SHARE ERIC
SURVEY OF HEALTH, AGEING AND RETIREMENT IN EUROPE

SHARE WORKING PAPER SERIES

Family size and vaccination among older individuals: The case of COVID-19 vaccine

Eric Bonsang, Elizaveta Pronkina

Working Paper Series 76-2021
DOI: 10.17617/2.3361025

SHARE-ERIC | Amalienstr. 33 | 80799 Munich | Germany | share-eric.eu

MAX PLANCK INSTITUTE FOR SOCIAL LAW AND SOCIAL POLICY
Munich Center for the Economics of Aging

This project has received funding from the European Union under grant agreements VS/2019/0332, VS/2020/0313 and the European Union's Horizon 2020 research and innovation programme under grant agreements No 870628, No 101015924.

SPONSORED BY THE Federal Ministry of Education and Research

Supported by the NIH National Institute on Aging

About the SHARE Working Paper Series

The series is designed to provide a timely discussion of results based on SHARE data within the SHARE family, i.e., members of the SHARE Country Teams, Area Coordination Teams and other SHARE bodies. The papers are not peer reviewed; the authors are solely responsible for the scientific content and the graphical layout of their submissions. The respective Country Team Leaders and Area Coordinators are encouraged to look over the submissions by their team members.

The publisher (SHARE ERIC) checks working papers in this series for formal issues such as proper acknowledgements to the funders of SHARE. The publisher takes no responsibility for the scientific content of the paper.

Acknowledgements

Research in this article is a part of the EU Horizon 2020 SHARE-COVID19 project (Grant agreement ID: 101015924).

This publication is based on preliminary SHARE wave 9 COVID-19 Survey 2 release 0 data (Börsch-Supan 2021). Therefore, the analyses, conclusions and results are preliminary. Please see Scherpenzeel et al. (2020) for methodological details. In addition, this paper uses data from SHARE Waves 1, 2, 4, 5, 6, 7 and 8 (DOIs: 10.6103/SHARE.w1.710, 10.6103/SHARE.w2.710, 10.6103/SHARE.w3.710, 10.6103/SHARE.w4.710, 10.6103/SHARE.w5.710, 10.6103/SHARE.w6.710, 10.6103/SHARE.w7.711, 10.6103/SHARE.w8caintd.100, 10.6103/SHARE.w8ca.100, 10.6103/SHARE.w8.100), see Börsch-Supan et al. (2013) for methodological details. The SHARE data collection has been funded by the European Commission, DG RTD through FP5 (QLK6-CT-2001-00360), FP6 (SHARE-I3: RII-CT-2006-062193, COMPARE: CIT5-CT-2005-028857, SHARELIFE: CIT4-CT-2006-028812), FP7 (SHARE-PREP: GA N°211909, SHARE-LEAP: GA N°227822, SHARE M4: GA N°261982, DASISH: GA N°283646) and Horizon 2020 (SHARE-DEV3: GA N°676536, SHARE-COHESION: GA N°870628, SERISS: GA N°654221, SSHOC: GA N°823782) and by DG Employment, Social Affairs & Inclusion through VS 2015/0195, VS 2016/0135, VS 2018/0285, VS 2019/0332, and VS 2020/0313. Additional funding from the German Ministry of Education and Research, the Max Planck Society for the Advancement of Science, the U.S. National Institute on Aging (U01_AG09740-13S2, P01_AG005842, P01_AG08291, P30_AG12815, R21_AG025169, Y1-AG-4553-01, IAG_BSR06-11, OGHA_04-064, HHSN271201300071C, RAG052527A) and from various national funding sources is gratefully acknowledged (see www.share-project.org).

Family size and vaccination among older individuals: The case of COVID-19 vaccine[1]

Eric Bonsang and Elizaveta Pronkina

Université Paris-Dauphine, PSL Research University, CNRS, IRD, LEDa

December 17, 2021

Abstract

While vaccination is generally considered an efficient way to protect against transmissible diseases, vaccine hesitancy is still widespread in many countries. In this paper, we investigate how individual-specific factors -the number of children- affect the probability of getting the vaccine against COVID-19. To answer this research question, we focus on the population aged 50 year-old or older as the risk of developing severe symptoms is the highest among them and use the Survey of Health, Ageing and Retirement in Europe (SHARE) Corona wave conducted in Summer 2021. To identify the effect of family size on vaccination, we exploit an exogenous variation in the probability to have more than two children due to the sex composition of the two firstborns. We document that having more than two children increases the probability of getting the COVID-19 vaccine among older individuals. This impact is economically and statistically significant. We propose a potential mechanism behind this result - the higher probability of being exposed to the disease proxied by knowing someone who tested positive with the Coronavirus or had symptoms similar to it, and by network size and number of contacts with children before the outbreak of Corona.

JEL codes: I18, I15, J13

Keywords: Vaccination, Family size, Exposure to virus, COVID-19

[1]We highly appreciate feedback from Jesús M. Carro, Juan Jose Dolado, Clémentine Garrouste, Anne Laferrère and Pedro Mira. We also would like to thank feedback from participants at the LEGOS seminar at Université Paris-Dauphine. This publication is based on preliminary SHARE wave 9 COVID-19 Survey 2 release 0 data. Therefore, the analyses, conclusions and results are preliminary. Research in this article is a part of the H2020 SHARE-COVID19 project (Grant agreement No. 101015924). For the full SHARE data use acknowledgments see Appendix.

1. Introduction

Getting a vaccine against COVID is crucial to protect oneself, decrease the spread of the disease and to return to a pre-pandemic lifestyle (WHO, November 2021).[2] Anti-COVID vaccines have been almost universally available to older individuals in Europe by Summer 2021. Yet, vaccination hesitancy is spread across many countries despite its documented benefits specifically for older individuals. What can explain within-country differences in vaccine take-up? This question is crucial to understand how to quickly reach herd immunity and to contain the spread of the disease. In this paper, we focus on one factor - family size, which is likely to affect, among other mechanisms, the differential exposure to the virus, and, in turn, may impact the vaccination decision.

We estimate the causal impact of family size on the probability of getting the COVID vaccine. Leaving aside the "quality" of children, we consider the role of the number of children. To do so, we use data from the Second Wave of Survey of Health, Ageing and Retirement in Europe (SHARE) Corona Survey conducted between June and August 2021.[3] The Corona Survey is a part of SHARE and is designed to estimate the impact of the COVID-19 pandemic on individuals above 50 years old in Europe and Israel. The second wave includes the question: *"Have you been vaccinated against Covid-19?",* that allows us to study revealed preferences about the COVID-19 vaccination. We combine questions from the SHARE Corona Survey with fertility history from the pre-pandemic survey waves to investigate the role of family links on vaccination. Thus, in our study, we focus on the elderly who were targeted as a priority group for vaccination in many countries.

To get a causal effect of the family size on the probability of getting the COVID vaccine, we use parental preferences for mixed-sex offspring as a source of exogenous variation for having more than two children in line with Angrist and Evans (1998). Accordingly, the instrument is based on the sex composition of the two firstborn biological children. Results show that the number of children has a significant positive impact on the probability of getting the COVID vaccine.

