# Is having more children beneficial for mothers' mental health in later life?
# Causal evidence from the National Health and Aging Trends Study

Thijs van den Broek, Ph.D.

Erasmus School of Health Policy and Management, Erasmus University Rotterdam, Posbus 1738, 3000 DR Rotterdam, The Netherlands; +31 10 408 85 77;

vandenbroek@eshpm.eur.nl

# Is having more children beneficial for mothers' mental health in later life? Causal evidence from the National Health and Aging Trends Study


**Abstract**

*Objectives.* Members of the baby boom cohorts had fewer children than their parents. Given that adult children are an important source of social support in later life, this may have implications for the mental health of new cohorts of older people. This study investigates whether having additional children protects white mothers aged 65 and older against mental health problems.

*Method.* Data are from Wave 1 and Wave 5 of the National Health and Aging Trends Study (n=3,858). An instrumental variable approach exploiting the preference for mixed-sex offspring is used to estimate the causal effect of additional children on the risk of elevated depression and anxiety symptomatology.

*Results.* The estimated instrumental variable model shows that additional children reduce the risk of suboptimal mental health among white mothers aged 65 and older.

*Conclusion.* Results suggest that declines in higher-order births may put new cohorts of older women at increased risk of suboptimal mental health.

## Introduction

Older persons – and in particular older women – rely heavily on their children, not only for informal care and practical support, but also for confiding and reassurance (Antonucci & Akiyama, 1987; Van der Pas, Van Tilburg, & Knipscheer, 2007). The parents of the baby boom generation were characterized by their high fertility, but subsequent cohorts had considerably fewer children (Cherlin, 2010; Devolder, González, & Gavino, 2002; Kirmeyer & Hamilton, 2011). This gives rise to concerns about the welfare of new cohorts of older people (e.g., Marcil-Gratton & Légaré, 1992; Ryan, Smith, Antonucci, & Jackson, 2012). When older women have fewer children, they may miss important social support resources, and this may, in turn, harm their mental health.

Drawing on data from the National Health and Aging Trends Study, the current study investigates the causal impact of the number of children on mothers' mental health in later life in the United States. Rather than comparing the mental health of mothers and childless women, as has often been done in earlier research (e.g., Bures, Koropeckyj-Cox, & Loree, 2009; Huijts, Kraaykamp, & Subramanian, 2013; Koropeckyj-Cox, 1998; Zhang & Hayward, 2001), it focuses on the effect of additional children on mothers' mental health. A closer look at mental health differences between mothers of different parity is called for, because fertility declines in developed countries have to a substantial extent been driven by declines in higher-order births (Devolder et al., 2002). As explained in further detail later, the current study extends earlier work on the association between number of children and mother's mental health (e.g., Grundy, Van den Broek, & Keenan, 2019; Spence, 2008) by adopting an instrumental variable approach that is less prone to bias due to reverse causality and selection.

## Background and hypothesis

The task specificity model developed by Litwak (1985; Messeri, Silverstein, & Litwak, 1993) posits that older persons' adult children have particular characteristics, such as long-term internalized commitment, and that these characteristics make them likely providers of a range of different types of support, including emergency financial assistance and acute help during illness. Although it has been argued that the family has lost many of its functions since the 1960s (e.g., Popenoe, 1993), children remain a very important source of support for aging parents (Hareven, 1994; Van der Pas et al., 2007; Wolff & Kasper, 2006). The social support that children provide can buffer the negative mental health impact of stressful events that many people experience in later life, such as the onset of disability (Taylor & Lynch, 2004). Also, current cohorts of older people are not less likely than preceding cohorts to have frequent contacts with adult children (Steinbach, Mahne, Klaus, & Hank, 2019; Treas & Gubernskaya, 2012). Older persons with fewer children may be at greater risk of social isolation (Marcil-Gratton & Légaré, 1992), which is harmful for mental health (Hawton et al., 2011). Children may be particularly important for the mental health of older women, as mothers are more likely than fathers to keep in frequent contact with their children (Greenwell & Bengtson, 1997; Steinbach et al., 2019; Van der Pas et al., 2007) and to rely on them for emotional support (Antonucci & Akiyama, 1987; Van der Pas et al., 2007).

Although null-results have also been reported (Spence, 2008; Tosi & Grundy, 2019), studies tend to show that a higher number of children is associated with better mental health for older mothers. Drawing on Norwegian population register data, Kravdal et al. (2017) showed that, after controlling for age at first birth, women with three or more children were less likely to use anti-depressant medication than their counterparts with two children.

Grundy et al. (2019) found that mothers of four or more children had fewer depressive symptoms than mothers of two in a group of five Eastern-European countries included in the Generation and Gender Surveys, but they did not find a similar difference in a group of four Western-European countries. Henretta et al.'s (2008) analyses of US Health and Retirement Study data showed that, after controlling for education, having more children was associated with fewer depressive symptoms among women aged 51-61. However, in a smaller sample of the British National Survey of Health and Development the authors did not find a similar effect.

