% of children under age 3 received some childcare from sources other than their parents or formal childcare in 2020, with substantial variation across countries (Eurostat 2022).¹ Grandparents providing childcare can reduce the cost of childrearing for young parents by substituting for formal care or own childcare provision, in particular in contexts with strong family ties (Battistin et al. 2014; Chen et al. 2000). Even in contexts where formal childcare is both available and affordable, grandparents often make important contributions by offering a flexible alternative source of childcare, e.g., in case of illness or during school holidays.

While childcare provided by grandparents is highly beneficial to parents (Dimova and Wolff 2011; Compton 2015; Bratti et al. 2018), the consequences for grandparents themselves are less clear. Looking after grandchildren might provide grandparents with physical and mental stimulation, thereby helping to maintain their health in old age. This would imply that providing childcare can be considered as "active ageing," i.e., an activity that benefits both older individuals and wider society. Yet, keeping up with young children can also be physically strenuous and stressful. The negative health effects of informal care provision by older parents or spouses have been documented extensively in the literature (Bom et al. 2018; Bom and Stöckel 2021; Heger 2017; Schmitz and Westphal 2015). It seems possible that grandparents find caring for young children similarly demanding. Therefore, the overall effect of grandchild care provision on health of grandparents is ambiguous. In this paper, we empirically estimate the effects of providing childcare for grandparents' health.

Previous studies on grandchild care provision and grandparents' health report contradictory findings. Several studies report that grandparents caring for their grand-children are in better health and have fewer mobility limitations and fewer depressive symptoms (Danielsbacka et al. 2019; Di Gessa et al. 2016; Ku et al. 2013; Wang et al. 2020; Zeng et al. 2021). Yet, interpreting these estimates as causal is challenging, because the transition to grandparenthood is not random. Lai et al. (2021) report that older adults expecting to become grandparents in the future are healthier than those who do not expect this transition. A possible explanation is that healthier individuals are, *ceteris paribus*, able to have more children than individuals in poor health, which in turn means that they are more likely to have grandparents and their lifespan overlaps longer with their grandchildren's lifespan (Margolis and Verdery 2019). Health is also an important precondition for all activities in old age such that grandparents in poor health.

A few previous studies address the endogeneity of grandchild care provision. Ates (2017) finds that the positive association between grandparents' childcare provision and health in Germany disappears when introducing individual fixed effects. While fixed effects can resolve bias from selection on time-invariant unobservable characteristics (e.g., long-term health conditions or family size), it does not address potential

¹ Other sources of childcare here includes grandparents, other household members, and professional child carers.

reverse causality introduced by an unexpected health shock that reduces a grandparent's capacity to provide childcare. Brunello and Rocco (2019) and Ku et al. (2012) use instrumental variables (IVs) to address such endogeneity. Brunello and Rocco (2019) use data on European grandparents from the Survey of Health, Ageing and Retirement in Europe (SHARE). Their IV strategy exploits variation in the propensity of grandparents to provide childcare due to the random timing of the survey and changes in the likelihood of grandchild care provision by the age of grandchildren. They find a sizable increase in depressive symptoms for grandparents providing childcare. Ku et al. (2012) examine Taiwanese grandparents and use marital status of parents and the number of grandchildren as IVs. Their findings indicate that even after addressing the endogeneity, grandchild care provision is beneficial for Taiwanese grandparents' health.

This study examines the causal effect of grandparental childcare provision on grandparents' health in the United States (US). We use the sex ratio (defined as the number of daughters relative to the total number of children) as an instrument to address the endogeneity of grandchild care provision. Our IV is motivated by two insights from the demographic literature on grandparenthood: parents of daughters transition to grandparenthood earlier than parents of sons, and grandparents are more likely to provide childcare for grandchildren born to a daughter than for those born to a son. Previous studies on the labor market consequences of grandparenthood have used the sex of the first-born child as an instrument for becoming a grandparent (Rupert and Zanella 2018). Using the sex ratio as an instrument for grandchild care provision (previously used, e.g., by Salm et al. (2021)) follows similar considerations as the sex of the first-born child,² but exploits more variation in the data.³ We conduct extensive diagnostic checks of the IV assumptions. While our instrument is strongly related to grandparental childcare provision, we also find weaker associations with marital status (Kabátek and Ribar 2021) and informal care receipt (Van Houtven and Norton 2004), which raise concerns about the exclusion restriction. We therefore implement the "plausibly exogenous" approach proposed by Conley et al. (2012) to derive bounds for our estimates that allow for reasonable violations of the exclusion restriction.

We use data from the Health and Retirement Study (HRS) with detailed health information on the number of functional limitations, self-rated health status, and depressive symptoms. Our results indicate that the effects of grandchild care provision on health are negative, implying that (similar to informal caregivers) grandparents looking after their grandchildren do so despite the impact it has on their health. We therefore conclude that grandparental childcare provision should not be considered as "active ageing."

Considering potential mechanisms, we show that the effects increase with the intensity of grandchild care provision. We find no evidence for activity substitution, i.e., grandparents caring for their grandchildren do not reduce their engagement in activities that are beneficial for their health (e.g., exercise). Heterogeneity analyses suggest

 $^{^2}$ The sex of a child is determined randomly at conception, and daughters have grandchildren earlier than sons.

 $^{^{3}}$ For individuals with one child, both instruments are identical. For individuals with two or more children, the sex of the first-born child only distinguishes two groups in the data (first-born daughter vs. first-born son), whereas the sex ratio distinguishes between three or more groups (two daughters, one daughter, no daughters for families with two children).

that these effects are driven primarily by grandmothers, which likely reflects their higher rates of grandchild care provision.

This study contributes to the literature by estimating a credibly identified causal effect of grandchild care provision on grandparents' health. We use an established instrument motivated by the demographic literature on grandparenthood, we conduct a battery of tests and falsification exercises of the IV assumptions, and we derive treatment bounds that allow for possible violations of the exclusion restriction. Our paper is also the first study to provide causal evidence in the US context. The contradictory findings by Brunello and Rocco (2019) and Ku et al. (2012) suggest that the health effects of grandparenting might be context-dependent. The US is a particularly interesting context characterized by both expensive formal childcare compared to some of the European countries examined by Brunello and Rocco (2019) and weaker family ties compared to East Asian societies (Ku et al. 2012).

The remainder of the paper is organized as follows: Section 2 describes the data, dependent and treatment variables, and sample statistics. Section 3 first motivates our instrument, describes the estimation strategy, discusses the assumptions for our IV model, and addresses potential violations of IV assumptions. Section 4 presents our main results, addresses IV validity issues, examines the robustness of our estimates, explores potential mechanisms and heterogeneity, and discusses the external validity of our results. Section 5 discusses our findings and concludes.

2 Data

2.1 Sample description

We use data from the HRS, a nationally representative longitudinal study of Americans aged 51 and above. Respondents are surveyed every other year since 1992. The survey includes different birth cohorts who enter the study as they become eligible. The core cohort, the HRS cohort, has been followed and interviewed since 1992. Since 1993, the HRS has included the Study of Assets and Health Dynamics Among the Oldest Old (AHEAD) cohort of individuals born before 1924, the Children of the Depression Age (CODA) cohort of people born between 1942 and 1930, and the War Babies cohort (WB) of individuals born between 1942 and 1947. An additional Early Baby Boomers (EBB) cohort of people born between 1948 and 1953 was added to the sample in 2004, and the Mid-Baby Boomers cohort of individuals born between 1954 and 1959 was added in 2010.

The HRS asks respondents (including cohabiting spouses) detailed information about their own demographic characteristics, health, employment, financial situation, and intergenerational transfers as well as demographic information about their family members such as children and parents. To explore the effect of grandchild care provision on grandparents' health, we restrict our working sample to HRS respondents aged between 50 to 80 who have at least one child. On the one hand, we aim to include as many potential grandparents as possible to maximize the sample size. On the other hand, we are concerned about the validity of our instrument if we include individuals older than 80. This older population is more likely to be frail and dependent and thus not able to provide grandchild care. We check the sensitivity of our results to this age restriction in Section 4.3 using a sample without age limits.⁴

Our study sample includes 120,066 observations (25,300 unique individuals) and covers the period from 1996 (wave 3) to 2014 (wave 12) in which the HRS asks respondents consistent questions on grandparents' childcare provision.

2.2 Dependent variables

The HRS includes detailed information on the health outcomes of respondents. We mainly focus on three dimensions of health: self-reported health status, physical functioning, and mental health.

First, the HRS asks respondents to self-report their general health status. Possible answers range from 1 for "excellent," 2 for "very good," 3 for "good," 4 for "fair," to 5 for "poor." While self-reported health is subjective and might be affected by reporting heterogeneity, it is a good predictor of mortality (Idler and Benyamini 1997; DeSalvo et al. 2006; Kuka 2020).

Second, we use more objective measures about physical health conditions. The HRS provides indices of functional limitations, such as limitations in Activities of Daily Living (ADLs) and limitations in Instrumental Activities of Daily Living (IADLs). The ADLs include items such as bathing, eating, dressing, getting in or out of bed, and walking across a room, and the IADLs assess difficulties in using the phone, managing money, taking medications, shopping for groceries, and preparing hot meals.⁵ All these indices range from 0 to 5. An index with a value of 5 means that an individual has difficulties with all activities considered, while a value of 0 means that the individual has no limitations.