Why does family size affect the decision to get the vaccine against COVID? On the one

[2]Retrieved on November 29, 2021
https://www.who.int/news-room/questions-and-answers/item/coronavirus-disease-(covid-19)-vaccines
[3]This publication is based on preliminary SHARE wave 9 COVID-19 Survey 2 release 0 data. For the full SHARE data use acknowledgments see the Appendix.

hand, a larger family size might lead to more personal contacts (e.g. Grundy and Read, 2012) and, in turn, a higher risk of contagion. Therefore, the decision to get the COVID vaccine can be interpreted as a preventive behaviour facing a greater risk of getting sick. In line with this hypothesis, before the COVID pandemic, Oster (2018) shows that disease outbreaks increase vaccination rates and Ahituv et al. (1996) document a higher condom use across the US in response to the local prevalence of AIDS. Differences in lifestyles before and after the outbreak of Corona could also matter for vaccination: this may be the case for grandparents who were used to meeting with their grandchildren in the past and to keep with those contacts would mean a high exposure to disease. The size of the network may also increase the probability of knowing someone who had the disease. Thus, it can increase awareness about the existence and negative effects of the virus. Funk et al. (2009) develop a theoretical model to show that the first-hand experience and diffusion of knowledge within the network can contain the transmissible disease. At the local level and related to the COVID pandemic, Giulietti et al. (2021) find a higher COVID vaccination rate in places with more deaths in the UK. The number of children could also increase the probability to have the needed information for getting vaccinated: the location and commuting to the vaccine centre, updates on eligibility criteria about vaccination, booking appointments on the online websites (e.g., for anecdotal evidence see the Washington Post, 2021).[4] On the other hand, there could be trade-offs between the number of children and quality of contacts among parents and children, and larger family size might lead to less strong parent-child social ties and, in turn, fewer social interactions (e.g. Ward et al. 2009). Next, knowing someone who had a mild form of COVID may downplay the perception of the contagion severity (Kreps and Kriner, 2021; Diament et al., 2021). Moreover, the negative externalities of not getting vaccinated might be lower for those individuals who are surrounded by younger family members who are less likely to develop a serious form of the disease.

Understanding whether family size has a positive or negative effect remains an empirical question. Based on additional modules in the SHARE Corona Survey and pre-COVID waves, we document a positive and statistically significant effect of family size on vaccination decisions. We show that having more than two children increases the probability of knowing someone who had the Corona symptoms or had been tested positive by Summer 2021 but does not affect the

[4]Refer to https://www.washingtonpost.com/technology/2021/02/02/vaccine-appointment-parents-kids/

probability of knowing someone who was hospitalised. Using pre-COVID information, we show that a larger family size directly leads to 1) a larger network size; 2) having more grandchildren; 3) being a great-grandparent; and it also 4) increases the pre-pandemic total number of contacts. All these variables imply that individuals with more than two children do face a higher risk of COVID-19 exposure and a higher utility from getting vaccinated compared to those with only two.

This paper contributes to the growing literature about factors that shape behavioural responses during the COVID-19 pandemic, and specifically, the individual determinants of vaccination and exposure to disease. The closest article to ours is Giulietti et al. (2021) that shows that individual intention to get vaccinated increases with the death rate at the local level. In our study, and differently from them, we directly estimate the impact of the individual-specific factors that lead to increase in the vaccine take-up due to a higher exposure rather than its proxy at an aggregate level. Moreover, our main hypothesis is in line with several studies documenting that the self-risk perception is one of the important drivers for the vaccine take up. Caserotti et al. (2021) conducted a survey about the intention to get vaccinated before, during, and after the confinement measure in Spring 2020 in Italy. They show that a higher risk perception of the virus increases the willingness to get vaccinated. This is confirmed by Wang et al. (2020) who use survey data in China and document that the risk perception of the virus increases vaccine acceptance. Solís Arce et al. (2021) documented that personal protection from getting the disease is the primary reason for willingness to get vaccinated. In this regard, we consider the vaccination decision instead of the intention to get vaccinated, and these two might differ as it is shown by Dai et al. (2021).

Next, we also contribute to the literature about family size. So far, most articles on the causal impact of the number of children on individual outcomes exploit the instrumental variable strategy based on the multiple births or the parental taste for a variety of child's gender. The studies based on the former, among others, include the seminal paper by Angrist and Evans (1998) (parental labour supply), Angrist, Lavy, & Schlosser (2010), and Black, Devereux, & Salvanes (2010) (quantity-quality trade-off); Conley and Glauber (2006) (schooling); Kruk and Reinhold (2014) (later-life depression); van den Broek and Tosi (2020) (later-life loneliness). Accordingly, with this paper, we contribute to the existing literature by showing that the number of children continues to impact individuals' health decisions during the pandemic.

The paper proceeds as follows. Section 2 explains the dataset and shows the descriptive findings. We discuss the identification strategy in Section 3. Section 4 provides the main findings, mechanism, and shows the analysis of heterogeneity. Section 5 discusses further the study. Additional robustness checks are in Appendix.

2. Data and Descriptive Findings

This Section first presents the survey data used in the study. Next, we discuss the descriptive statistics of the target sample.

2.1 SHARE data

This study exploits the Survey of Health, Ageing and Retirement in Europe (SHARE) and SHARE Corona Survey. SHARE provides socio-demographic information about individuals above 50 years old in 27 European countries and Israel. During the outbreak of Corona, the special questionnaire, SHARE Corona Survey, was designed to measure the consequences of the pandemic among older individuals. In this study, we consider the second Corona Survey fielded between June 2021 and August 2021 that contains information about vaccination decisions with most of the surveys taking place in July 2021. Additionally, we enrich information about respondents and construct variables of interest thanks to the longitudinal dimension of the SHARE. In total, we consider 25,026 individuals with two or more biological children from the following 28 countries: Austria, Germany, Sweden, Netherlands, Spain, Italy, France, Denmark, Greece, Switzerland, Belgium, Israel, Czech Republic, Poland, Luxembourg, Hungary, Portugal, Slovenia, Estonia, Croatia, Lithuania, Bulgaria, Cyprus, Finland, Latvia, Malta, Romania, and Slovakia.

Instrumental variables. The main instrument for the probability to have more than two children is the sex composition of the two firstborn children. Pre-pandemic SHARE waves ask each survey participant the year of birth and the sex of the children.[5] Accordingly, we can follow their fertility history and identify the sex of the two first born children. In the study, we restrict to individuals who consistently reported information about children's sex and year of birth to decrease the misreporting error.[6] Based on the sex composition of two firstborns, we

[5]This piece of information is available for more than 92 percent of respondents in the second SHARE Corona Survey.

[6]Appendix C, Table C.1 documents that our findings are robust for other selections: including individuals who reported different composition or number of children across surveys or across questions.

define the instrument as equal to one if the respondent has two daughters or two sons, and 0 if she/he has either a daughter-son or a son-daughter combination.[7]

Endogenous variables. The endogenous variable is a binary variable that is equal to one if the individual has more than two children and zero otherwise. In the study, we restrict to individuals who report the consistent number of children across the SHARE waves. Doing so, the target sample considers only surviving children before 2020, as they are the ones who compose the network during the outbreak of Corona.[8] We show that our results hold when we add all born children and/or non-biological children.[9]

Vaccination variables. The Second SHARE Corona Survey includes questions about COVID-19 vaccine. First, each respondent answers *"Have you been vaccinated against Covid-19?"*, and the ones who replied *"No"* received a follow-up question whether they already have made an appointment, want to get vaccinated, do not want to get vaccinated or are undecided.[10] Based on these questions, we define our main outcome variable: to be vaccinated or to have already made an appointment. It is equal to 0 if a respondent answers that she/he wants to get vaccinated but does not have an appointment, does not want to get vaccinated or is undecided. We show that our results remain the same when we pool together individuals who got vaccinated, have already made an appointment and want to get vaccinated as opposed to do not want and to be undecided.