As authors typically acknowledge, caution is called for when interpreting the results of most observational studies on the association between number of children and mothers' mental health in later life. This is because the presented results may be prone to bias due to unobserved confounders or reverse causality. Completed fertility is associated with a host of factors, including the timing of the transition to parenthood, educational attainment, religiosity, family background and socio-economic position (Hayford & Morgan, 2008; Isen & Stevenson, 2010; Kravdal & Rindfuss, 2008). Many of these factors may also have an impact on mental health in later-life, and failure to account for any confounding variable of this kind will bias the estimated effect of additional children on older mothers' mental health. Moreover, Mencarini et al. (2018) have recently shown that people with higher life satisfaction went on to have more children in several developed countries. This suggests that the causal link between high fertility and mental health may be bi-directional or even in the opposite direction of what is typically assumed.

In an innovative attempt to avoid potential bias due to confounding and reverse causality, Kruk and Reinhold (2014) adopted an instrumental variable approach that exploited people's preference for mixed-sex offspring and twin births to estimate the causal effect of additional children on depressive symptoms among older parents in Europe. Drawing on data from the Survey of Health, Ageing and Retirement in Europe, they found no evidence that additional children were protective against depressive symptoms in 13 European countries. Higher fertility due to twin births was even found to have adverse mental health implications.

It should be noted that Kruk and Reinhold's sample consisted of people aged 50 and older. The fact that their respondents were on average only 65 years old could have influenced results, because research has shown that networks of older persons become more centred on close relatives with increasing age (Cornwell, Laumann, & Schumm, 2008; Van Tilburg, 1998). The focus of the current study is on American mothers aged 65 and older. With an average age of 75, the sample to be studied here is thus likely to have fewer and weaker ties with persons outside the family than the younger sample studied by Kruk and Reinhold. This may make their mental health more strongly dependent on ties with adult children. Therefore, the hypothesis to be tested in the current study is that there is a protective causal effect of additional children on mothers' mental health in later life.

## Data and methods

### Data

This study draws on data from the National Health and Aging Trends Study (NHATS, see www.nhatsdata.org) (Kasper & Freedman, 2017), a panel study of a nationally representative sample of Medicare beneficiaries aged 65 and older in the United States. NHATS is funded by the National Institute on Aging and collected by the Bloomberg School of Public Health at Johns Hopkins University. Wave 1 data were collected in 2011 and follow-up information was collected annually. In 2015, when data collection for Wave 5 took place, a refreshment sample was added. Baseline response rates were 71% for the 2011 sample and 63% for the 2015 refreshment sample (DeMatteis, Freedman, & Kasper, 2016; Montaquila, Freedman, Edwards, & Kasper, 2012). After applying the supplied analytical weights to adjust for the designed oversampling of particular groups of older people and for bias arising from systematic non-response, the Wave 1 sample was representative for the population aged 65 and older in 2011 and the Wave 5 sample for the 65+ population in 2015.

In the current study, Wave 1 and Wave 5 data were pooled. The sample used was restricted to white female main respondents aged 65 and older with at least two children (4,275 observations nested in 3,180 women). The restriction to women of white ethnicity was due to the lack of preference for mixed-sex offspring among African Americans noted both in earlier research (Tian & Morgan, 2015) and in preliminary analyses of the data used here. As explained in further detail later, the feasibility of the current study's analytical approach hinges on such a preference. Observations with missing information on the birth year of at least one child (n=345) or on the outcome measure (n=27) were dropped. Also, mothers whose second child was not a singleton (n=45) were excluded, because for these respondents the instrument used could not be properly coded. The exclusions listed here resulted in a final analytical sample of 3,858 observations nested in 2,856 women. All models were estimated with robust standard errors to account for the nested nature of the data. As a robustness check, all analyses were repeated whereby one randomly chosen observation was dropped from the sample for all 1,002 respondents observed in both Wave 1 and Wave 5. Results of these analyses did not differ substantially from the results reported here (for full results, see Appendix A in the online supplemental material).

**Measures**

Mental health was measured with the four-item version of the Patient Health Questionnaire (PHQ-4). The PHQ-4 is a screener for depression and anxiety disorders that has been validated in clinical and general population samples (Kroenke, Spitzer, Williams, & Löwe, 2009; Löwe et al., 2010). Respondents were asked how often in the last month they had experienced two symptoms indicative of depressed mood ("had little interest or pleasure in doing things"; "felt down, depressed, or hopeless") and two symptoms indicative of anxiety ("felt nervous, anxious, or on edge"; "been unable to stop or control worrying"). For all four items, response categories ranged from 0 ("not at all") to 3 ("nearly every day"). Scores on the four items were summed into an internally consistent scale ranging from 0 to 12 (Cronbach's $\alpha$ = .73). Given that the scale was strongly positively skewed and zero-inflated, scores were dichotomized. Respondents with a PHQ-4 score of 3 or higher were coded as having suboptimal mental health. The threshold of 3 has been recommended to distinguish people with elevated depression and anxiety symptomatology from people with normal scores (Kroenke et al., 2009).