Third, we use information about respondents' mental health. The HRS asks respondents about their mental health using the Center for Epidemiologic Studies Depression (CESD) score. The CESD score captures the number of adverse sentiments a respondent experienced all or most of the time in the past 2 years, including whether an individual was depressed, felt alone, felt sad, had restless sleep, felt everything was an effort, could not get going, felt unhappy, and did not enjoy life. The CESD scale has been validated as an instrument to identify major depression in older adults (Irwin et al. 1999). For all health measures, we consider other operationalizations as robustness checks in Section 4.3. We consider the four health measures described here as our main outcomes, because these measures are frequently used in academic studies as well as in clinical practice. We report results for alternative health measures based on limitations in gross motor skills, mobility, large muscle use, and fine motor skills as well as a measure of cognitive functioning in the appendix.

⁴ The distribution of unlimited age of HRS respondents in Appendix Fig. A1 is almost symmetric around age 70.

⁵ More details on the construction of these measures can be found in Chien et al. (2015).

2.3 Treatment indicators

Our treatment variable of interest is whether individuals provide grandchild care or not. There are two relevant questions in the data: First, the HRS asks respondents whether they and their spouse spent 100 or more hours taking care of their grandchildren or great-grandchildren since the last wave.⁶ If the answer is yes, respondents are asked to which child they provided grandchild care. This question was not asked in waves 1 and 2.7 Second, the HRS asks respondents to estimate their childcare hours provided in the last 2 years. This question is asked separately for the respondent and the spouse.⁸ For those who cannot remember the hours or do not know the exact hours or refuse to give the number of hours, the HRS further asks the minimum and maximum values of hours of grandchild care provided.⁹ There are both advantages and disadvantages of using each question to construct our treatment indicator of grandchild care provision. The first question does not distinguish between grandchild care provided by respondents or their spouses, which would introduce measurement errors since we are interested in estimating the health effects on those who are actually looking after the grandchildren. On the other hand, the question only requires respondents to answer "yes" or "no" and might thus be less affected by recall bias than asking for the exact number of grandchild care hours provided over the last 2 years.

For our main analysis, we use the self-reported number of hours of grandchild care by respondents. Among those who are providing childcare, the majority of grandparents provide less than 1000 h over 2 years. Appendix Fig. A2 shows the distribution of grandchild care hours reported by HRS respondents for these grandparents. We construct a binary indicator of grandparental childcare provision, which indicates whether the respondent reported 100 or more hours of grandchild care over the last 2 years. If the number of hours is missing and the minimum and maximum values are above 100, we assume that the respondent is providing childcare. To examine the potential measurement errors in the treatment variable, we use the first question on childcare provided by the respondent and their partner to construct an alternative treatment indicator, which is 1 if the answer is "yes," i.e., the respondent and their partner provided at least 100h of childcare since the last wave. We also explore other cutoffs for the self-reported number of childcare hours as robustness checks in Section 4.3.

⁶ The question asked in the HRS is "Did you or your husband/wife/partner or your late husband/late wife/late partner spend 100 or more hours in total in the last two years taking care of great-grandchildren/grandchildren?"

⁷ In wave 2 of the HRS, the AHEAD cohort was asked whether grandchild care was provided for a year or longer. This question is no longer asked from wave 3. The question is "Which of your children is the parent of those grandchildren (or great-grandchildren)?"

⁸ The grandchild care hour question in the HRS is "Roughly how many hours altogether did you spend since the last wave?" for the respondent and "Roughly how many hours altogether did your husband/wife/partner spend since the last wave?" for respondent's spouse.

⁹ The quote in HRS is "Did it amount to a total of less than MAX BREAKPOINT, more than MIN BREAK-POINT, or what?" The MIN BREAKPOINTs are 0, 200, 201, 500, and 501. The MAX BEAKPOINTs are 199, 200, 499, 500, and 5,000.

2.4 Sample statistics

Table 1 presents the summary statistics of the working sample of HRS respondents who are between ages 50 and 80 in each survey year. The average age of the sample is around 66. About 58% of the sample are female. The average educational attainment of the sample is around 12 years. On average, each respondent has between three to four children. The oldest child is on average about 44 years old, and the youngest child is on average about 35 years old. About half of the respondents' children are daughters. The majority of the sample is married or living with a partner and white.

		Mean	S.D	Obs
Demographics:				
Age		65.84	7.93	120,204
Female		0.58	0.49	120,204
Education (years)		12.14	3.19	119,989
Number of children		3.72	2.07	120,204
Age of oldest child		43.76	8.53	119,881
Age of youngest child		34.75	9.63	119,881
Marital status				
	Married/partnered	0.71	0.45	120,105
	Separated/divorced	0.11	0.32	120,105
	Widowed	0.15	0.36	120,105
	Never married	0.01	0.11	120,105
Race/ethnicity				
	White	0.78	0.41	120,070
	Black/African-American	0.16	0.37	120,070
	Other	0.05	0.22	120,070
Instrumental variable:				
Sex ratio		0.49	0.29	120,066
Grandparent's provision of childcare (grandparenting):				
Grandparenting for at least one child (Q1)		0.33	0.47	120,101
Grandparenting for at least 100 h (Q2)		0.28	0.45	120,204
Heath variables:				
Self-report health status (1-5)		2.88	1.12	120,142
ADL limitations (0-5)		0.30	0.87	120,130
IADL limitations (0–5)		0.25	0.79	120,129
CESD score (0–8)		1.50	1.99	112,246

 Table 1
 Summary statistics of the sample

Notes: The data used are from the HRS 1996 to 2014 of individuals who are ages 50 to 80. Definitions of these variables can be found in Appendix Table A1

Approximately 28 to 33% of respondents provide some grandchild care according to the different definitions discussed earlier.

The average self-reported health status of respondents is good. The average ADL score and IADL score are both close to zero, which indicates that our sample is relatively healthy. The average CESD depression score is 1.5 out of 8. Detailed definitions of these variables are provided in Appendix Table A1.

3 Methods

In this section, we review findings from the demographic literature on grandparenthood and grandchild care to motivate the sex ratio as our IV for providing grandchild care. Then, we discuss the estimation strategy and provide evidence for the assumptions required for a causal interpretation in the IV framework.

3.1 Sex ratio as an IV for grandparental childcare provision

The transition to grandparenthood as well as the decision to provide grandchild care are endogenous choices, which depend on many factors that are plausibly related to health. For example, parents with larger families are more likely to become grandparents (Margolis and Verdery 2019), and parents who give birth earlier in life are more likely to become grandparents at younger ages. Family size and age at the first birth have been linked to health and mortality of mothers in particular (Mirowsky 2005; Wu and Li 2012), but they are also related to socioeconomic status (Adserá 2017). Beyond the transition to grandparenthood, grandparents' capacity to provide grandchild care depends, among other factors, on the proximity between grandparents and their adult children (Compton 2015), as well as their health (Eibich and Siedler 2020).

In this study, we address the endogeneity of grandchild care provision by using the sex ratio, defined as the number of daughters divided by the total number of children of a respondent, as an instrument.¹⁰ The instrument relies on two distinct mechanisms that link the sex ratio to grandparents' childcare provisio—(i) parents of daughters transition to grandparenthood earlier than parents of sons, and (ii) parents of daughters invest more in their offspring than parents of sons.

It is well-documented that women tend to give birth earlier than men (Margolis and Verdery 2019). This implies, *ceteris paribus*, that parents of daughters will transition to grandparenthood earlier than parents of sons. The gender of a child can be considered as good as randomly determined; thus, the gender of a person's first-born child might serve as a suitable instrument that predicts the transition into grandparenthood (and subsequently grandchild care provision) (Rupert and Zanella 2018). This is also borne out in our data: Fig. 1 shows the share of individuals who are providing grandchild

¹⁰ The total number of children is defined very broadly and potentially includes deceased as well as non-biological children (e.g., adopted or step-children). The number of living children is arguably a more relevant predictor of grandparenting; however, selective mortality among male children might bias our results. Reassuringly, our results remain robust using the total number of living children to define the sex ratio instrument (Appendix Table A2).



Fig. 1 Sex of the first-born child and grandchild care provision Notes: The data used are from the HRS 1996 to 2014 of individuals who

Notes: The data used are from the HRS 1996 to 2014 of individuals who are ages 50 to 80. The sample is limited to individuals with only 1 child and at least 1 grandchild. This graph shows the share of grandparents who provide some grandchild care by age for individuals with a daughter and those with a son, respectively. Grandchild care is defined as an indicator that equals 1 if the estimated grandchild care hours reported by respondents are at least 100 h.

care by age for individuals with a daughter and those with a son, respectively, for older adults with one child and at least one grandchild. At most younger ages, older adults



Fig. 2 Grandparenthood for older parents with two children

Notes: The data used are from the HRS 1996 to 2014 of individuals who are ages 50 to 80 and have 2 children. This graph shows the share of those individuals who are grandparents by age. Grandparenthood is defined as an indicator that equals 1 if the number of grandchildren reported by the respondent is at least 1.

with a daughter are much more likely to provide grandchild care than those with a son. This gap narrows substantially with age and mostly disappears beyond age 70.