Exposure to the Coronavirus. To investigate the mechanism, first, we exploit the SHARE Corona Survey. In particular, the questions: *"Did you or anyone close to you experience symptoms that you would attribute to the Covid illness, e.g. cough, fever, difficulty breathing, or loss of sense of taste or smell?"* and: *"Have you or anyone close to you been tested for the Corona virus and the result was positive?"* and: *"Have you or anyone close to you been hospitalized due to an infection from the Corona virus?"*. These questions are followed by

[7]In the study, we do not use the multiple birth instrument because of the non-random assignment of the twin-birth discussed in the recent article by Bhalotra and Clarke (2020).
[8]Our results remain the same when we include children from the first wave participated in.
[9]Even though there are few non-biological children in our sample. In the preferred specification we exclude them to minimise any endogenous marriage decision or adoption depending on preferences for gender that might in turn impact the vaccination choice. Anyways, our results hold when we include the non-biological children.
[10]The refusal rate is very low, below 0.15 percent. A follow up question is *"Do you want to get vaccinated against Covid-19?"*. The answers include *"Yes, I already have a vaccination scheduled"* (6 percent); *"Yes, I want to get vaccinated"* (16 percent); *"No, I do not want to get vaccinated"* (43 percent); and *"I'm still undecided"* (35 percent).

"Who was it?". Combining these answers, we can identify if a respondent knew someone who is not herself/himself who has had symptoms similar to the Coronavirus disease or tested positive for the virus or was hospitalised.

Next, we use pre-pandemic questions to derive information about the network size of each individual, the number of grandchildren, and having a great-grandchild. Finally, focusing on children, we define the total number of contacts that a respondent had with them before the outbreak of Corona.[11]

Demographic and socio-economic variables. The main analysis controls only for gender, age, educational attainment and the country of residence. We define three educational categories: less than high school, high school and tertiary education. To further validate our identification strategy by showing that our instrument is unrelated to predetermined characteristics, we draw retrospective information from the SHARELIFE module about individual characteristics before having the first child: place of birth (city, rural area), being a bad student at mathematics or language at age 10, household size at age 10, number of rooms in the dwelling at age 10, number of books in house at age 10, number of chronic disease at age 10, having mental health problem at age 10, vaccinated before age 16, age at child's birth and age at marriage.

2.2 Descriptive findings

Table 1 reports the descriptive statistics for our target sample - respondents with two or more children. 82.5 percent of individuals already got or scheduled a COVID vaccine by the moment of the survey. This share is in line with the almost universal access to vaccines against COVID among the older individuals in our sample by Summer 2021. Using external information from The Oxford COVID-19 Government Response Tracker (OxCGRT), we show that the vaccine in the selected countries became available to almost all individuals way before the survey was fielded.[12] Moreover, since we focus on respondents aged above 50 years old, the universal availability for this group was likely reached even earlier. Looking at the fertility history, the

[11]We sum up all the contacts with children to get the total number of days per year. The question asks about the frequency of contacts on a categorical scale. Accordingly, we redefine daily frequency as 365 days, several times a week as 104 days, once a week as 52 days, about every two weeks as 26 days, about once a month as 12, less than once a month as 5, and never as 0.

[12]The results matched between the interview date and the OxCGRT vaccination policy for the countries in this study are available upon request.

average number of children is 2.54, and 36 percent of respondents report having more than two children. Thus, the probability to be treated is equal to 0.36. As for the instrumental variable, 50.8 percent of individuals have two firstborns of the same sex. This is in line with the stylized fact that having a firstborn boy is more likely than a girl (0.51 versus 0.49). Accordingly, the share of respondents with the two firstborn sons is slightly higher than with daughters, 0.26 and 0.24, respectively, as in the seminal paper by Angrist and Evans (1998).

========== About here Table 1 ==========

The average age of individuals in our sample is 71.6 years and they already completed fertility potentially long before the outbreak of Corona in 2019.[13] The sample includes a slightly larger share of women, 0.59, in line with higher female life expectancy. By Summer 2021, 38.8 percent of respondents knew someone who had symptoms similar to the Coronavirus. On average, survey participants report having 2.8 people in their network and 3.3 grandchildren before the pandemic. About 13 percent of respondents report being a great-grandparent, and they had contact with their children, on average, 362 days per year.

3. Empirical approach and identification strategy

This article aims to estimate the effect of family size on the probability to be vaccinated against COVID-19. The model can be written as:

$$y_i = \beta_0 + \beta_1 C_i + X'_i \beta_2 + \varepsilon_i, \quad (1)$$

where y_i is a binary variable that is equal to one if the individual *i* is vaccinated and zero otherwise, C_i is a binary variable that is equal to one if the individual *i* has more than two children and zero if she/he has two children, X_i is a column vector of a set of predetermined factors that are assumed to affect the probability to be vaccinated (it includes five-year age dummies, gender, the educational attainment and country fixed effects), ε_i is the error term, and

[13]The family size for the respondents in our study is unlikely to be affected by the outbreak of Corona. Moreover, we draw this piece of information from pre-pandemic waves to alleviate concern about the direct impact of pandemic on family formation.

β_0, β_1, and β_2 are parameters to be estimated. These parameters could be consistently estimated by the Ordinary Least Squares (OLS) under the assumption that the number of children is mean independent of the error term. This condition is unlikely to be met in the present context. The decision of having more children is clearly non-random and depends on a large number of unobserved characteristics that are likely to be correlated with the probability to be vaccinated (e.g. risk aversion, time preferences, health, financial resources). For example, let's consider one common unobserved factor in both decisions - religion. It has been extensively documented that being more religious leads to the refusal of the use of contraception resulting in more children (e.g. Adsera, 2006). At the same time, recent articles have also shown that being religious is a direct predictor of a lower vaccine acceptance (e.g. Callaghan et al., 2021). Then, the simple OLS estimator would be downward biased. To mitigate this issue, we follow the literature starting with Angrist and Evans (1998) and take the advantage of parental preferences for mixed-sex offspring, which allows to use the sex composition of the first two children as a source of exogenous variation in the probability to have more than two children. The first-stage regression is written as follows:

$$C_i = \alpha_0 + \alpha_1 Z_i + X'_i \alpha_2 + \nu_{i,} \quad (2)$$

where Z_i equals one if the two firstborns are either two boys or two girls, and zero if the two firstborns are of different sex, ν_i is the error term that is assumed to be mean independent of Z_i and X_i.

Given that it is unlikely that the treatment effect is homogeneous across the population of interest (i.e. the effect of having more than two children on the probability to be vaccinated might differ across individuals), the two-stage least squares (2SLS) estimator identifies a Local Average Treatment Effect (LATE). Therefore, the instrument needs to satisfy four conditions (Imbens and Angrist, 1994). First, the instrument needs to be related to the probability of having more than two children (*the relevance assumption*). Second, it should be as good as randomly assigned (i.e. it should be independent of the potential outcomes and potential treatment assignments) (*the independence assumption*). Third, while it may have no effect on the probability of having more than two children for some individuals, all those who are affected

should be affected in the same way (*the monotonicity assumption*). Finally, it should affect the probability to be vaccinated only through its impact on the probability of having more than two children (*the exclusion restriction assumption*).

The sex composition of the first two children meets the relevance assumption: parents generally prefer to have at least one son and one daughter as opposed to two children of the same sex (see e.g. Ben-Porath and Welch, 1976), and parents who have either two daughters or two sons are more likely to have a third child than parents who have one son and one daughter (e.g. Angrist and Evans, 1998). The results of the first stage regression will confirm that it holds for our target sample.