The explanatory variable of interest is number of biological children. The instrument used to predict this plausibly endogenous variable is the sex composition of respondents' two firstborn children. This variable is dichotomous and distinguishes between mothers of whom the two firstborn children have identical sexes (daughter-daughter or son-son) and mothers of whom the two firstborn children have different sexes (daughter-son or son-daughter). Given that it is randomly assigned by nature whether or not the sex of the second child is identical to the sex of the first, this instrument is exogenous. Consequently, there is no need to include additional covariates to account for potential confounding.

**Analytical approach**

The current study adopts an instrumental variable approach to avoid potential bias due to confounding or reverse causality (Martens, Pestman, De Boer, Belitser, & Klungel, 2006). Following Kruk and Reinhold (2014), the instrumental variable approach taken in the current study exploits the preference for mixed-sex offspring that has been noted in the United States, in particular among white women (Tian & Morgan, 2015). Due to this preference, mothers of two children are more likely to have a third child when the two firstborn children are either both daughters or both sons than when they are a daughter and a son. Consequently, the former group of mothers has higher completed fertility than the latter. Given that this fertility difference is attributable to whether the sex of the second child is different from the sex of the first child, it is effectively randomly assigned by nature and therefore exogenous. In the approach taken in the current study, this exogenous fertility difference is used to estimate the causal effect on mental health of having more children. This is done in a two-stage approach. Equation 1 presents the first stage model:

$$X_i = \alpha_0 + \alpha_1 Z_i + \varepsilon_i \tag{1}$$

In the first stage, the exogenous instrument *Z*, i.e. whether or not the two firstborn children of mother *i* are of the same sex, is used to predict mother *i*'s number of children *X*. Estimate $\alpha_1$ denotes the difference in the total number of children between mothers whose two firstborn children are of the same sex and mothers whose two firstborn children are of different sexes. As shown in equation 2, this exogenous fertility difference is subsequently used to estimate the causal effect of having more children on mental health.

$$\Pr(Y_i = 1) = \Phi(\beta_0 + \beta_1 \hat{X}_i + u_i) \qquad (2)$$

In the second stage probit regression, the probability of suboptimal mental health $Y$ for mother $i$ is regressed on $\hat{X}$, i.e. the number of children as predicted in the first stage.

## Results

<Table 1 here>

Descriptive statistics for the analytical sample are provided in Table 1. On average, the mothers included in the sample had 3 children. As was to be expected due to the preference for mixed-sex offspring noted in earlier research (Tian & Morgan, 2015), women whose two firstborn children were both sons or both daughters had a higher total number of children than their counterparts who had both a son and a daughter among their two firstborn children. The former group also less often had suboptimal mental health as indicated by a score of 3 or higher on the PHQ-4 scale.

Table 2 shows the results of the naïve probit and instrumental variable probit analyses predicting suboptimal mental health. The naïve probit model did not show any association between number of children and the risk of suboptimal mental health. It is important to note, however, that this model assumes that number of children is an exogenous variable. This may be an invalid assumption, given the aforementioned social patterning of fertility patterns and the effect that mental health may have on fertility. Therefore, true effects of number of children may be suppressed in this model.

<Table 2 here>

The first stage of the instrumental variable model shows that women whose two firstborn children were both sons or both daughters had on average 0.2 children more than their counterparts with both a son and a daughter among their two firstborn children. This difference was statistically significant and the F-statistic greatly exceeded 10, indicating that the instrument used in the first stage (sex composition of the two firstborn children) strongly predicted the instrumented variable (total number of children) ($F(1, 3856) = 23.3$, $p < .001$). The exogenous fertility difference between both groups of mothers was used in the second stage to estimate the causal effect of number of children on mental health. As hypothesized, the model showed that an additional child had a statistically significant protective effect against suboptimal mental health. A Wald test indicated the estimates of the instrumental variable probit model were to be preferred over those of the naïve probit model, because the assumption of exogeneity underlying the latter model did not hold ($\chi^2(1, n=3,858) = 4.04$, $p < .05$).

<Figure 1 here>

To facilitate an easier interpretation of the magnitude of the instrumental variable probit results, the predicted probability of suboptimal mental health for mothers with different numbers of children are presented in Figure 1. The predicted probabilities shown in this figure can be interpreted as the estimated average risk of suboptimal mental health for older mothers if they all were to have two, three, four or five children, respectively.

## Discussion

The current study addressed the question whether there are mental health benefits to having more children for older white mothers in the United States. With the notable exception of a study by Kruk and Reinhold (2014), earlier work on the association between number of children and mental health has tended to adopt a descriptive approach. Such an approach comes with the risk that findings are biased due to selection or reverse causality. In contrast to the commonly taken descriptive approach, the current study adopted an instrumental variable approach exploiting the known preference for mixed-sex offspring to estimate the causal effect of additional children on mental health. Results indicate that additional children reduce the risk of suboptimal mental health among white mothers aged 65 and older.