While the sex of the first-born child is a plausibly exogenous instrument, it also relies on very limited variation. Our sex ratio instrument exploits that parents of daughters (on average) transition to grandparenthood earlier than parents of sons, regardless of birth order. Figure 2 shows the likelihood of grandparenthood for older adults with two children. Until about age 60, the likelihood of becoming a grandparent for parents with two daughters is considerably higher than for all other groups. The likelihood is very similar for parents with one daughter and one son, regardless of the birth order. The likelihood of becoming a grandparent tends to be the lowest for parents of two sons, although the differences between groups largely vanish from age 70 onward. This pattern is likely driven by the lower variation in age at first birth among women than among men (Margolis and Verdery 2019),¹¹ which implies that, conditional on family size, the number of daughters is predictive of an earlier transition to grandparenthood.¹²

The sex of a child not only affects their parents' likelihood to become grandparents, but also the extent of their involvement with the grandchild. Maternal grandparents tend to invest more time into grandchild care than paternal grandparents (Compton and Pollak 2011; Danielsbacka et al. 2011). The literature has proposed three possible explanations for this difference: First, maternal grandparents share a longer lifetime with their grandchildren (Margolis and Verdery 2019). As discussed above, they tend to become grandparents earlier in life and are consequently younger and on average healthier than paternal grandparents. Hence, they can invest more into their grandchildren than paternal grandparents. Second, from an evolutionary perspective, grandparents invest into their grandchildren to ensure the survival of their kin. Since there is more uncertainty around paternal kinship, grandparents will invest preferentially into their daughters' offspring rather than into their sons' (Danielsbacka et al. 2011). Third, mothers tend to share stronger bonds with their daughters (Somary and Strieker 1998) and might therefore invest more into their daughters' children. While the first mechanism suggests that any differences in grandchild care provision are driven by differences in the timing of the transition to grandparenthood, the second and third explanations imply that children's sex is also predictive of grandparental childcare provision conditional on the timing of grandparenthood. Figure 3 shows that from age 60 onwards until about age 75, grandparents with two daughters are most likely to provide grandchild care, whereas grandparents with two sons tend to be the least likely group to provide grandchild care.¹³

The third mechanism (emotional bonds between mothers and daughters) also raises concerns about instrument validity, since these bonds might either affect other outcomes or are in turn affected by unobserved confounders. For example, Somary and Strieker (1998) report few differences in grandparents' behavior across lineage, but note that they control for proximity between grandparents and grandchildren, which

¹¹ The lower variation in age at first birth among women implies that parents with a first-born son and a second-born daughter will in many cases become grandparents due to the first birth of their daughter.

¹² While family size is in itself predictive of grandparenthood, it is also endogenous, and we therefore condition on family size.

¹³ Fig. 2 shows the probability of grandparenthood, whereas Fig. 3 shows the probability of grandparental childcare provision conditional on grandparenthood.

Page 11 of 32 Are the grandparents alright? The health... 71 6 Probability of Grandchild Care Provision .5 .4 .3 .2 1 55 60 65 70 50 80 75 Age of Respondent --- Daughter-Son Daughter-Daughter Son-Daughter Son-Son

Fig. 3 Grandparental childcare for older parents with two children Notes: The data used are from the HRS 1996 to 2014 of individuals who are ages 50 to 80 and have 2 children and at least 1 grandchild. This graph shows the share of those individuals who provide grandchild care by age. Grandparental childcare is defined as an indicator that equals 1 if the estimated grandchild care hours reported by respondents are at least 100 h.

plays an important role for grandchild care investments (Compton 2015). It seems plausible that the proximity between grandparents and their children might also affect grandparents' health through other mechanisms than the provision of grandchild care. We discuss such concerns in more detail in Section 3.3.2

In summary, the literature suggests that maternal grandparents invest more into their grandchildren, because (i) they become grandparents earlier in life, and (ii) grandparents prefer to invest in their daughters' offspring. Importantly, it is not just the first child that matters-having one or more daughters is predictive of grandchild care provision regardless of birth order. However, family size (and consequently the absolute number of daughters) is endogenous and might be correlated with, e.g., the health of the older parents (Wu and Li 2012). We therefore use the sex ratio as our instrument to exploit random variation in the number of daughters born to a respondent conditional on family size.¹⁴ Compared to the sex of the first-born child ("birth order instrument") used in previous studies, the sex ratio instrument draws on more variation for respondents with two or more children. For example, for parents with one child, the birth order instrument and the sex ratio instrument exploit the same variation-parents of a daughter as compared to parents of a son. For parents with two children, the birth order instrument only distinguishes between parents with a first-born daughter and parents with a first-born son, regardless of the sex of the second child. The sex ratio instrument distinguishes parents with two sons from parents with one daughter and those with two daughters, under the assumption that parents with two daughters are most likely to provide grandchild care and parents with two sons are least likely to provide grandchild care. Since 88% of the HRS sample has two or more children, the

¹⁴ In contrast to the absolute number of daughters, the sex ratio is not correlated with family size ($\rho = -0.0007$ in our sample).

sex ratio instrument provides much more variation than the birth order instrument. Therefore, we expect that the sex ratio instrument should be stronger than the birth order instrument.¹⁵

3.2 Model specification

We estimate the first stage of our IV model as follows:

$$GCcare_{it} = \delta Sexratio_{it} + X'_{it}\beta + \epsilon_{it} \tag{1}$$

where $GCcare_{it}$ is a binary indicator for individuals providing grandchild care in year *t*. Sexratio_{it} is the ratio of the number of daughters to the number of children of an individual *i* in year *t*. X_{it} is a vector of covariates. In our preferred specification, we control for individual demographic characteristics such as age (quadratic polynomial), race/ethnicity, religion, gender, birth place and census region fixed effects as well as year fixed effects, fixed effects for the year of birth of the first-born child, age of the youngest child, cohort fixed effects of the individual, and family size. These covariates can all be considered as predetermined, and they capture important demographic differences in fertility and grandparenthood. While the validity of the instrument does not depend on the inclusion of these covariates, they might help to improve the precision of our estimates.¹⁶ ϵ_{it} are the standard errors clustered at the individual level which allow for correlation within individuals across waves.

We estimate the effect of grandchild care provision on health in the second stage of the model as follows:

$$Y_{it} = \alpha GC care_{it} + X_{it}^{'} \eta + \mu_{it}$$
⁽²⁾

where Y_{it} is an indicator of the health status of individual *i* in year *t*. The other controls X_{it} are the same as in Eq. 1. We estimate our IV model using linear two-stage least squares estimation (2SLS). Although we use longitudinal data, the model does not include individual fixed effects for two reasons: (*i*) If the required IV assumptions hold, the inclusion of individual fixed effects is not necessary for causal identification; (*ii*) we aim to exploit variation in the sex ratio between individuals arising from the random assignment of sex at conception of the child rather than variation in the instrument over time for individuals that occurs due to new births, deaths of children, or misreporting.¹⁷

¹⁵ We note that both instruments require the assumption that becoming a grandparent should not have a direct effect on health for older adults that do not provide care for their grandchildren. We discuss this assumption in more details later when considering the validity of the exclusion restriction, but here we note that we are generally comfortable with this assumption.

¹⁶ We include fixed effects for the year of birth of the first-born and age of the youngest child, because our sample also includes individuals with only one child. Table A3 shows that our results are robust to including fixed effects for year of birth for the oldest and the youngest child.

¹⁷ In our sample, 16.5% of individuals show variation in the sex ratio over time. Of these, 14% (2.4% of all individuals in the sample) experience the loss of a child, 28% (4.7% of individuals in our sample) appear to misreport the number of children in at least one wave, and the remaining 58% are new births or adoptions. Our results remain robust if we exclude individuals with changes in the sex ratio over time (Table A4) or if we use an alternative measure based on the number of living children (Table A2).

3.3 IV assumptions

The interpretation of our IV estimates as causal effects requires three assumptions: (i) reliability, i.e., the sex ratio should be correlated with grandparental childcare provision; (ii) validity, i.e., the sex ratio should be as good as randomly assigned and should not affect health through any other mechanisms than through grandchild care provision; and (iii) monotonicity, i.e., the sex ratio should affect the likelihood of providing grandchild care in the same direction (non-negative in this case) for all observations in our sample. In this section, we will discuss the plausibility of these assumptions in details.

3.3.1 Reliability

Table 2 shows estimates of the first stage using Eq. 1. In column 1, we regress our indicator of grandparental childcare provision only on the sex ratio instrument. The estimate suggests that—in line with the demographic literature on grandparenthood *ceteris paribus* older adults that only have daughters (i.e., a sex ratio of 1) are 5 percentage points more likely to provide grandchild care than older adults with only sons (i.e., a sex ratio of 0). For parents with two children, this would imply that every daughter increases the likelihood of grandchild care provision by 2.5 percentage

	(1)	(2)	(3)	(4)	(5)
Sex ratio	0.050***	0.050***	0.050***	0.048***	0.048***
	(0.007)	(0.007)	(0.007)	(0.007)	(0.007)
Year FE + Birth year FE of first- born	No	Yes	Yes	Yes	Yes
Cohort FE + Birth year FE of youngest-born	No	No	Yes	Yes	Yes
Demographics	No	No	No	Yes	Yes
Family size	No	No	No	No	Yes
Mean of dependent variable	0.28	0.28	0.28	0.29	0.29
Number of clusters	25,300	25,227	25,202	25,055	25,055
Observations	120,066	119,968	119,880	119,293	119,293
1st stage Kleibergen-Paap F- statistic	46.82	49.48	49.94	47.16	47.10

 Table 2
 First-stage estimates for grandparenting

Notes: The data used are from the HRS 1996 to 2014 of individuals who are ages 50 to 80. Each cell reports estimates from a separate specification using Eq. 1. The sex ratio is defined as the number of daughters divided by the total number of children of an individual. Grandparental childcare provision (grandparenting) is defined as an indicator that equals 1 if the estimated grandchild care hours reported by respondents are at least 100 h. Column 1 reports estimates without any controls. Column 2 adds year fixed effects and fixed effects for the year of birth of the first-born child of an individual. Column 3 adds age of the youngest child and cohort fixed effects of the individual. Column 4 includes individual demographics such as age (quadratic polynomial), race, religion, gender, birth place, and census region fixed effects. Column 5 further controls for the number of children of individuals. Standard errors are clustered at the individual level. * p < 0.10; **p < 0.05; ***p < 0.01

points. In columns 2–5, we successively add control variables to account for standard demographic characteristics. The point estimate of the sex ratio instrument is barely affected by the introduction of these controls. The Kleibergen-Paap F-statistic on the strength of the excluded instrument is larger than 46 in all models, which exceeds thresholds that have traditionally been used as a rule-of-thumb. This suggests that the sex ratio is indeed a sufficiently strong predictor of grandchild care provision, i.e., the reliability assumption holds.