The instrument also meets the independence assumption: the sex composition of the first two children is plausibly randomly assigned, and there is no reason to believe that the sex composition of the first two children would affect any other characteristics related to the probability to be vaccinated (e.g. health, preferences). Naturally, we cannot test this assumption directly, but we can provide suggestive evidence based on events that occurred before having the firstborn. To do so, we exploit early-life sociodemographic information. Appendix A, Table A.1 reports the balance test if parents who have same-sex children and mixed-sex children differ along with early-life socio-demographic characteristics, place of birth, and childhood health. The difference is not statistically significant at a 5 percent level across any variable. Accordingly, it suggests that the instrument is as good as random.

In order for the sex composition of the first two children to meet the monotonicity assumption, there should not be anyone who would prefer to have more children after having one son and one daughter, but not after having two children of the same sex. While naturally we cannot test this assumption in the data, De Chaisemartin (2017) show that violations of the monotonicity assumption do not affect the results as long as the LATEs of "compliers" (people who react to the instrument in a typical way) and "defiers" (people who react in the opposite way) do not differ too much, or if a subgroup of compliers accounts for the same percentage of the population as the percentage of defiers and has the same LATE, then the 2SLS estimator still allows identification of a LATE.

Finally, to meet the exclusion restriction assumption, the sex composition of the first two children should affect the probability to be vaccinated against COVID-19 only via its impact on the probability to have more than two children. Angrist and Evans (1998) discuss one potential

threat to the exclusion restriction assumption which is related to a slightly higher probability of having a boy than a girl (Markle, 1974), and this difference results in a higher probability to have two children of the same sex when the first child is a boy. Children's sex, in turn, is related to their own health outcomes (MacLean et al., 2013), parents' labour market outcomes (Lundberg and Rose, 2002) and parents' probability of divorce (Dahl and Moretti, 2008; Kabátek and Ribar, 2021). The sex composition of the first two children may therefore potentially affect the probability to be vaccinated through other channels besides its impact on fertility. Still, the potential correlation between the sex composition of the first two children and children's sex can be taken into account by controlling for the sex of the first- and second-born child.

As an additional check, instead of using the fact that the two first children have the same sex as the instrument, we also use two instruments: a dummy variable that is equal to one when the first two children are girls and another dummy variable that is equal to one when the first two children are boys. As explained by Angrist and Evans (1998), this alternative identification strategy allows us to perform an overidentification test to check whether the sex of the children might bias the results. The motivation for using this test is that the bias due to an effect of child sex on vaccine take up should be different according to the instrument used, i.e. two boys or two girls. The null hypothesis of the overidentification test is that the two boys and the two girls instruments give the same estimates when used separately. The statistic used for this test consists in the Hansen J statistic that is assumed to be distributed as a Chi-square with one degree of freedom under the null.

Another potential threat to the exclusion restriction assumption has been mentioned in the literature. Having same- as opposed to mixed-sex children might affect child rearing costs, as parents may prefer to have daughters and sons in separate rooms. Rosenzweig and Wolpin (2000) found evidence that having mixed-sex children was associated with higher childrearing costs in rural India, but Bütikofer (2011) found no meaningful differences in richer countries such as the United Kingdom, Switzerland, Mexico, Bulgaria and Albania. Since we focus on Europe, in this context the sex composition of the first two children is therefore unlikely to meaningfully affect childrearing costs. Moreover, it is not clear whether the induced costs of child rearing would affect the probability to get vaccinated against COVID-19 later in life, especially given that vaccination was free-of-charge in all the countries considered in this

study.[14]

In sum, the sex composition of the first two children, therefore, offers a reasonable instrument for establishing the causal effect of having three or more children versus exactly two children on the probability to be vaccinated.

4. Results

In this Section, we present the estimates of the first-stage equation and the main results about the impact of family size on vaccination decisions. Next, we discuss the potential mechanism which explains these findings followed by heterogeneity analysis.

4.1 First-stage results

Table 2 documents the results for the first stage regression described in Equation 2 above. Individuals who have two firstborns of the same-sex have a 6 percentage points higher probability of having more than two children (Table 2).[15] The magnitude of the estimate is similar to Angrist and Evans (1998) and more recent studies based on European data (Kruk and Reinhold, 2014; van den Broek and Tosi, 2020).[16]

========== About here Table 2 ==========

Table 2 shows that the estimate is stable across specifications and sets of controls. Column 1 includes only a constant, age, gender and country dummies. Column 2 adds educational attainment to increase the precision of the estimate in the second stage. Next, in column 3, we control for the sex of the two firstborns to mitigate the potential omitted variable bias due to the difference in probability of having a son or a daughter. The findings remain

[14]Refer to https://ec.europa.eu/info/live-work-travel-eu/coronavirus-response/safe-covid-19-vaccines-europeans/questions-and-answers-covid-19-vaccination-eu_en#vaccination the distribution of vaccines might vary across countries. But, checking country-by-country we do not find any country in our sample in which access to vaccines is not free of charge.

[15]When we replace the dummy for more than two children with the number of children, the findings are naturally very similar. Having two firstborns of the same-sex increases the number of children by 0.08. Since results are very similar we do not report them.

[16]Kruk and Reinhold (2014) also use the SHARE data, compared to them our target sample includes ten-year later-born cohorts and covers 28 instead of 12 countries.

unchanged. Then, we use two separate instruments for the two firstborns being boys or girls to check that our results are not driven by the preference for the specific gender of children. Column 4 reports no evidence that only one child's gender drives the results in the first stage regression: both sons and daughters positively affect the number of children. Moreover, the magnitudes of the estimates are not statistically different (p-value=0.345). This finding is in line with the initial hypothesis that parents have a taste for diversity of children's gender, and we conclude that the specific gender does not drive the results of the first-stage equation. In the analysis, we also estimated the equation adding further controls for events before childbearing like early-life socioeconomic characteristics and all specifications lead to very similar findings.[17]

4.2 Family size and COVID vaccination

Table 3 reports our main results on how the family size affects the probability of getting a vaccine against COVID. Column 1 is the OLS estimates and shows that having more than two children is associated with a 2.8 percentage points lower probability of being vaccinated. When we correct for endogeneity and apply the same sex of two firstborn children as an instrument, having more than two children increases the getting COVID vaccine by about 14.8 percentage points (column 2).[18],[19] Next, as expected with evidence in Section 3, the magnitude and significance remains the same when we control for parental educational attainment (column 3) which is in line with as good as random assignment of the same sex instrument. In all specifications the instrument is relevant, and F-test on the excluded instrument is above 100, larger than the conventional rule of thumb for the weak instrument. Next, we include controls for the sex of each child to reduce the potential omitted variable bias that might be related to a differential impact of children's gender on vaccination decisions. Column 4 shows that the main estimate of having more than two children does not change and remains statistically significant. The estimates of the first child being a boy and the second child being a boy are both very small compared to the 2SLS estimate. Finally, Column 5 replaces our same-sex instrument with two

[17]These results are available upon request.

[18]In case, when we use the number of children as an endogenous variable, the 2SLS estimate on the probability to get the vaccine is equal to 0.096 (p-value=0.050). In the main analysis, we prefer to focus on the indicator for having more than two children as an endogenous variable because it captures an increase from two to more children rather than a per child effect as noted in Angrist and Evans (1998).