Adopting an analytical approach largely similar to the one taken in the current study, Kruk and Reinhold (2014) did not find any evidence that having more children was protective against depressive symptoms among mothers aged 50 and older in 13 European countries.

As argued earlier, it is important to note that the sample of the current study was considerably older than the sample studied by Kruk and Reinhold. Given that older people's social networks become more strongly centred on close family with increasing age (Cornwell et al., 2008; Van Tilburg, 1998), this may explain why Kruk and Reinhold's analyses did not show the beneficial mental health effects of additional children reported here. Future research is needed to assess whether there is also a causal protective effect of additional children on the mental health of mothers in the older age groups in European contexts.

An alternative explanation for the differences between the findings of Kruk and Reinhold and those of the current study could be that the strong norms of family obligation in the United States (Cooney & Dykstra, 2011) and the limited security that the country's pension and long-term care systems provide (Gage, 2014; Hinrichs & Lynch, 2010; OECD, 2017) may make older women depend relatively strongly on offspring for their welfare. Consequently, children may be more important for the mental health of older women in the United States than in some of the European countries studied by Kruk and Reinhold. It should also be noted that Kruk and Reinhold used the EURO-D scale to measure depressive symptoms, as opposed to the PHQ-4 used to measure mental health in the current study. Courtin et al. (2015) have noted that several known risk factors for depression were less strongly associated with EURO-D scores than with scores on the commonly used CES-D scale. This could possibly also explain Kruk and Reinhold's null results.

In contrast to Kruk and Reinhold's study, the current study did not use twin births as an alternative instrument for total number of children. This choice was made because twin births are not randomly assigned. They are associated with factors such as maternal age, weight, and height (Basso, Nohr, Christensen, & Olsen, 2004; Beemsterboer et al., 2006) that may also be associated with later-life mental health. Therefore, twin births arguably do not constitute a valid instrumental variable.

Several sensitivity analyses were performed to check the robustness of the reported findings (see online supplemental material, Appendix A-D). Results of analyses performed on a sample with only one randomly chosen observation for each of 1,002 respondents observed in both Wave 1 and Wave 5 (Appendix A) and of unweighted data (Appendix B) were very similar to those presented in Table 2. In the model presented in Appendix C an alternative instrument specification was used that distinguished mothers with two sons and mothers with two daughters as their firstborn children. However, this specification did not perform better than the more parsimonious specification used in the main analyses and the estimated causal mental health effect of additional children did not change substantially. This is not surprising, given that the total number of children of mothers whose two firstborn children both were daughters did not differ significantly from their counterparts whose two firstborns both were sons. Finally, it is plausible that the number of daughters, rather than the number of children, are important for older mothers' mental health. This is because mothers tend to prefer the caregiving of daughters (Suitor & Pillemer, 2006). Additional analyses were therefore performed exploiting the sex of the firstborn child as an exogenous instrument (cf. Oswald & Powdthavee, 2010). This variable was obviously strongly associated with total number of daughters, but not associated with total number of children in the sample analyzed here. The results did not show a causal effect of number of daughters on mothers' mental health (See Appendix D). There is thus no evidence that additional daughters are more important than additional sons.

Some limitations of the current study should be considered. First, a key assumption underlying the instrumental variable approach taken here is that the instrument (here: sex composition of two firstborn children) only affects the outcome via the instrumented variable (here: total number of children). However, given the known preference for mixed sex offspring (Tian & Morgan, 2015), having both a son and a daughter among the two firstborn children is likely to be more in line with mother's preferences than having two sons or two daughters, and therefore possibly beneficial for mental health. This could imply that the protective effect of additional children against suboptimal mental health might be underestimated in the current study.

Second, the results reported here can only be generalized to white mothers with at least two children. This is because only respondents with two or more children could be included in the analytical sample given the analytical approach adopted. Moreover, fathers and African American mothers could not be included, because the instrument used was insufficiently predictive of number of children in these groups. Although preliminary analyses showed that fathers whose two firstborn children were both sons or both daughters had slightly more children than their counterparts with both a son and a daughter among their two firstborn children, the sex composition of the two firstborn children was insufficiently predictive of total number of children for the instrument to meet conventional strength standards ($F(1, 2630) = 0.9$, $p = .349$). In addition to the smaller sample size of the father sample (2,632 observations nested in 1,970 fathers), it should be considered that the men in the father sample tended to have children with women from later birth cohorts than the women included in the mother sample, and that higher order births were less common for women in these later birth cohorts (Kirmeyer & Hamilton, 2011). African American mothers included in NHATS did not have a higher number of children if the two firstborn children were both sons or both daughters as opposed to a daughter and a son. This is consistent with findings from earlier research (Tian & Morgan, 2015).