3.3.2 Validity

The validity assumption consist of two parts—exogeneity of the instrument and the exclusion restriction. Exogeneity of the instrument implies that the instrument should be as good as randomly assigned. This does not require that the probability of giving birth to a son is the same as the probability of giving birth to a daughter, rather it means that the probability of having a daughter should not be correlated with any characteristics of the parents. We argue that this assumption is highly plausible. The sex of a child is randomly determined at conception, and sex-selective abortion or miscarriage rates are unlikely to play a major role in the context of this study. The exclusion restriction requires that the sex ratio should not affect older parents' health through any other pathway than its effect on grandparental childcare provision. As noted earlier, one concern is that our instrument identifies variation in grandchild care provision that arises because older parents of daughters transition to grandparenthood earlier than older parents of sons. If the transition to grandparenthood itself affects health in other ways than through childcare provision, the exclusion restriction would be violated. We are not very concerned about such direct effects of grandparenthood on health. While previous studies report effects of grandparenthood on labor market outcomes (Frimmel et al. 2022), changes in labor force participation are likely driven by (anticipated) grandchild care provision. It is possible that becoming a grandparent might have other effects on older adults, but we would expect that such effects are either unrelated to health (e.g., financial transfers to support their children) or have limited, positive effects (e.g., increasing life satisfaction or improvements in health behavior), in which case our estimated negative health effects can be regarded as a conservative upper bound of the true effect of grandchild care provision.¹⁸

The exclusion restriction will also be violated if the sex of a child affects individuals' lives beyond their transition to grandparenthood. Previous studies have documented that parents of daughters differ in several aspects from parents of sons, e.g., parents of daughters are at a higher risk of divorce (Kabátek and Ribar 2021), they are more likely to vote for left-wing political parties (Oswald and Powdthavee 2010), and daughters provide more informal care (Dahlberg et al. 2018). It is plausible that some of these differences (such as divorce or informal care receipt) might have consequences for

¹⁸ We cannot directly test whether grandparenthood affects health. Our instrument partly affects grandchild care provision through an earlier transition to grandparenthood, and we therefore cannot disentangle the effect of becoming a grandparent from the effect of providing grandchild care. However, in Tables A5 and A6 in the online appendix, we show that becoming a grandparent has no effects on a wide range of different social activities. We argue that this provides suggestive evidence for our assumption that becoming a grandparent does not affect health directly.





Notes: The data used are from the HRS 1996 to 2014 of individuals who are ages 50 to 80. The graph shows the point estimates and confidence intervals of the coefficients on the sex ratio using the regression model with controls including year fixed effects, fixed effects for the year of birth of the first-born child of an individual, age of the youngest child, cohort fixed effects of the individual, and individual demographics such as age (quadratic polynomial). The sex ratio is defined as the number of daughters divided by the total number of children of an individual.

health in later life. In particular, we would expect that a previous divorce is related to worse health and health behavior, whereas the health effects of informal care receipt may depend on the counterfactual.¹⁹ This means that it is not clear in which direction such violations of the exclusion restriction would affect our results.

We conduct two diagnostic checks for the validity of our instrument. First, we check for covariate balancing. If the validity assumption holds, we would expect that the distribution of covariates that are not affected by the treatment should be similar across the different values of the instrument, or put differently, there is no significant correlation between such covariates and our instrument. Note that failure to detect such correlations does not in turn imply that the validity assumption holds, since unobserved confounders remain a concern. We regress a battery of covariates as dependent variables on our sex ratio IV controlling for year and first-born fixed effects, cohort fixed effects for the individual, age fixed effects of the youngest child as well as individual demographic characteristics. Figure 4 shows the point estimates and confidence intervals of the regression coefficients on the sex ratio for each dependent variable listed on

¹⁹ Informal care receipt would likely improve health if the alternative is an unmet need for care, whereas the health effects are less clear if the alternative is formal care provided by a paid professional.

the vertical axis. The sex ratio appears to be uncorrelated with demographic characteristics of the respondent, implying that it is as good as randomly assigned. However, the sex ratio is negatively associated with being married and positively related to divorce, implying that (in line with the literature) more daughters increase the risk of divorce. Similarly, older adults with more daughters are also more likely to receive informal care in our data. Interestingly, older adults with more daughters are less likely to live with any of their children. There are no significant differences in the proximity to their children. Nevertheless, correlations between the sex ratio and marital status as well as informal care receipt raise concerns about the validity of our instrument. We note, however, that both the conditional and unconditional correlations between our instrument and these potential confounders (see Table A7) are considerably smaller than our first-stage estimates (Table 2) and our reduced form estimates (Table A8). This suggests that these mechanisms lead to no or at worst minor violations of instrument validity. We address such potential violations in more detail in Section 3.3.4.

We conduct another falsification exercise following (Angrist et al. 2010) to detect potential violations of the exclusion restriction. We estimate the reduced form regression for a subsample of HRS respondents that are not yet grandparents, but for whom we observe a later transition to grandparenthood. In this subsample, the sex ratio instrument is not supposed to predict grandchild care provision (i.e., the treatment) since none of the respondents have grandchildren. The exclusion restriction requires that the instrument is associated with the outcome only through its effect on the treatment. This implies that there should be no significant relationships between the sex ratio instrument and the health outcomes in the reduced form regression for individuals without grandchildren, because there is no valid first stage in this subsample. A significant reduced form estimate signals a violation of IV validity because it would suggest that the sex ratio is related to health through pathways other than grandchild care. Table 3 reports the reduced form estimates from our preferred model specification for the subsample of respondents who are not yet grandparents. None of the estimates is statistically significant, and the point estimates are very small and close to zero (the reduced form estimates for our working sample are shown in Appendix Table A8 for comparison).

Table A9 repeats this exercise for the full sample of respondents without grandchildren, which includes individuals that might never become grandparents.²⁰ The results are in line with our earlier findings, i.e., we fail to detect any violations of the exclusion restriction.

3.3.3 Monotonicity assumption

We test the monotonicity assumption by re-estimating the first stage regression of Eq. 1 for different subgroups within our working sample. The monotonicity assumption is satisfied if our instrument affects treatment in the same direction for all observations in the sample. This implies that the estimated effect of the sex ratio on grandparental

 $^{^{20}}$ Individuals that never become grandparents are plausibly a selective group, and we therefore argue that the earlier results presented in Table 3 are more informative.

Dependent variable	ADL	IADL	Self-reported health	CESD score
Sex ratio	-0.003	0.007	0.027	0.008
	(0.034)	(0.030)	(0.041)	(0.081)
Mean of dependent variable	0.32	0.27	2.98	1.66
Number of clusters	7990	7988	7993	7320
Observations	8638	8636	8640	7915

Table 3 Falsification exercise: reduced-form regressions

Notes: The data used are from the HRS 1996 to 2014 of the subsample of respondents who are ages 50 to 80 and who do not have grandchildren now but have later. ADL and IADL are the number of limitations reported by an individual with the range from 0 to 5. Self-reported health is general health status reported by HRS individuals, with values of 1 for "excellent," 2 for "very good," 3 for "good," 4 for "fair," to 5 for "poor." CESD score is the number of depressive symptoms reported by an individual with the range from 0 to 8. The details of each health outcome can be referred to Appendix Table A1. All models control for year fixed effects, fixed effects of the individual, individual demographics such as age (quadratic polynomial), race, religion, gender, birth place and census region fixed effects, and the number of children of individuals. Standard errors are clustered at the individual level. * p < 0.05; *** p < 0.01

childcare provision should be positive or zero for any arbitrarily defined subsample within our working sample. A significant negative effect in the first stage would imply a violation of monotonicity.²¹

Table 4 shows estimates of the first-stage regression for ten different subsamples defined by demographic characteristics. We note that the size of the first-stage estimate varies considerably—between 3.3 percentage points for men and 8.9 percentage points for older adults with more than four siblings. For all subsamples, the sex ratio instrument increases the likelihood of grandparental childcare provision and (with one exception) estimates are strongly significant.

Additionally, we also examine the robustness of the first-stage estimate across random subsamples from our working sample. We repeatedly draw a random 25% subsample from our working sample and re-estimate the first-stage regression in this subsample. Figure 5 shows the distribution of the point estimates across 1000 random draws. While there is considerably variation in the magnitude of the effect, the point estimate is positive across all subsamples.

In summary, these tests and falsification exercises suggest that the reliability and monotonicity assumptions hold, but the exclusion restriction may be violated because the sex ratio also affects other characteristics that are plausibly linked to health.

3.3.4 Addressing violations of the exclusion restriction

We address these potential violations of the exclusion restriction using the "plausibly exogenous" approach proposed by Conley et al. (2012). In their framework, violations

²¹ A negative but insignificant point estimate might either reflect a true zero effect (which does not violate monotonicity) or a violation of monotonicity.