[19]Appendix C, Table C.1 confirms that our results are robust to alternative ways to derive the variables of interest and the target sample.

instruments for having two boys and two girls. In this case, we can perform the overidentification test for differences in the impact of having two boys versus having two girls on vaccination. The p-value of this test is equal to 0.46. Moreover, the 2SLS estimate does not change compared to the conventional specification in Column 3. These two facts together suggest that our main estimate is unlikely to be driven by gender and captures an increase in the overall number of children. Accordingly, the 2SLS estimate documents a large and statistically significant impact of family size on the COVID vaccination decision among older individuals during the pandemic. And, this impact holds when we combine together individuals with revealed and stated preferences about vaccination, i.e. who already got vaccinated and who want to take it in case they did not take it yet (see Column 1 in Table D.1. Appendix D).

========== About here Table 3 ==========

So far, we discussed the 2SLS estimate of family size which disregards the binary nature of the vaccination decision and having more than two children. However, when we impose an additional structure on the joint process of getting vaccinated and having more than two children, we can derive the model parameters in both equations. In particular, if we assume that the underlined process is a bivariate normal with two errors following the joint normal distribution we can identify the average partial effect (APE) of having more than two children as the difference between the probability of being vaccinated against COVID-19 if a respondent has more than two children and the probability of being vaccinated if she/he has exactly two children. Appendix B, Table B.1 shows the results. Assuming normality of the error terms and the model structure, the APE equals 0.113 and is statistically significant at a 1 percent level. In our case, the LATE from 2SLS is remarkably close to the APE suggesting that the estimation method does not seem to drive the results. Accordingly, both identification strategies lead to the same result: family size significantly increases the vaccination decision. Still, for the question of space and the flexibility to the distribution of the error terms, we prefer to use the linear IV estimator in the main analysis.

4.3 Mechanism

Why does a larger family size result in a higher vaccination rate against COVID-19? From the very beginning of the pandemic, it has been documented that the SARS-CoV-2 virus can be easily transmitted through air droplets and maintaining social distancing is necessary (WHO, March 2020).[20] As a result, minimising personal contacts has been encouraged in all countries since Spring 2020, and before the availability of the vaccine, it was considered almost the main tool to stop the spread of the virus. As the vaccine against COVID became almost universally available to all individuals in our sample by the moment of the survey, getting a jab became necessary to reduce the risk of getting the disease. Our data allows us to check, specifically, if the differential exposure to the virus due to the number of children and social interactions with them before the pandemic can partially explain the increase in vaccine uptake.

First, let's consider questions from the SHARE Corona Survey. Table 4 repeats the 2SLS identification strategy considering different outcome variables. By using the sex composition of the first two children as instrument, we document that family size increases by 22 percentage points the probability of knowing someone who had the COVID symptoms (column 1) and by 17 percentage points - who tested positive (column 2). These two estimates are large and statistically significant suggesting the higher awareness about the virus among individuals with more children.

========== About here Table 4 ==========

As for the first-hand experience about the severe form of Corona illness, the SHARE Corona Survey asks if a respondent knows someone apart from her/him who was hospitalised due to an infection (column 3). When we use this variable as an additional outcome, family size does not affect it. It suggests that even though the overall awareness about the virus increases due to the number of children, it does not translate into knowing more individuals who went to the hospital, which might be in line with a lower probability to develop a severe form among younger individuals. However, the limitation of this question is that it only captures the severity

[20]Refer to "Modes of transmission of virus causing COVID-19: implications for IPC precaution recommendations" https://www.who.int/news-room/commentaries/detail/modes-of-transmission-of-virus-causing-covid-19-implications-for-ipc-precaution-recommendations

of illness by hospitalisation but a respondent could have known someone who experienced severe disease without being hospitalised or could develop the long-term health effects after having the disease that in turn leads to a first-hand negative experience.

Next, we switch to pre-pandemic information. On average, having more than two children leads to a statistically significant increase in the network size (column 1), the number of children (column 2), and the probability of being a great-grandparent (column 3). These estimates correspond with an almost mechanical expected increase due to the presence of an additional child. Even though these results support our initial hypothesis about higher exposure, one can argue that knowing more people, in general, or specifically having more children, does not imply being in contact with them and meeting them regularly (e.g. Ward et al. (2009) on a discussion about quality of relationship). Thus, we further exploit pre-pandemic information and use the number of contacts with children. By considering behaviour before the outbreak of Corona, we can rule out any potential adjustment in the number of contacts during the pandemic as a result of the differential fear of the virus. Column 4 documents that a large family size increases by 110 days the total number of contacts with children during the year.

========== About here Table 5 ==========

One caveat about our results in Table 5 is worth noting. Even though the 2SLS estimates do not change and remain significant after controlling for the gender of the first and the second child, the null hypothesis of the overidentification test associated with using two instruments -having two boys and having two girls, separately- is rejected for all four outcomes, differently from our main results on vaccination. This fact implies that the sex of the children affects these outcomes beyond its effect on the probability to have more than two children. Such a difference related to network size, the number of grandchildren, and being a great grandparent can be explained by the fact that women become parents at a younger age than men.[21] The difference in the total number of contacts can be attributed to higher attachment to daughters, e.g. Crespo and Mira (2014) on evidence about adult daughters and care provision. Given that the probability to have the two firstborns with the same sex is higher if they are boys (because of the probability to

[21]Previous studies, Rupert and Zanella (2018) using the Panel Study of Income Dynamics and Backhaus and Barslund (2021) using the SHARE data, documented this age difference resulting from having a girl versus having a boy at the first birth.

be a boy is slighly higher), it means that the instrument "same sex" is negatively correlated with the average number of daughters. It implies that the 2SLS estimates for these outcomes are biased downward.

We also check if the number of children impacts the probability of coresidence with a child and household size as within-family interactions have been considered the main source of getting the transmissible disease (Bouckaert et al., 2020). The results indicate the absence of any statistically significant impact on these outcomes. That is in line with an older average age of respondents in the SHARE Corona Survey, 72 years old, and by that age, children often have left the parental house.

4.4 Cross-country heterogeneity

Our results imply the significant effect of the family size for the vaccination decision. However, how does this impact vary across countries? First, we group countries depending on the average excess mortality from March to June 2020 based on Eurostat data. Column 1 in Table 5 reports the results for six badly hit countries with excess mortality above 14 percent by June 1, 2020: Belgium, France, Italy, the Netherlands, Spain, and Sweden. Column 2 pools together the rest of twenty two countries. The magnitude of the impact of family size on the probability of getting a COVID-19 vaccine is not statistically different across badly hit countries and less hit countries. Still, the point estimate is statistically significant from zero in countries with lower mortality during the first wave of the pandemic. This result suggests that first-hand experience and individual factors matter more for making vaccination decisions in places with relatively lower mortality at the country level as perhaps in the badly hit countries the awareness about disease increases regardless of family size.

========== About here Table 6 ==========

Next, we split the sample into three country groups based on the geographical classification which is also related to established social norms: 1) Nordic or Central countries: Sweden, Netherlands, Denmark, Ireland, Finland, Belgium, Austria, Germany, France, Switzerland, Luxembourg, and Israel, and 2) Southern countries: Spain, Italy, Greece, Portugal, Cyprus, and Malta; and 3) Eastern group: Czech Republic, Poland, Hungary, Slovenia, Estonia,

Croatia, Lithuania, Bulgaria, Latvia, Romania, and Slovakia. Our conjecture is that the family-specific exposure and the impact of family size on vaccination should be larger in countries with more contacts with children as individuals might face a higher probability to get the virus during the personal interactions, and a stronger willingness to be back to their number of contacts that were interrupted by the outbreak of Corona. The magnitude of the effect of having more than two children on vaccination almost doubles in Southern countries (column 4). Despite the loss of precision due to smaller sample size, this impact is statistically significant at a 5 percent level, and it is different from the 2SLS estimate in Nordic or Central countries. Southern countries in our sample are typically characterised by strong family ties and common intergenerational interactions within the family (Bolin et al., 2008). Indeed, when we look at the total number of contacts with children in pre-pandemic, the increase in Southern countries is particularly large compared to other groups.[22]

Beyond geographical comparison, we also checked heterogeneity across socio-demographic individual factors, like age, gender, predetermined health, and others. This analysis shows that the main impact is homogeneous in our target sample.