Also, it should be noted that the instrument used does not predict unplanned additional children, given that it exploits American mothers' stronger disposition to plan to have an additional child when their two firstborn children have the same sex. This comes with the caveat that the causal effect of additional children estimated in the instrumental variable model presented here is not valid for additional unplanned pregnancies, which may have negative mental health implications that last until old age (Herd, Higgins, Sicinski, & Merkurieva, 2016). Finally, the dichotomous nature of the instrument used made it unfeasible to model potential non-linearity in the effect of additional children. It could therefore not be tested whether there were diminishing mental health returns to higher-order additional children.

Many studies have shown that children remain a very important source of support and social contact for aging parents (e.g., Hareven, 1994; Steinbach et al., 2019; Treas & Gubernskaya, 2012; Wolff & Kasper, 2006). In line with these findings, the current study showed that children are beneficial for older women's mental health. The finding that having an additional child had, on average, a causal protective effect against elevated depression and anxiety symptomatology in later life for white mothers should be considered in the light of the decline in the likelihood of higher-order births that – albeit to a smaller extent than in European contexts – has been noted in the United States (Devolder et al., 2002; cf. Kirmeyer & Hamilton, 2011). The results reported here suggest that this demographic development may come with increased risks of suboptimal mental health for new cohorts of older women.

## References

Antonucci, T. C., & Akiyama, H. (1987). An examination of sex differences in social support among older men and women. *Sex Roles*, *17*(11), 737–749. https://doi.org/10.1007/BF00287685

Basso, O., Nohr, E. A., Christensen, K., & Olsen, J. (2004). Risk of twinning as a function of maternal height and body mass index. *JAMA*, *291*(13), 1564–1566. https://doi.org/10.1001/jama.291.13.1564-c

Beemsterboer, S. N., Homburg, R., Gorter, N. A., Schats, R., Hompes, P. G. A., & Lambalk, C. B. (2006). The paradox of declining fertility but increasing twinning rates with advancing maternal age. *Human Reproduction*, *21*(6), 1531–1532. https://doi.org/10.1093/humrep/del009

Bures, R. M., Koropeckyj-Cox, T., & Loree, M. (2009). Childlessness, parenthood, and depressive symptoms among middle-aged and older adults. *Journal of Family Issues*, *30*(5), 670–687. https://doi.org/10.1177/0192513X08331114

Cherlin, A. J. (2010). Demographic trends in the United States: A review of research in the 2000s. *Journal of Marriage and Family*, *72*(3), 403–419. https://doi.org/10.1111/j.1741-3737.2010.00710.x

Cooney, T. M., & Dykstra, P. A. (2011). Family obligations and support behaviour: A United States–Netherlands comparison. *Ageing and Society*, *31*(6), 1026–1050. https://doi.org/10.1017/S0144686X10001339

Cornwell, B., Laumann, E. O., & Schumm, L. P. (2008). The social connectedness of older adults: A national profile. *American Sociological Review*, *73*(2), 185–203. https://doi.org/10.1177/000312240807300201

Courtin, E., Knapp, M., Grundy, E., & Avendano-Pabon, M. (2015). Are different measures of depressive symptoms in old age comparable? An analysis of the CES-D and Euro-D scales in 13 countries. *International Journal of Methods in Psychiatric Research*, *24*(4), 287–304. https://doi.org/10.1002/mpr.1489

DeMatteis, J., Freedman, V. A., & Kasper, J. D. (2016). *National Health and Aging Trends Study round 5 sample design and selection. NHATS technical paper #16*. Baltimore: Johns Hopkins University School of Public Health.

Devolder, D., González, M. J., & Gavino, K. (2002). *Trends in fertility by parity in Europe, USA and Canada. CED Working Paper 203*. Barcelona: CED Centre for Demographic Studies.

Gage, B. (2014). Building a long-term services and supports delivery system that works. *Public Policy & Aging Report*, *24*(2), 56–59. https://doi.org/10.1093/ppar/pru011

Greenwell, L., & Bengtson, V. L. (1997). Geographic distance and contact between middle-aged children and their parents: The effects of social class over 20 years. *The Journals of Gerontology: Series B*, *52B*(1), S13–S26. https://doi.org/10.1093/geronb/52B.1.S13

Grundy, E., Van den Broek, T., & Keenan, K. (2019). Number of children, partnership status, and later-life depression in Eastern and Western Europe. *The Journals of Gerontology, Series B*, *74*(2), 353–363. https://doi.org/10.1093/geronb/gbx050

Hareven, T. K. (1994). Aging and generational relations: A historical and life course perspective. *Annual Review of Sociology*, *20*(1), 437–461.

https://doi.org/10.1146/annurev.so.20.080194.002253

Hawton, A., Green, C., Dickens, A. P., Richards, S. H., Taylor, R. S., Edwards, R., … Campbell, J. L. (2011). The impact of social isolation on the health status and health-related quality of life of older people. *Quality of Life Research*, *20*(1), 57–67. https://doi.org/10.1007/s11136-010-9717-2