Table 4 Falsification exercise: monot	onicity									
Subsample	Male	Female	≤ 4 Children	> 4 Children	\leq 4 Siblings	> 4 Siblings	Married	Not married	White	Not white
Sex ratio	0.027**	0.063***	0.050^{***}	0.031	0.040^{***}	0.079***	0.045***	0.056***	0.041^{***}	0.072***
	(0.010)	(0.00)	(0.007)	(0.019)	(0.008)	(0.015)	(0.009)	(0.011)	(0.008)	(0.015)
Mean of dependent variable	0.25	0.31	0.29	0.27	0.28	0.30	0.30	0.24	0.28	0.30
Number of clusters	10,873	14,182	19,091	7266	20,032	6639	19,049	9939	18,812	6243
Observations	49,665	69,628	86,829	32,461	92,318	26,326	84,183	35,012	93,429	25,864
1st stage Kleibergen-Paap F-statistic	6.45	45.68	45.96	2.50	25.89	28.33	28.19	26.37	27.14	24.34
Notes: The data used are from the Hi first-stage estimates from our preferrufirst-born child of an individual, age c gender, birth place and census region $p < 0.01$	XS 1996 to ed model s _l if the young fixed effect	2014 of inc pecification gest child, cc s, and the nu	lividuals who au in Eq. 1 for eac ohort fixed effec umber of childre	e ages 50 to 80 h subsample. A ts of the individ n of individuals	. The depender Il models contr ual, individual . Standard erro	it variable is gri ol for year fixeo demographics s rs are clustered	andparental d effects, fix uch as age (at the indivi	childcare proved effects for t quadratic poly dual level. $*p$ -	ision. Each the year of nomial), rac < $0.10; **p$	cell shows birth of the :e, religion, <0.05; ***



Fig. 5 Monotonicity assumption: distribution of first-stage estimates Notes: The data used are from the HRS 1996 to 2014 of individuals who are ages 50 to 80. This graph shows the distribution of the point estimates of the first stage using Eq. 1 across 1000 random subsamples of the working sample.

of the exclusion restriction are represented by a direct effect of the instrument on the outcome. They propose four different methods that can be used to construct a confidence interval around an estimated treatment effect which remains valid even in the presence of (small) violations of the exclusion restriction. The methods differ in their assumptions about the plausible values of the direct effects of the instrument.

Here, we implement the "union of confidence intervals" approach, which imposes the fewest assumptions and yields the largest interval. This approach only requires that we specify the support for the direct effect of the instrument on the outcome. We consider the following second stage regression:

$$Y_{it} = \alpha GC care_{it} + X'_{it}\eta + \gamma Sexratio_{it} + \mu_{it}$$
(3)

In Eq. 3, γ represents a direct effect of the instrument on the outcome, which violates the exclusion restriction. If the true value of γ was known as γ_0 , we could simply estimate

$$(Y_{it} - \gamma_0 Sexratio_{it}) = \alpha GC care_{it} + X_{it}\eta + \mu_{it}$$
(4)

to conduct valid inference on α . If γ_0 is unknown, but the support of γ , *G*, is limited and known, then we can construct valid confidence intervals for each element of *G* by assuming that $\gamma = \gamma_0$ and estimating (4). The union of these confidence intervals then forms a valid confidence interval for α for any $\gamma \in G$.

Dependent variable	ADL	IADL	Self-reported health	CESD score
Grandparenting	-0.109***	-0.111***	-0.108^{***}	-0.140***
	(0.007)	(0.006)	(0.010)	(0.018)
Mean of dependent variable	0.30	0.25	2.88	1.49
Number of clusters	25,045	25,044	25,052	24,072
Observations	119,222	119,222	119,234	111,449

Table 5 OLS estimates

Notes: The data used are from the HRS 1996 to 2014 of individuals who are ages 50 to 80. Each cell shows OLS estimates from our preferred model specification with full controls for each dependent variable. ADL and IADL are the number of limitations reported by an individual with the range from 0 to 5. Self-reported health is general health status reported by HRS individuals, with values of 1 for "excellent," 2 for "very good," 3 for "good," 4 for "fair," to 5 for "poor." CESD score is the number of depressive symptoms reported by an individual with the range from 0 to 8. The details of each health outcome can be referred to Appendix Table A1. All models control for year fixed effects, fixed effects for the year of birth of the first-born child of an individual, age of the youngest child, cohort fixed effects of the individual, individual demographics such as age (quadratic polynomial), race, religion, gender, birth place and census region fixed effects, and the number of children of individuals. Standard errors are clustered at the individual level. * p < 0.05; *** p < 0.01

We argue that the falsification exercises reported in Tables 3 and A9 represent our best estimates of a direct effect of the sex ratio on older adults' health.²² The point estimates reported in these tables are small and insignificant, but the standard errors indicate that we cannot exclude the possibility of larger direct effects that are comparable in magnitude to the reduced form effects reported in Table A8. We therefore construct confidence intervals for our causal effect that allow for direct effects of the instruments that as large as the standard errors reported in Table 3 for each outcome.

4 Results

4.1 The effect of grandparental childcare provision on health

Before estimating the causal effect of grandparents' childcare provision on health using the sex ratio instrument, we examine this relationship using ordinary least squares regression (OLS). Comparing OLS and 2SLS estimates will provide an indication of the size and direction of the bias caused by the endogeneity of childcare provision. Table 5 shows the results for our preferred model specification for all four health outcomes. Note that for all health indicators, higher values represent worse health outcomes.

²² This might not be the case if the magnitude of the direct effect is correlated with treatment status, since the falsification exercises only include untreated individuals. However, we would expect that individuals experiencing large negative health effects (e.g., due to the effect of the instrument on divorce) are less likely to provide grandchild care than individuals experiencing no or milder consequences. In this scenario, the estimated direct effects in these falsification exercises would be larger than potential direct effects in the full sample, which would support our decision to use these estimates to define upper bounds on potential violations of the exclusion restriction.

Overall, the results of Table 5 show that grandchild care provision is associated with better health across all dimensions of health considered here. The estimates suggest that grandparents caring for their grandchildren have fewer limitations in ADL and IADL, they rate their subjective health as better, and they show fewer depressive symptoms. Although these results seemingly support the notion that active ageing is beneficial for older adults' health, they should not be interpreted as causal effects. It is plausible that grandparents in good health are more likely to provide childcare than grandparents with poor health, which implies that these estimates may reflect reverse causality rather than a causal effect of childcare provision on health.

The results from our 2SLS regressions in Table 6 confirm the presence of such reverse causality. The point estimates for all four health outcomes are positive, suggesting that grandchild care leads to worse health in the form of more functional limitations, more depressive symptoms, and worse self-reported health. Estimates for limitations in ADL and IADL as well as estimates for self-reported health are significant at the 5% level, whereas the effect on the CESD score is not significant. The Anderson-Rubin test (AR) reported in Table 6 is robust to potential problems caused by weak instruments and confirms the significance of these effects.

Tables A10-A13 in the online appendix show that these results are overall robust to the inclusion or exclusion of covariates. The magnitude of these effects ranges from 0.78 standard deviations (SD) for ADLs, 0.86 SD for IADLs, to 1.05 SD for self-reported health. These are substantial effect sizes, which suggest that childcare provision can be a strenuous activity for grandparents. We report estimates for alternative health outcomes measuring limitations in mobility and physical activity as well as cognitive functioning in Appendix Table A14. All estimates are statistically signifi-

Dependent variable	ADL	IADL	Self-reported health	CESD score
Grandparenting	0.683**	0.683**	1.178**	0.938
	(0.335)	(0.292)	(0.489)	(0.729)
Mean of dependent variable	0.30	0.25	2.88	1.49
S.D. of dependent variable	0.87	0.79	1.12	1.99
Number of clusters	25,045	25,044	25,052	24,072
Observations	119,222	119,222	119,234	111,449
Kleibergen-Paap F-statistic	46.89	46.82	47.41	50.77
AR F statistic	4.80	6.60	7.00	1.75
AR <i>p</i> -value	0.029	0.010	0.008	0.186

Table 6	2SLS	estimates
---------	------	-----------

Notes: The data used are from the HRS 1996 to 2014 of individuals who are ages 50 to 80. Each cell shows 2SLS estimates from our preferred model specification with full controls in Eq. 2 for each dependent variable. ADL and IADL are the number of limitations reported by an individual with the range from 0 to 5. Self-reported health is general health status reported by HRS individuals, with values of 1 for "excellent," 2 for "very good," 3 for "good," 4 for "fair," to 5 for "poor." CESD score is the number of depressive symptoms reported by an individual with the range from 0 to 8. The details of each health outcome can be referred to Appendix Table A1. All models control for year fixed effects, fixed effects for the year of birth of the first-born child of an individual, age of the youngest child, cohort fixed effects of the individual, individual demographics such as age (quadratic polynomial), race, religion, gender, birth place and census region fixed effects, and the number of children of individuals. Standard errors are clustered at the individual level. * p < 0.10; ** p < 0.05; *** p < 0.01

cant and indicate that grandparental childcare provision is detrimental to grandparents' health.

4.2 IV validity and "plausibly exogenous" estimates

We address potential violations of the exclusion restriction by constructing confidence intervals for the estimates in Table 6 using the "union of confidence intervals" approach developed by Conley et al. (2012). Following (3)–(4), we plot these confidence intervals against a sensitivity parameter δ , which defines the support of direct effect of the instrument on the outcome as [0, 2δ]. This means that for a given value of δ , the shown confidence interval remains valid as long as the violation of the exclusion restriction (i.e., the direct effect of the instrument on the outcome) does not exceed 2δ .