5. Discussion

This study documents the impact of the family size on the COVID-19 vaccine uptake in Summer 2021. Beyond vaccination, we analyse the effect of having more children on arguably less costly preventive measures against the Coronavirus. We have information about roughly ten different types of preventive measures like wearing a mask, maintaining social distance, hands washing, the frequency of going to public places, and others, in Summer 2020 and Summer 2021. Following Bertoni et al. (2021) we reduce the dimensionality of the questions by using polychoric principal component analysis (PCA). In Summer 2020 and Summer 2021, two and three components, respectively, have eigenvalues larger than one and explain about 70 percent of the total variance of all preventive measures. We use the components defined in this analysis as the main outcome variables in Table E.1 Appendix E. Even though the 2SLS estimates of family size are positive for all measures, they are not statistically different from zero. No effect can be related to the relatively lower perceived protection of other measures against getting the Corona

[22]Having more than two children increases the number of contacts by about 75 days (se 107) per year in Nordic or Central countries, by 195 days (se 88) per year in Southern countries and decreases by 25 days (se 120) per year in Eastern countries.

illness compared to vaccination. Moreover, preventive behaviour in Summer 2021 might be already affected by taking the COVID-19 vaccine (Anderson et al., 2021).

Next, when we switch to other vaccines apart from COVID, flu and pneumonia, we document consistent evidence that such an increase in the willingness to get the shot due to a large family is COVID-vaccine-specific (see Table F.1, Appendix F).[23] It is also likely to be related to the pandemic time and the uncertainty that the outbreak of Corona caused about the probability of getting illness and about its consequences. The importance of vaccination against influenza and pneumonia received less coverage and attention, which might mitigate the impact of family size. Moreover, COVID-19 vaccine was distributed free of charge across countries to rule out any supply or income factors in getting access to it. Thus, the effect of family size on vaccination choice could be particularly pronounced for the pandemic-caused illness and uncertainty of the time after the outbreak.

We conclude this Section by mentioning some limitations of our study. Even though our data allows us to test potential mechanisms -the risk of getting the virus, exposure to the illness- there can be other plausible channels beyond the one we can check. For example, if children were helping in booking an appointment for vaccination or commuting to the vaccination centre as shown in anecdotal evidence, then the documented increase in the vaccination rate can be, in part, also attributed to it. The other mechanism might be related to the differential drop in the lifestyle compared to pre-pandemic time depending on the family size. If parents with more children used to spend more time together increasing not only the number of contacts (as we show in the data) but also the duration of them and the type of activities, then vaccination becomes crucial to return to old habits. Related to variables in the Corona Survey, we can only consider the extensive margin of exposure during the pandemic: knowing someone who had COVID. However, the intensive margin can also matter: how many people had COVID among whom a respondent knows; the closeness of relationship between a respondent and a person who got the virus; the form of disease and others.

Related to the econometric strategy, as in any IV strategy with heterogeneous treatment effects, our estimation method captures effects on individuals affected by the instrument (Angrist et al., 1996), which has some implications for the external validity of our estimates. In our

[23]We take information about vaccination decisions from the second SHARE Corona Survey in Summer 2021. Before the outbreak of Corona, there has been no completed survey wave asking about vaccines.

specific case, it implies that our method identifies the causal effect of a planned change in fertility for parents who have preferences for mixed-sex offspring. Several studies interested in the impact of family size compared their findings based on the sex composition of two firstborns to the ones using the multiple births at the first or the second birth as an instrumental variable. However, the recent study by Bhalotra and Clarke (2020) documents the selection of women into multiple births that leads to the not-random assignment of the twin instrument, and, in turn, causes the failure of the identification strategy. Accordingly, in our article, we focus only on the analysis based on the sex composition of the two firstborns.

Moreover, using the sex composition of the first two children as an instrument means that the analyses do not determine the causal effect of having more children at lower parities (e.g. having one versus no children, or having a second child) on the probability to be vaccinated. Nevertheless, this group of individuals corresponds with more than 70 percent of the full sample, and analysing the effect of having three or more versus two children is highly relevant in the European context, given that much of the change in European fertility over the past decades has been due to a decrease in the proportion of people having three or more children (Sobotka and Beaujouan, 2014).

6. Conclusion

The COVID-19 pandemic led to a dramatic increase in the number of deaths since the initial outbreak of the virus. Vaccination has become the main tool to protect oneself and others from getting the disease. In the time of the COVID-19 pandemic, studying the factors that drive vaccination decisions against the Coronavirus is key to reaching herd immunity, in particular, due to the availability of shots in most countries. In the study, we investigate the role of personal factors and focus on family size.

To answer this research question, we study older individuals for whom the Coronavirus was considered particularly deadly and who were included in the priority groups of the COVID-19 vaccination campaigns in many countries. Since the fertility decision is endogenous, we apply an instrumental variable strategy and exploit the parental preferences for mixed-sex children as an exogenous variation. We find that a larger family size leads to a higher probability to get vaccinated against COVID. The potential mechanism behind this result is - a higher exposure to disease and the probability of getting a contagion.

Our results have important policy implications. First, we directly show that the number of children increases the uptake of the COVID-19 vaccine which has been distributed at no cost in our sample. Still, due to the existence of such a gap between families with more children as opposed to two, policymakers should increase the awareness about the Coronavirus in the whole population and implement easy-to-get vaccination campaigns to reach herd immunity quicker. As for the future, in light of the recent fall in fertility across Europe and for preparation for other pandemics, policymakers should be cautious that family size as a driver for vaccination decisions could be more limited in some decades.

References

Adsera, A. (2006). Religion and changes in family-size norms in developed countries. *Review of Religious Research*, 271–286.

Ahituv, A., Hotz, V. J., & Philipson, T. (1996). The responsiveness of the demand for condoms to the local prevalence of aids. *Journal of Human Resources*, 869–897.

Andersson, O., Campos-Mercade, P., Meier, A. N., & Wengström, E. (2021). Anticipation of covid-19 vaccines reduces willingness to socially distance. *Journal of Health Economics*, *80*, 102530.

Angrist, J., & Evans, W. (1998). Children and their parents' labor supply: Evidence from exogenous variation in family size. *The American Economic Review*, *88*(3), 450–477.

Angrist, J., Imbens, G. W., & Rubin, D. B. (1996). Identification of causal effects using instrumental variables. *Journal of the American Statistical Association*, *91*(434), 444–455.

Angrist, J., Lavy, V., & Schlosser, A. (2010). Multiple experiments for the causal link between the quantity and quality of children. *Journal of Labor Economics*, *28*(4), 773–824.

Backhaus, A., & Barslund, M. (2021). The effect of grandchildren on grandparental labor supply: evidence from europe. *European Economic Review*, *137*, 103817.

Ben-Porath, Y., & Welch, F. (1976). Do sex preferences really matter? *The Quarterly Journal of Economics*, *90*(2), 285–307.

Bergmann, M., Hannemann, T.-V., Bethmann, A., & Schumacher, A. (2021). Determinants of SARS-CoV-2 Vaccinations in the 50+ Population. *MEA Discussion Paper*.