Hayford, S. R., & Morgan, S. P. (2008). Religiosity and fertility in the United States: The role of fertility intentions. *Social Forces*, *86*(3), 1163–1188. https://doi.org/10.1353/sof.0.0000

Henretta, J. C., Grundy, E. M. D., Okell, L. C., & Wadsworth, M. E. J. (2008). Early motherhood and mental health in midlife: A study of British and American cohorts. *Aging & Mental Health*, *12*(5), 605–614. https://doi.org/10.1080/13607860802343084

Herd, P., Higgins, J., Sicinski, K., & Merkurieva, I. (2016). The implications of unintended pregnancies for mental health in later life. *American Journal of Public Health*, *106*(3), 421–429. https://doi.org/10.2105/AJPH.2015.302973

Hinrichs, K., & Lynch, J. F. (2010). Old-age pensions. In F. G. Castles, S. Leibfried, J. Lewis, H. Obinger, & C. Pierson (Eds.), *The Oxford handbook of the welfare state* (pp. 353–366). Oxford: Oxford University Press.

Huijts, T., Kraaykamp, G., & Subramanian, S. V. (2013). Childlessness and psychological well-being in context: A multilevel study on 24 European Countries. *European Sociological Review*, *29*(1), 32–47. https://doi.org/10.1093/esr/jcr037

Isen, A., & Stevenson, B. (2010). *Women's education and family behavior: Trends in* *marriage, divorce and fertility. NBER working paper 15725*. Cambridge, MA: National Bureau of Economic Research.

Kasper, J. D., & Freedman, V. A. (2017). *National Health and Aging Trends Study user guide: Rounds 1-6 final release*. Baltimore: Johns Hopkins University School of Public Health.

Kirmeyer, S. E., & Hamilton, B. E. (2011). *Childbearing differences among three generations of U.S. women. NCHS data brief 68*. Hyattsville MD: National Center for Health Statistics.

Koropeckyj-Cox, T. (1998). Loneliness and depression in middle and old age: Are the childless more vulnerable? *The Journals of Gerontology Series B: Psychological Sciences and Social Sciences*, *53B*(6), S303–S312. https://doi.org/10.1093/geronb/53B.6.S303

Kravdal, Ø., Grundy, E., & Skirbekk, V. (2017). Fertility history and use of antidepressant medication in late mid-life: A register-based analysis of Norwegian women and men. *Aging & Mental Health*, *21*(5), 477–486. https://doi.org/10.1080/13607863.2015.1118010

Kravdal, Ø., & Rindfuss, R. R. (2008). Changing relationships between education and fertility: A study of women and men born 1940 to 1964. *American Sociological Review*, *73*(5), 854–873. https://doi.org/10.1177/000312240807300508

Kroenke, K., Spitzer, R. L., Williams, J. B. W., & Löwe, B. (2009). An ultra-brief screening scale for anxiety and depression: The PHQ–4. *Psychosomatics*, *50*(6), 613–621. https://doi.org/10.1016/S0033-3182(09)70864-3

Kruk, K. E., & Reinhold, S. (2014). The effect of children on depression in old age. *Social Science and Medicine*, *100*, 1–11. https://doi.org/10.1016/j.socscimed.2013.09.003

Litwak, E. (1985). *Helping the elderly: The complementary roles of informal networks and formal systems*. New York: Guilford Press.

Löwe, B., Wahl, I., Rose, M., Spitzer, C., Glaesmer, H., Wingenfeld, K., … Brähler, E. (2010). A 4-item measure of depression and anxiety: Validation and standardization of the Patient Health Questionnaire-4 (PHQ-4) in the general population. *Journal of Affective Disorders*, *122*(1), 86–95. https://doi.org/10.1016/j.jad.2009.06.019

Marcil-Gratton, N., & Légaré, J. (1992). Will reduced fertility lead to greater isolation in old age for tomorrow's elderly? *Canadian Journal on Aging*, *11*(1), 54–71. https://doi.org/10.1017/S0714980800014525

Martens, E. P., Pestman, W. R., De Boer, A., Belitser, S. V, & Klungel, O. H. (2006). Instrumental variables: Application and limitations. *Epidemiology*, *17*(3), 260–267. https://doi.org/10.1097/01.ede.0000215160.88317.cb

Mencarini, L., Vignoli, D., Zeydanli, T., & Kim, J. (2018). Life satisfaction favors reproduction. The universal positive effect of life satisfaction on childbearing in contemporary low fertility countries. *PLOS ONE*, *13*(12), e0206202. https://doi.org/10.1371/journal.pone.0206202

Messeri, P., Silverstein, M., & Litwak, E. (1993). Choosing optimal support groups : A review and reformulation. *Journal of Health and Social Behavior*, *34*(2), 122–137. https://doi.org/10.2307/2137239

Montaquila, J., Freedman, V. A., Edwards, B., & Kasper, J. D. (2012). *National Health and Aging Trends Study round 1 sample design and selection. NHATS technical paper #1*.