Figure 6 shows the confidence interval for the estimated effect of grandchild care provision on ADL limitations. The upper grey line shows the upper limit of the 95% confidence interval, which remains constant as we only consider violations of the exclusion restriction that would bias our estimates towards zero. The lower grey line represents the lower limit of the 95% confidence interval as a function of the sensitivity parameter δ . For any given value of δ , the shown confidence interval will be robust to a violation of the exclusion restriction that is at most of magnitude 2δ . The red vertical line shows the value of δ at which the largest permissible violation of the exclusion restriction corresponds to the size of the standard error on the estimate for ADLs in Table 3. Even in the presence of violations of the exclusion restriction that are as large as or slightly larger than the standard error for the reduced form estimate in the falsification exercise in Table 3, the 95% confidence interval for our estimate would still exclude zero. This includes potentially very large direct effects of the instrument. For example, a value of $\delta = 0.02$ is equivalent to a direct effect of the instrument (0.04) that





Notes: The data used are from the HRS 1996 to 2014 of individuals who are ages 50 to 80. This figure presents 95% confidence intervals for the effect of grandchild care provision on the ADL outcome for violations of the exclusion restriction that do not exceed 2δ . The intervals were constructed using the "union of confidence intervals" by Conley et al. (2012) from Eqs. 3–4. The vertical line corresponds to a value of γ equivalent to the size of the standard errors in Table 3.



Fig. 7 95% interval estimates on IADL

Notes: The data used are from the HRS 1996 to 2014 of individuals who are ages 50 to 80. This figure presents 95% confidence intervals for the effect of grandchild care provision on the IADL outcome for violations of the exclusion restriction that do not exceed 2δ . The intervals were constructed using the "union of confidence intervals" by Conley et al. (2012) from Eqs. 3–4. The vertical line corresponds to a value of γ equivalent to the size of the standard errors in Table 3.

is larger than our reduced form estimate for our working sample in Table A8 (0.033). The link between the sex ratio and marital status (-0.020) or informal care receipt (+0.019, see Table A7) is weaker than the first stage reported in Table 2 (+0.048), and therefore we would expect that any direct effects of the sex ratio on health that occur due to a higher risk of divorce or a higher likelihood of informal care receipt should plausibly be smaller than the reduced form effect observed in our working sample.²³ Existing studies on the health effects of divorce and informal care receipt support this assumption. For example, Zulkarnain and Korenman (2019) estimate that divorce increases the risk of reporting bad subjective health by around 5 percentage points, which would correspond to a violation of the exclusion restriction by $\gamma = 0.0024$.²⁴ This is considerably smaller than the maximum permissible violation of the exclusion restriction considered here. For informal care receipt, Barnay and Juin (2016) and Wang and Yang (2022) actually report positive health effects are larger.²⁵ Fig. 6 thus

 $^{^{23}}$ The Wald IV estimator can be derived as the ratio of the reduced form effect to the first stage effect, which in turn implies that the reduced form effect is the product of the first stage effect and the treatment effect. To illustrate our argument, consider the extreme case where the entire reduced form effect documented in Table A8 can be attributed to a violation of the exclusion restriction, because the sex ratio affects, e.g., marital status. Since the link between the sex ratio and marital status is less than half of the first-stage estimate in Table 2, the treatment effect of marital status on health would need to be more than twice as large than the effect of grandchild care provision on health to result in the same reduced form effect.

 $^{^{24}}$ We first translate the effect size reported by Zulkarnain and Korenman (2019) into a change in standard deviations (0.11), then multiply this by the effect of the instrument on divorce (0.02) reported above, and finally multiply the result with the standard deviation of our subjective health variable (1.12) to account for the differences in measurement.

²⁵ Ideally, we would consider the joint impact across all potential pathways that could violate the exclusion restriction. Yet, there are no empirical studies considering both divorce and informal care receipt, and it does not seem plausible to assume that these effects are independent.

implies that our estimate is robust to larger violations of the exclusion restriction than what we would consider plausible.

Similarly, Figs. 7 and 8 show that even in the presence of considerable violations of the exclusion restriction (denoted by the red line), our estimates of the negative health effects of grandparental childcare provision remain statistically significant. While we cannot rule out that the exclusion restriction for the sex ratio instrument is violated, the estimated confidence intervals suggest that our conclusions remain robust even to moderately sized direct effects of the sex ratio on health.

4.3 Robustness

We assess the sensitivity of our results in a series of further robustness check. First, we re-estimate our models using the full sample of HRS respondents regardless of age. The results (panel A in Appendix Table A15) are qualitatively similar to those from our main specification in Table 6. Second, we exclude respondents without grandchildren from the working sample. Our sex ratio instrument identifies variation in both the timing of grandparenthood and the likelihood to provide grandchild care for existing grandchildren. This implies that the control group in our main specification consists of both older adults that are not (yet) grandparents as well as grandparents that do not provide childcare for their grandchildren. Excluding older adults without grandchildren from the sample means that both treatment and control group consist exclusively of grandparents and our instrument identifies variation in the likelihood to provide grandchild care. The results in panel B of Appendix Table A15 are both quantitatively and qualitatively similar to our main results in Table 6.

Appendix Table A16 reports the estimates using several alternative definitions of grandchild care provision to assess whether our results might be affected by measure-





Notes: The data used are from the HRS 1996 to 2014 of individuals who are ages 50 to 80. This figure presents 95% confidence intervals for the effect of grandchild care provision on the self-reported health outcome for violations of the exclusion restriction that do not exceed 2δ . The intervals were constructed using the "union of confidence intervals" by Conley et al. (2012) from Eqs. 3–4. The vertical line corresponds to a value of γ equivalent to the size of the standard errors in Table 3.

ment errors in the treatment indicator. The definition used in our main specification is a binary indicator whether respondents report providing at least 100h of grandchild care over the past 2 years. The corresponding estimates (Table 6) are repeated in the panel A of Appendix Table A16 to facilitate comparisons. We consider two alternative definitions: (*i*) an alternative indicator for whether the respondent and their spouse or partner provided at least 100h of grandchild care over the past 2 years as discussed in Section 2.3, and (*ii*) a binary indicator that is based on the same information as our main specification but excludes observations who do not know the number of care hours and reported a maximum number of hours of childcare provision.²⁶ The results in panels B and C of Appendix Table A16 confirm that our results are robust to these changes across different treatment definitions.

In addition, we check the sensitivity of our estimates to different outcome definitions. Appendix Table A17 reports the results for the health indicators defined with different cutoffs.²⁷ The baseline column shows our main estimates in Table 6. The sign of the estimates on different health indicators is consistent with the baseline model. while the magnitude and significance varies across health outcome definitions. For ADLs, grandchild care provision significantly increases the likelihood of reporting between one and three limitations. For IADLs and subjective health, grandchild care is consistently harmful across different definitions. The magnitude of these effects generally decreases as the threshold increases, which indicates that most grandparents experience relatively mild limitations. For mental health, none of the estimates is statistically significant, although the sign indicates an increase in depressive symptoms across specifications. Overall, our estimates of the effects of grandchild care provision on health are robust across different outcome definitions. Finally, we also present first stage estimates using the first-born instrument in Appendix Table A18. As the estimates show, the sex of the first-born child is only weakly or not at all related to the provision of grandchild care, and we therefore do not show 2SLS estimates based on this instrumental variable.

4.4 Mechanisms and heterogeneity

Finally, we consider potential mechanisms that might explain why childcare provision has adverse effects on grandparents' health. We also examine effect heterogeneity across demographic subgroups, since differences in the magnitude of these effects may also provide insights towards such mechanisms. Caring for small children can be physically and mentally demanding, even more so for older adults who might experience declines in physical functioning. If grandchild care itself is a strenuous

 $^{^{26}}$ As discussed in Section 2.3, for respondents who report missing grandchild hours, the HRS further asks the minimum and maximum of the bracket range of care hours. We exclude 3644 respondents who reported 0 minimum care hours and 199 maximum care hours (3339) and 200 maximum care hours (305) since the actual care hours for this sample are ambiguous.

 $^{^{27}}$ All regressions are estimated using linear probability models. The results are robust to nonlinear IV estimation (results available on request).

activity, we would expect that the effects of grandchild care provision increase with intensity. To test this, we re-estimate our IV regressions using indicators based on different threshold values to distinguish between low- and high-intensity childcare provision—any care provision, >50 h, >200 h, >500 h, and >1000 h of grandchild care over the past 2 years. Panels D–H in Table A16 show that effect sizes are generally larger for more intensive grandchild care provision—when we consider only older adults as treated if they provide at least 500 h of grandchild care in the past 2 years, effect sizes are roughly twice as large as in our main specification using a cut-off of 100 h.

We also consider specifications using the actual hours of grandchild care provision in Appendix Tables A19 and A20. In line with the estimates in Table A16, estimates for grandchild care hours suggest that an increase in hours spent on grandchild care has negative effects on grandparents' health. The first stage estimates in Appendix Table A19 suggest that the instrument is only weakly related to hours, which is not surprising given the distribution of hours shown in Appendix Fig. A2.

Second, it is possible that grandchild care provision crowds out time investment in other activities that are beneficial to grandparents' health. We test this hypothesis by estimating our preferred IV specification using engagement in a range of different activities as outcome variables. We note that the information on social participation is only available for a smaller subsample of the HRS data.²⁸ The results in Table A21 remain inconclusive—we find negative effects of grandchild care provision on the frequency of watching TV and writing, but no significant effects on, e.g., social participation (e.g., volunteering, charity work) or exercise.