Bertoni, M., Celidoni, M., Dal Bianco, C., & Weber, G. (2021). How did European retirees respond to the COVID-19 pandemic? *Economics Letters*, *203*, 109853.

Bhalotra, S., & Clarke, D. (2020). The twin instrument: Fertility and human capital investment. *Journal of the European Economic Association*, *18*(6), 3090–3139.

Black, S. E., Devereux, P. J., & Salvanes, K. G. (2010). Small family, smart family? Family size and the IQ scores of young men. *Journal of Human Resources*, *45*(1), 33–58.

Bolin, K., Lindgren, B., & Lundborg, P. (2008). Your next of kin or your own career?: Caring and working among the 50+ of europe. *Journal of Health Economics*, *27*(3), 718–738.

Börsch-Supan, A. (2020a). *Survey of Health, Ageing and Retirement in Europe (SHARE) Wave 1. Release version: 7.1.0. SHARE-ERIC. Data set.* doi:10.6103/SHARE.w1.710

Börsch-Supan, A. (2020b). *Survey of Health, Ageing and Retirement in Europe (SHARE) Wave 2. Release version: 7.1.0. SHARE-ERIC. Data set.* doi:10.6103/SHARE.w2.710

Börsch-Supan, A. (2020c). *Survey of Health, Ageing and Retirement in Europe (SHARE) Wave 3 - SHARELIFE. Release version: 7.1.0. SHARE-ERIC. Data set.* doi:10.6103/SHARE.w3.710

Börsch-Supan, A. (2020d). *Survey of Health, Ageing and Retirement in Europe (SHARE) Wave 4. Release version: 7.1.0. SHARE-ERIC. Data set.* doi:10.6103/SHARE.w4.710

Börsch-Supan, A. (2020e). *Survey of Health, Ageing and Retirement in Europe (SHARE) Wave 5. Release version: 7.1.0. SHARE-ERIC. Data set.* doi:10.6103/SHARE.w5.710

Börsch-Supan, A. (2020f). *Survey of Health, Ageing and Retirement in Europe (SHARE) Wave 6. Release version: 7.1.0. SHARE-ERIC. Data set.* doi:10.6103/SHARE.w6.710

Börsch-Supan, A. (2020g). *Survey of Health, Ageing and Retirement in Europe (SHARE) Wave 7. Release version: 7.1.1. SHARE-ERIC. Data set.* doi:10.6103/SHARE.w7.711

Börsch-Supan, A. (2020h). *Survey of Health, Ageing and Retirement in Europe (SHARE) Wave 8. COVID-19 Survey 1 Interview Date. Release version: 1.0.0. SHARE-ERIC. Data set.* doi:10.6103/SHARE.w8caintd.100

Börsch-Supan, A. (2020i). *Survey of Health, Ageing and Retirement in Europe (SHARE) Wave 8. COVID-19 Survey 1. Release version: 1.0.0. SHARE-ERIC. Data set.* doi:10.6103/SHARE.w8ca.100

Börsch-Supan, A. (2021a). *Survey of Health, Ageing and Retirement in Europe (SHARE) Wave 8. Release version: 1.0.0. SHARE-ERIC. Data set.* doi:10.6103/SHARE.w8.100

Börsch-Supan, A. (2021b). *Survey of Health, Ageing and Retirement in Europe (SHARE) Wave 9. COVID-19 Survey 2. Release version: 0 SHARE-ERIC. Preliminary data set.*

Börsch-Supan, A., Brandt, M., Hunkler, C., Kneip, T., Korbmacher, J., Malter, F., … Zuber, S. (2013). Data resource profile: the Survey of Health, Ageing and Retirement in Europe (SHARE). *International Journal of Epidemiology*, *42*(4), 992–1001. doi:10.1093/ije/dyt088

Bouckaert, N., Gielen, A. C., & Van Ourti, T. (2020). It runs in the family–Influenza vaccination and spillover effects. *Journal of Health Economics*, *74*, 102386.

Bütikofer, A. (2011). Sibling sex composition and cost of children. *Unpublished manuscript, Department of Economics, Norwegian School of Economics, Bergen, Norway*.

Callaghan, T., Moghtaderi, A., Lueck, J. A., Hotez, P., Strych, U., Dor, A., ... Motta, M. (2021). Correlates and disparities of intention to vaccinate against covid-19. *Social Science & Medicine*, *272*, 113638.

Caserotti, M., Girardi, P., Rubaltelli, E., Tasso, A., Lotto, L., & Gavaruzzi, T. (2021). Associations of covid-19 risk perception with vaccine hesitancy over time for italian residents. *Social Science & Medicine*, *272*, 113688.

Conley, D., & Glauber, R. (2006). Parental educational investment and children's academic risk estimates of the impact of sibship size and birth order from exogenous variation in fertility. *Journal of Human Resources*, *41*(4), 722–737.

Crespo, L., & Mira, P. (2014). Caregiving to elderly parents and employment status of European mature women. *Review of Economics and Statistics*, *96*(4), 693–709.

Dahl, G. B., & Moretti, E. (2008). The demand for sons. *The Review of Economic Studies*, *75*(4), 1085–1120.

Dai, H., Saccardo, S., Han, M. A., Roh, L., Raja, N., Vangala, S., ... Croymans, D. M. (2021). Behavioural nudges increase covid-19 vaccinations. *Nature*, *597*(7876), 404–409.

De Chaisemartin, C. (2017). Tolerating defiance? Local average treatment effects without monotonicity. *Quantitative Economics*, *8*(2), 367–396.

Diament, S. M., Kaya, A., & Magenheim, E. B. (2021). Frames that matter: Increasing the willingness to get the covid-19 vaccines. *Social Science & Medicine*, 114562.

Funk, S., Gilad, E., Watkins, C., & Jansen, V. A. (2009). The spread of awareness and its impact on epidemic outbreaks. *Proceedings of the National Academy of Sciences*, *106*(16), 6872–6877.

Giulietti, C., Vlassopoulos, M., & Zenou, Y. (2021). When reality bites: Local deaths and vaccine take-up. *Available at SSRN: https://papers.ssrn.com/sol3/papers.cfm?abstract_id=3980243*.

Grundy, E., & Read, S. (2012). Social contacts and receipt of help among older people in England: are there benefits of having more children? *Journals of Gerontology Series B: Psychological Sciences and Social Sciences*, *67*(6), 742–754.

Imbens, G. W., & Angrist, J. D. (1994). Identification and estimation of local average treatment effects. *Econometrica*, *62*(2), 467–475.

Kabátek, J., & Ribar, D. C. (2021). Daughters and divorce. *The Economic Journal*, *131*(637), 2144–2170.

Kreps, S., & Kriner, D. (2021). Factors influencing covid-19 vaccine acceptance across subgroups in the united states: Evidence from a conjoint experiment. *Vaccine*.

Kruk, K. E., & Reinhold, S. (2014). The effect of children on depression in old age. *Social Science & Medicine*, *100*, 1–11.

Lundberg, S., & Rose, E. (2002). The effects of sons and daughters on men's labor supply and wages. *Review of Economics and Statistics*, *84*(2), 251–268.

MacLean, A., Sweeting, H., Egan, M., Der, G., Adamson, J., & Hunt, K. (2013). How robust is the evidence of an emerging or increasing female excess in physical morbidity between childhood and adolescence? Results of a systematic literature review and meta-analyses. *Social Science & Medicine*, *78*, 96–112.

Markle, G. E. (1974). Sex ratio at birth: values, variance, and some determinants. *Demography*, *11*(1), 131–142.