Baltimore: Johns Hopkins University School of Public Health.

OECD. (2017). *Pensions at a glance 2017. OECD and G20 indicators*. Paris: OECD Publishing.

Oswald, A. J., & Powdthavee, N. (2010). Daughters and left-wing voting. *The Review of Economics and Statistics*, *92*(2), 213–227. https://doi.org/10.1162/rest.2010.11436

Popenoe, D. (1993). American family decline, 1960-1990: A review and appraisal. *Journal of Marriage and Family*, *55*(3), 527–542. https://doi.org/10.2307/353333

Ryan, L. H., Smith, J., Antonucci, T. C., & Jackson, J. S. (2012). Cohort differences in the availability of informal caregivers: Are the boomers at risk? *The Gerontologist*, *52*(2), 177–188. https://doi.org/10.1093/geront/gnr142

Spence, N. J. (2008). The long-term consequences of childbearing: Physical and psychological well-being of mothers in later life. *Research on Aging*, *30*(6), 722–751. https://doi.org/10.1177/0164027508322575

Steinbach, A., Mahne, K., Klaus, D., & Hank, K. (2019). Stability and change in intergenerational family relations across two decades: Findings from the German Ageing Survey, 1996–2014. *The Journals of Gerontology: Series B*. https://doi.org/10.1093/geronb/gbz027

Suitor, J. J., & Pillemer, K. (2006). Choosing daughters: Exploring why mothers favor adult daughters over sons. *Sociological Perspectives*, *49*(2), 139–161. https://doi.org/10.1525/sop.2006.49.2.139

Taylor, M. G., & Lynch, S. M. (2004). Trajectories of impairment, social support, and depressive symptoms in later life. *The Journals of Gerontology Series B: Psychological*

*Sciences and Social Sciences*, *59*(4), S238–S246. https://doi.org/10.1093/geronb/59.4.S238

Tian, F. F., & Morgan, S. P. (2015). Gender composition of children and the third birth in the United States. *Journal of Marriage and Family*, *77*(5), 1157–1165. https://doi.org/10.1111/jomf.12218

Tosi, M., & Grundy, E. (2019). Intergenerational contacts and depressive symptoms among older parents in Eastern Europe. *Aging & Mental Health*, *23*(6), 686–692. https://doi.org/10.1080/13607863.2018.1442412

Treas, J., & Gubernskaya, Z. (2012). Farewell to moms? Maternal contact for seven countries in 1986 and 2001. *Journal of Marriage and Family*, *74*(2), 297–311. https://doi.org/10.1111/j.1741-3737.2012.00956.x

Van der Pas, S., Van Tilburg, T., & Knipscheer, K. (2007). Changes in contact and support within intergenerational relationships in the Netherlands: A cohort and time-sequential perspective. *Advances in Life Course Research*, *12*(Supplement C), 243–274. https://doi.org/10.1016/S1040-2608(07)12009-8

Van Tilburg, T. (1998). Losing and gaining in old age: Changes in personal network size and social support in a four-year longitudinal study. *The Journals of Gerontology: Series B*, *53B*(6), S313–S323. https://doi.org/10.1093/geronb/53B.6.S313

Wolff, J. L., & Kasper, J. D. (2006). Caregivers of frail elders: Updating a national profile. *The Gerontologist*, *46*(3), 344–356. https://doi.org/10.1093/geront/46.3.344

Zhang, Z., & Hayward, M. D. (2001). Childlessness and the psychological well-being of older persons. *The Journals of Gerontology Series B: Psychological Sciences and Social Sciences*, *56B*(5), S311–S320. https://doi.org/10.1093/geronb/56.5.S311

Table 1. Sample characteristics; means and percentages

| | All mothers | Two firstborn children of identical sex | Two firstborn children of different sexes |
|---|---|---|---|
| Number of children | 3.0 | 3.2 | 2.9 |
| (Standard deviation) | (1.3) | (1.4) | (1.2) |
| Suboptimal mental health (PHQ-4 >= 3) | 29.0% | 27.3% | 30.6% |
| Age | 75.3 | 75.5 | 75.2 |
| (Standard deviation) | (7.2) | (7.2) | (7.1) |
| | | | |
| Number of observations | 3,858 | 1,904 | 1,954 |
| Number of respondents | 2,856 | 1,439 | 1,417 |

Table 2. Results of naïve probit and instrumental variable (IV) probit regression models predicting suboptimal mental health (PHQ-4 >=3)

| | Naïve probit | | IV probit | | | |
|---|---|---|---|---|---|---|
| | | | First stage | | Second stage | |
| | Coeff. | (SE) | Coeff. | (SE) | Coeff. | (SE) |
| Number of children | 0.001 | (0.018) | | | -0.378* | (0.157) |
| Sex composition two firstborn children: | | | | | | |
| Identical | | | Ref. | | | |
| Different | | | -0.217*** | (0.045) | | |
| Constant | -0.556*** | (0.060) | 3.161*** | (0.034) | 0.673 | (0.544) |