Finally, we examine effect heterogeneity by splitting our sample by gender and race/ethnicity.²⁹ Table A22 suggests that the effects in our baseline specification are primarily driven by grandmothers. This is not surprising—grandmothers tend to provide more childcare than grandfathers, and consequently, our instrument is only weakly related to grandchild care provision for men. It is therefore not clear whether the differences in the estimates for grandfathers and grandmothers reflect genuine treatment effect heterogeneity, or whether they simply show differences in the strength of the instrument. We find few differences between White and Black/Hispanic respondents.

4.5 Complier and external validity

The IV estimates in Table 6 represent a local average treatment effect (LATE), i.e., the effect of grandparental childcare provision on health for individuals whose decision to care for their grandchildren is determined by the sex ratio. Determining the size or characteristics of this complier population is not possible with our multi-valued instrument; however, based on the discussion of the mechanisms connecting the sex

²⁸ Social participation is covered in the Psychosocial and Lifestyle Questionnaire as a left-behind survey in HRS, which was introduced in 2006 and is given at every wave to a 50% subsample of core respondents. The consistent activity questions in this questionnaire are available from 2008.

²⁹ We choose to estimate regressions on separate samples rather than modeling interactions to allow the effects of covariates to also differ across subsamples.

ratio to grandchild care provision, we can draw some tentative conclusions. As noted in Section 3.1, complier with more daughters are more likely to provide grandchild care, because they transition to grandparenthood earlier and because they invest more in their daughters (e.g., due to stronger emotional bonds). This would suggest that complier providing childcare are on average younger and have a stronger bond with their children than "always taker" (i.e., grandparents providing childcare regardless of the values of the instrument).

If these younger grandparents are still active on the labor market, they might either have to reconcile their childcare provision with their working hours, creating a double burden, or they might choose to retire early (Rupert and Zanella 2018) with possible negative consequences for their health (Fitzpatrick and Moore 2018). It also seems plausible that younger grandparents (who have experienced less physical and cognitive decline) and those with a stronger bond to their children are more likely to continue providing grandchild care even if they perceive care provision to negatively affect their health. This suggests that treatment effects on the complier might be larger than effects in the general population. On the other hand, younger grandparents might be more resilient, because they have experienced less or no decline in physical and cognitive functioning. Table A23 suggests that the negative effects of grandchild care provision might be larger for grandparents aged 67 and older. However, we cannot draw definitive conclusions, because our instrument does not predict grandchild care provision as well among older grandparents. Our estimates are also likely not representative of grandparents that have to provide intensive childcare because the parents are absent (e.g., deceased, incarcerated, or labor migrants). In this case, we would expect the negative effects of grandchild care provision to be larger, because grandparents provide more intensive childcare and can obtain less respite. Moreover, the circumstances of the parents' absence might cause additional stress for grandchildren and grandparents. Finally, the US has neither particularly strong family ties nor extensive, affordable public childcare. In settings with stronger family ties (e.g., in Southern Europe or East Asia), grandparents might provide more intensive childcare. Yet, looking after one's grandchildren might also be perceived as more fulfilling and pleasant. It is therefore not clear whether we would expect stronger or weaker effects of grandchild care provision on health. In settings with extensive and affordable public childcare (e.g., Northern Europe), grandparents will likely provide less intensive childcare and have more opportunities for respite, and we would therefore expect effects on health to be less pronounced.

We provide some tentative empirical evidence by re-estimating our IV regressions in a marginal treatment effect (MTE) framework (Heckman and Vytlacil 2007). The MTE is estimated under the same assumptions as standard IV models and measures the expected treatment effect as a function of an individual's unobserved resistance to treatment. Intuitively, individuals select into treatment based on their observable characteristics (incl. the instrument) and an unobserved resistance to treatment (e.g., based on their expected gains from treatment). In this setting, the instrumental variable is used to estimate the propensity score of treatment, which in turn is used to estimate the marginal treatment effect at different values of the unobserved resistance to treatment (Brinch et al. 2017).³⁰ One attractive feature of the MTE framework is that the estimated MTEs can be used to derive different treatment effects, including the LATE or the average treatment effect.

Figures A3–A6 show the results for our four health outcomes. We note that for ADLs and subjective health, the marginal treatment effects are broadly stable across the distribution of the resistance to treatment, which suggests that the estimated effects from our IV regressions are likely to hold more broadly beyond the complier population. In line with the discussion above, we observe that for IADLs, the marginal treatment effect is a downward-sloping function, i.e., individuals who are more likely to be treated are more heavily affected than individuals who are less likely to be treated. For the CESD score, we observe an upward-sloping pattern. This suggests that even though our 2SLS estimate for mental health is not significant, some grandparents may indeed experience negative effects on their mental health (in line with findings by Brunello and Rocco (2019)). Overall, we interpret these MTE estimates as suggestive evidence that our estimated effects for ADLs, IADLs, and subjective health might have some validity beyond the complier population.

5 Discussion

This study examines the effect of childcare provision on grandparents' health in the US. We use the sex ratio as an instrument for grandparental childcare provision, drawing on insights from the demographic literature on grandparenthood. Our sex ratio instrument measures the share of daughters among all children born to a person, which captures that parents of daughters transition on average earlier into grandparenthood and grandparents are more likely to provide care for grandchildren born to their daughters than to grandchildren born to their sons. We conduct several tests and falsification exercises that suggest that the exclusion restriction for the sex ratio may not hold, because having a daughter is linked to other characteristics such as marital status and the receipt of informal care. We address these violations of the exclusion restriction by deriving 95% confidence intervals that remain valid in the presence of small or moderately sized direct effects of the sex ratio on health using the "plausibly exogenous" approach by Conley et al. (2012).

Our OLS results are in line with earlier studies suggesting that grandchild care provision is indeed positively associated with grandparents' health, but this association is likely driven by reverse causality. Once we address such endogeneity using the sex ratio as an IV, we find that effects of grandchild care provision on health are predominantly negative.

We find that grandparental childcare provision leads to an increase in ADLs by 0.79 standard deviations (SD), an increase in IADLs by 0.86 SD, and worsens self-reported health by 1.05 SD. These are substantial negative effects. However, we argue that the magnitude of these effects should be interpreted with caution. First, as we discuss in Section 3.3.4, it is possible that the exclusion restriction might be violated.

 $^{^{30}}$ Under the assumption that individuals are treated if their propensity score is larger than their resistance to treatment.

In particular, a higher likelihood of divorce would likely exert a negative influence on health, thus biasing our results away from zero. We construct confidence intervals that are robust to moderate violations of the exclusion restriction, yet this means that our results are set- and not point-identified. In other words, the magnitude of the effects might be lower than our 2SLS point estimates suggest. Second, it is possible that the LATE identified for the complier population might be larger than an average treatment effect. We provide some tentative evidence that for ADLs and self-reported health, the marginal treatment effects are stable; however, for IADLs, we indeed find that the marginal treatment effects follow a downward-sloping curve. Third, we note that all our health measures are discrete, and a 1-point change is the smallest possible change that an individual can experience. An estimated effect size of 1.2 for selfreported health (or 1.05 SD) does therefore not necessarily mean that some treated individuals experience a substantial health shock. Instead, it could also be that most treated individuals experience a small 1-point change in their health status. For these reasons, we focus in our interpretation primarily on the qualitative direction of these health effects rather than their magnitude.

We provide some tentative evidence that the negative effects of grandchild care provision on health are stronger for high-intensity childcare provision, whereas activity substitution does not seem to play an important role. The effects are more pronounced for grandmothers than for grandfathers. Given the relatively stable MTEs, this difference likely reflects differences in the strength of the instrument rather than effect heterogeneity between men and women.

Our findings suggest that previously reported positive associations between grandchild care and health are biased, likely due to reverse causality. Yet, it is likely that the context and methodology of the study also matter. Our IV regressions identify a LATE, which is determined by a complier population that is likely younger and has stronger bonds with their children than the general population of grandparents. We provide some tentative evidence that the health effects for older grandparents might be larger, yet the results are not entirely conclusive, because our instrument does not predict grandchild care provision well among older grandparents. Our estimated effects are also unlikely to be valid for grandparents that have to care for their grandchildren because the parents are absent. Studies using a different methodology are needed to examine the effects in these populations. Findings may also differ based on, e.g., the role of the family and the strength of family ties. The US is an interesting setting with neither particularly strong family ties (compared to, e.g., East Asia) nor with extensive subsidized formal childcare places (e.g., as in Northern Europe). It seems possible that health effects in this setting are very different from those reported, e.g., for China (Choi and Zhang 2021; Wang et al. 2020).

We also acknowledge a few limitations of our study. In particular, exploring potential mechanisms in more detail would require a more reliable measure of actual care hours, e.g., based on time use diaries. Similarly, it seems plausible that the effects of grandchild care may differ based on the tasks taken over by grandparents. For example, taking care of infants during the day or when parents are close-by may be much less demanding than looking after these children overnight. Unfortunately, such data is not available in the HRS. We alternatively considered engagement in a range of activities as potential mechanisms. While we find no evidence for activity substitution, this should only be considered as suggestive, since the results are based on a much smaller sample than our main findings.

In summary, our results show that grandparental childcare provision does not improve the health of grandparents, rather it may be detrimental. Good health is an important precondition for grandparents to provide childcare, and this reverse causality causes the frequently observed positive associations documented in the literature. This implies that childcare provision should not be considered as "active ageing" a socially desirable activity that preserves or improves older adults' health. Instead, childcare provision appears to be an activity that older adults engage in to help their family even though it may be detrimental to their own health or well-being. Consequently, family policies that improve the availability and affordability of formal childcare may generate positive externalities as grandparents may feel less obliged to help out their adult children to the detriment of their own health. Rather, they may enjoy the grandparents' privilege and pass on the baton when the going gets rough.

Supplementary Information The online version contains supplementary material available at https://doi.org/10.1007/s00148-024-01044-5.

Acknowledgements We would like to thank editor Xi Chen and three anonymous reviewers for helpful comments and suggestions that have improved the paper. We would like to thank Benedicte Apouey, Pedro Pita Barros, Judith Bom, Eric Bonsang, Katharina Janke, Vladimir Kozlov, Tim Riffe, Jeroen Spijker, Hansel Teo, and Christine Valente for their helpful comments and suggestions. We also thank seminar participants at the Department of Economics, University of Padova, the EuHEA seminar, the workshop "Behaviours, Health and Ageing" at Université Paris Dauphine, the EAPS Health, Morbidity and Mortality Working Group in Budapest and conference participants at the Society for Longitudinal and Lifecourse Studies conference (SLLS) in Munich and PAA for their helpful comments. Carlotta Dietrich and Lisa Rieker provided excellent research assistance.

Funding Open Access funding enabled and organized by Projekt DEAL. Zai was supported by grants to the Max Planck - University of Helsinki Center from the Jane and Aatos Erkko Foundation (#210046), the Max Planck Society (#5714240218), the University of Helsinki (#77204227), and the cities of Helsinki, Vantaa, and Espoo.

Data Availability The data used in the paper are public available from Health and Retirement Study (HRS) website which is managed by the Survey Research Center in the University of Michigan: https://hrsdata.isr.umich.edu/

Declarations

Conflict of Interest The authors declare no competing interests.

Open Access This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit http://creativecommons.org/licenses/by/4.0/.

References

- Adserá A (2017) Education and fertility in the context of rising inequality. Vienna Yearb Popul Res 15:63–92 Age UK (2017) 5 million grandparents take on childcare responsibilities
- Angrist J, Lavy V, Schlosser A (2010) Multiple experiments for the causal link between the quantity and quality of children. J Law Econ 28(4):773–824
- Ates M (2017) Does grandchild care influence grandparents' self-rated health? Evidence from a fixed effects approach. Soc Sci Med 190(C):67–74. Publisher: Elsevier
- Barnay T, Juin S (2016) Does home care for dependent elderly people improve their mental health? J Health Econ 45:149–160
- Battistin E, De Nadai M, Padula M (2014) Roadblocks on the road to grandma's house: fertility consequences of delayed retirement. IZA Discussion Paper 8071, Institute for the Study of Labor (IZA)
- Bom J, Bakx P, Schut F, van Doorslaer E (2018) The impact of informal caregiving for older adults on the health of various types of caregivers: a systematic review. Gerontologist 59(5):e629–e642
- Bom J, Stöckel J (2021) Is the grass greener on the other side? The health impact of providing informal care in the UK and the Netherlands. Soc Sci Med 269:113562
- Bratti M, Frattini T, Scervini F (2018) Grandparental availability for child care and maternal labor force participation: pension reform evidence from Italy. J Popul Econ 31:1239–1277
- Brinch CN, Mogstad M, Wiswall M (2017) Beyond LATE with a discrete instrument. J Polit Econ 125(4):985–1039
- Brunello G, Rocco L (2019) Grandparents in the blues. The effect of childcare on grandparents' depression. Rev Econ Househ 17(2):587–613. Publisher: Springer
- Chen F, Short SE, Entwisle B (2000) The impact of grandparental proximity on maternal childcare in China. Popul Res Policy Rev 19(6):571–590
- Chien S, Campbell N, Hayden O, Hurd M, Main R, Mallett J, Martin C, Meijer E, Miu A, Moldoff M et al (2015) RAND HRS data documentation, version N. RAND Corporation
- S-wE Choi, Zhang Z (2021) Caring as curing: grandparenting and depressive symptoms in China. Soc Sci Med 289:114452
- Compton J (2015) Family proximity and the labor force status of women in Canada. Rev Econ Household 13(2):323–358
- Compton J, Pollak RA (2011) Family Proximity, Childcare, and Women's Labor Force Attachment. National Bureau of Economic Research Working Paper Series, No. 17678. featured in NBER digest on 2012-03-30
- Conley TG, Hansen CB, Rossi PE (2012) Plausibly exogenous. Rev Econ Stat 94(1):260-272
- Dahlberg L, Berndt H, Lennartsson C, Schön P (2018) Receipt of formal and informal help with specific care tasks among older people living in their own home. national trends over two decades. Soc Policy Adm 52(1):91–110
- Danielsbacka M, Tanskanen AO, Coall DA, Jokela M (2019) Grandparental childcare, health and well-being in Europe: a within-individual investigation of longitudinal data. Social Science & Medicine
- Danielsbacka M, Tanskanen AO, Jokela M, Rotkirch A (2011) Grandparental child care in Europe: evidence for preferential investment in more certain kin. Evolutionary Psychology, 9(1):147470491100900102. Publisher: SAGE Publications Inc
- DeSalvo KB, Bloser N, Reynolds K, He J, Muntner P (2006) Mortality prediction with a single general self-rated health question. J Gen Intern Med 21(3):267–275
- Di Gessa G, Glaser K, Tinker A (2016) The health impact of intensive and nonintensive grandchild care in Europe: new evidence from SHARE. J Gerontol Series B 71(5):867–879
- Dimova R, Wolff F-C (2011) Do downward private transfers enhance maternal labor supply? Evidence from around Europe. J Popul Econ 24:911–933
- Eibich P, Siedler T (2020) Retirement, intergenerational time transfers, and fertility. Eur Econ Rev 124:103392
- Eurostat (2022) Living conditions in Europe childcare arrangements
- Fitzpatrick MD, Moore TJ (2018) The mortality effects of retirement: evidence from social security eligibility at age 62. J Public Econ 157:121–137
- Frimmel W, Halla M, Schmidpeter B, Winter-Ebmer R (2022) Grandmothers' labor supply. J Hum Resour 57(5):1645–1689

- Heckman JJ, Vytlacil EJ (2007) Econometric evaluation of social programs, part II: using the marginal treatment effect to organize alternative econometric estimators to evaluate social programs, and to forecast their effects in new environments. Handb Econ 6:4875–5143
- Heger D (2017) The mental health of children providing care to their elderly parent. Health Economics, 26(12):1617–1629. Publisher: John Wiley & Sons, Ltd
- Idler EL, Benyamini Y (1997) Self-rated health and mortality: a review of twenty-seven community studies. Journal of Health and Social Behavior, pp 21–37
- Irwin M, Artin KH, Oxman MN (1999) Screening for depression in the older adult: criterion validity of the 10-item Center for Epidemiological Studies Depression Scale (CESD). Arch Intern Med 159(15):1701–1704
- Kabátek J, Ribar DC (2021) Daughters and divorce. Econ J 131(637):2144-2170
- Ku L-J, Stearns S, Houtven C, Holmes G (2012) The health effects of caregiving by grandparents in Taiwan: an instrumental variable estimation. Review of Economics of the Household, 10(4):521–540. Publisher: Springer
- Ku L-JE, Stearns SC, Van Houtven CH, Lee S-YD, Dilworth-Anderson P, Konrad TR (2013) Impact of Caring for Grandchildren on the Health of Grandparents in Taiwan. J Gerontol Series B 68(6):1009– 1021
- Kuka E (2020) Quantifying the benefits of social insurance: unemployment insurance and health. Rev Econ Stat 102(3):490–505
- Lai DWL, Li J, Bai X (2021) To be or not to be: relationship between grandparent status and health and wellbeing. BMC Geriatr 21(1):204
- Laughlin L (2013) Who's minding the kids? Child care arrangements: Spring 2011. Technical Report P70-135, US Census Bureau
- Margolis R, Verdery AM (2019) A cohort perspective on the demography of grandparenthood: past, present, and future changes in race and sex disparities in the United States. Demography 56(4):1495–1518
- Mirowsky J (2005) Age at first birth, health, and mortality. J Health Soc Behav 46(1):32–50. Publisher: SAGE Publications Inc
- Oswald AJ, Powdthavee N (2010) Daughters and left-wing voting. Rev Econ Stat 92(2):213-227
- Rupert P, Zanella G (2018) Grandchildren and their grandparents' labor supply. J Public Econ 159:89-103
- Salm M, Siflinger BM, Xie M (2021) The effect of retirement on mental health: indirect treatment effects and causal mediation
- Schmitz H, Westphal M (2015) Short- and medium-term effects of informal care provision on female caregivers' health. J Health Econ 42:174–185
- Somary K, Strieker G (1998) Becoming a grandparent: a longitudinal study of expectations and early experiences as a function of sex and lineage1. Gerontologist 38(1):53–61
- Van Houtven C, Norton EC (2004) Informal care and health care use of older adults. J Health Econ 23(6):1159–1180
- Wang H, Fidrmuc J, Luo Q (2020) Grandparenting and well-being of the elderly in China. Technical report, Bank of Finland, Institute for Economies in Transition. Issue: 18/2020
- Wang Y, Yang W (2022) Does receiving informal care lead to better health outcomes? Evidence from China Longitudinal Healthy Longevity Survey. Res Aging 44(7–8):510–518
- Wu X, Li L (2012) Family size and maternal health: evidence from the One-Child policy in China. J Popul Econ 25(4):1341–1364
- Zeng Y, Chen Y-C, Lum TYS (2021) Longitudinal impacts of grandparent caregiving on cognitive, mental, and physical health in China. Aging & Mental Health, 25(11):2053–2060. Publisher: Routledge
- Zulkarnain A, Korenman S (2019) Divorce and health in middle and older ages. Rev Econ Household 17(4):1081–1106

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