Notes: Data are from the National Health and Aging Trends Study, Wave 1 and Wave 5; n=3,858;
Weighted; Robust standard errors
* $p < .05$, ** $p < .01$, *** $p < .001$

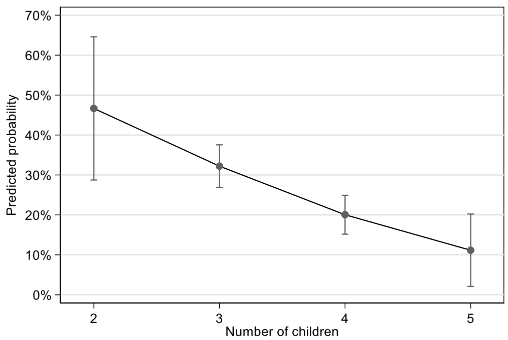


Figure 1. Risk of suboptimal mental health (PHQ-4 >=3) by number of children

Appendix A. Results of naïve probit and instrumental variable (IV) probit regression models predicting suboptimal mental health (PHQ-4 >=3) with 1 observation per respondent only

| | Naïve probit | | IV probit | | | |
|---|---|---|---|---|---|---|
| | | | First stage | | Second stage | |
| | Coeff. | (SE) | Coeff. | (SE) | Coeff. | (SE) |
| Number of children | 0.013 | (0.021) | | | -0.400* | (0.203) |
| Sex composition two firstborn children: | | | | | | |
| Identical | | | Ref. | | | |
| Different | | | -0.186*** | (0.051) | | |
| Constant | -0.570*** | (0.070) | 3.112*** | (0.038) | 0.756 | (0.698) |

Notes: Data are from the National Health and Aging Trends Study, Wave 1 and Wave 5; n=2,856;
Weighted
* $p$ < .05, ** $p$ < .01, *** $p$ < .001

Appendix B. Results of naïve probit and instrumental variable (IV) probit regression models predicting suboptimal mental health (PHQ-4 >=3) (Not weighted)

| | Naïve probit | | IV probit | | | |
|---|---|---|---|---|---|---|
| | | | First stage | | Second stage | |
| | Coeff. | (SE) | Coeff. | (SE) | Coeff. | (SE) |
| Number of children | -0.017 | (0.016) | | | -0.357* | (0.149) |
| Sex composition two firstborn children: | | | | | | |
| Identical | | | Ref. | | | |
| Different | | | -0.210*** | (0.044) | | |
| Constant | -0.494*** | (0.053) | 3.242*** | (0.033) | 0.636 | (0.526) |

Notes: Data are from the National Health and Aging Trends Study, Wave 1 and Wave 5; n=3,858;
Not weighted; Robust standard errors
* $p < .05$, ** $p < .01$, *** $p < .001$

Appendix C. Results of naïve probit and instrumental variable (IV) probit regression models predicting suboptimal mental health (PHQ-4 >=3)

| | Naïve probit | | IV probit | | | |
|---|---|---|---|---|---|---|
| | | | First stage | | Second stage | |
| | Coeff. | (SE) | Coeff. | (SE) | Coeff. | (SE) |
| Number of children | 0.001 | (0.018) | | | -0.413* | (0.163) |
| Sex composition two firstborn children: | | | | | | |
| Daughter and son | | | Ref. | | | |
| Two sons | | | 0.237*** | (0.052) | | |
| Two daughters | | | 0.185** | (0.063) | | |
| Constant | -0.556*** | (0.060) | 2.946*** | (0.030) | 0.793 | (0.571) |

Notes: Data are from the National Health and Aging Trends Study, Wave 1 and Wave 5; n=3,858;
Weighted; Robust standard errors
* *p* < .05, ** *p* < .01, *** *p* < .001

Appendix D. Results of naïve probit and instrumental variable (IV) probit regression models predicting suboptimal mental health (PHQ-4 >=3)

| | Naïve probit | | IV probit | | | |
|---|---|---|---|---|---|---|
| | | | First stage | | Second stage | |
| | Coeff. | (SE) | Coeff. | (SE) | Coeff. | (SE) |
| Number of daughters | 0.023 | (0.022) | | | 0.046 | (0.048) |
| Sex firstborn child: | | | | | | |
| Son | | | Ref. | | | |
| Daughter | | | 0.989*** | (0.034) | | |
| Constant | -0.588*** | (0.042) | 1.017*** | (0.023) | -0.623*** | (0.076) |

Notes: Data are from the National Health and Aging Trends Study, Wave 1 and Wave 5; n=3,858;
Weighted; Robust standard errors
* $p < .05$, ** $p < .01$, *** $p < .001$