Oster, E. (2018). Does disease cause vaccination? Disease outbreaks and vaccination response. *Journal of Health Economics*, *57*, 90-101.

Rosenzweig, M. R., & Wolpin, K. I. (2000). Natural" natural experiments" in economics. *Journal of Economic Literature*, *38*(4), 827–874.

Rupert, P., & Zanella, G. (2018). Grandchildren and their grandparents' labor supply. *Journal of Public Economics*, *159*, 89–103.

Scherpenzeel, A., Axt, K., Bergmann, M., Douhou, S., Oepen, A., Sand, G., ... Börsch-Supan, A. (2020). Collecting survey data among the 50+ population during the COVID-19 pandemic: The Survey of Health, Ageing and Retirement in Europe (SHARE). *Survey Research Methods*, 217-221. doi:10.18148/srm/2020.v14i2.7738

Sobotka, T., & Beaujouan, É. (2014). Two Is best? The persistence of a two-child family ideal in Europe. *Population and Development Review*, *40*(3), 391–419.

Solís Arce, J. S., Warren, S. S., Meriggi, N. F., Scacco, A., McMurry, N., Voors, M., … Omer, S. B. (2021). COVID-19 vaccine acceptance and hesitancy in low- and middle-income countries. *Nature Medicine*, *27*(8), 1385–1394. doi:10.1038/s41591-021-01454-y

van den Broek, T., & Tosi, M. (2020). The more the merrier? The causal effect of high fertility on later-life loneliness in Eastern Europe. *Social Indicators Research*, *149*(2), 733–748.

Wang, J., Jing, R., Lai, X., Zhang, H., Lyu, Y., Knoll, M. D., & Fang, H. (2020). Acceptance of covid-19 vaccination during the covid-19 pandemic in china. *Vaccines*, *8*(3), 482.

Ward, R. A., Spitze, G., & Deane, G. (2009). The more the merrier? Multiple parent-adult child relations. *Journal of Marriage and Family*, *71*(1), 161–173.

Article: "Family size and vaccination among older individuals: The case of COVID-19 vaccine"

Tables, Figures and Appendix

Table 1: Descriptive statistics

	Statistics				
	Mean	Std.Dev.	Min	Max	Obs.
Outcome:					
Got COVID-19 vaccine	0.825	0.380	0	1	25026
Endogenous variables:					
Number of children	2.543	0.940	2	14	25026
Dummy more than two	0.363	0.481	0	1	25026
Instrumental variables:					
First two same-sex	0.508	0.500	0	1	25026
Two firstborns - boys	0.262	0.439	0	1	25026
Two firstborns - girls	0.247	0.431	0	1	25026
Firstborn is a boy	0.512	0.500	0	1	25026
Secondborn is a boy	0.503	0.500	0	1	25026
Controls:					
Age	71.602	8.786	50	103	25026
Female	0.589	0.492	0	1	25026
Education:					
Less than high school	0.354	0.478	0	1	25026
High school	0.423	0.494	0	1	25026
Tertiary education	0.223	0.416	0	1	25026
Mechanism:					
Anyone had COVID-19 symptoms	0.388	0.487	0	1	24814
Anyone tested positive	0.379	0.485	0	1	24777
Anyone hospitalized	0.113	0.317	0	1	24920
Network size	2.803	1.644	0	7	23921
Number of grandchildren	3.320	2.829	0	20	20892
Has a greatgrandchild	0.132	0.338	0	1	20886
Total contacts with a child per year	362.328	331.732	0	4015	17166

Note: Respondents with two or more children. Preliminary SHARE wave 9 COVID-19 Survey 2 release 0 data.

Table 2: The first stage regression

	(1)	(2)	(3)	(4)
Two firstborns same-sex	0.0603***	0.0607***	0.0608***	
	(0.00592)	(0.00590)	(0.00590)	
Firstborn is a boy			-0.00304	
			(0.00590)	
Secondborn is a boy			-0.00490	
			(0.00589)	
Two firstborns - boys				0.0568***
				(0.00715)
Two firstborns - girls				0.0648***
				(0.00733)
Individual controls	No	Yes	Yes	Yes
F-test (IV)	103.6509	105.852	106.234	53.3181
Prob(Two boys = Two girls)				.3447
R2	0.0557	0.0647	0.0647	0.0647

Note: Number of individuals is 25026. All columns constant, a set of five-year age dummies, female, and country fixed effects. The list of individual controls includes educational attainment. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$. Preliminary SHARE wave 9 COVID-19 Survey 2 release 0 data.

Table 3: Family size and probability to get vaccinated against COVID-19

	OLS	2SLS			
	(1)	(2)	(3)	(4)	(5)
Dummy more than two	-0.0282***	0.148**	0.146**	0.143*	0.141*
	(0.00469)	(0.0747)	(0.0738)	(0.0736)	(0.0733)
Firstborn is a boy				-0.00414	
				(0.00446)	
Secondborn is a boy				0.00873*	
				(0.00448)	
Individual controls	No	No	Yes	Yes	Yes
F-test (IV)		103.6509	105.852	106.234	53.318
Hausman test (p-value)		.0152	.0188	.0206	.0222
Hansen test (p-value)		.	.	.	.46

Note: Number of individuals is 25026. All columns control for constant, a set of five-year age dummies, female, and country fixed effects. The list of individual controls includes educational attainment. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$. Preliminary SHARE wave 9 COVID-19 Survey 2 release 0 data.

Table 4: Family size and awareness about COVID-19

	Anyone had COVID-19 symptoms	Anyone had tested positive	Anyone hospitalized with COVID-19
	(1)	(2)	(3)
Dummy more than two	0.226**	0.169*	-0.000445
	(0.103)	(0.101)	(0.0662)
Individual controls	Yes	Yes	Yes
F-test (IV)	100.832	101.611	103.34
Hausman test (p-value)	.057	.185	.85
N	24833	24796	24940

Note: All columns control for constant, a set of five-year age dummies, female, educational attainment, and country fixed effects.* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$. Preliminary SHARE wave 9 COVID-19 Survey 2 release 0 data.

Table 5: Number of children and network size

	Network size	Number of grandchildren	Has a greatgranchild	Number of contacts per year
	(1)	(2)	(3)	(4)
Dummy more than two	0.601*	2.035***	0.112*	113.4*
	(0.316)	(0.482)	(0.0623)	(62.72)
Individual controls	Yes	Yes	Yes	Yes
F-test (IV)	111.268	102.187	102.257	101.103
Hausman test (p-value)	.163	.741	.337	.302
N	23993	20960	20954	17206

Note: All columns control for constant, female, a set of five-year age dummies, educational attainment and country fixed effects.* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

Table 6: Heterogeneity in the impact of family size on getting COVID-19 vaccine across countries

	Badly hit	Less hit	Nordic or Central	Southern	Eastern
	(1)	(2)	(3)	(4)	(5)
Dummy more than two	0.0939	0.163*	0.0269	0.282**	0.119
	(0.108)	(0.0902)	(0.0990)	(0.114)	(0.144)
Individual controls	Yes	Yes	Yes	Yes	Yes
F-test (IV)	20.2648	86.614	27.982	38.148	44.859
Hausman test (p-value)	.299	.033	.618	.004	.305
N	6352	18674	8007	6010	10268

Note: The list of controls includes constant, a set of five-year age dummies, female, educational attainment, and country fixed effects. Badly hit countries in Column 1 are Belgium, France, Italy, the Netherlands, Spain and Sweden. Column 3, Nordic or Central countries, includes Sweden, Netherlands, Denmark, Ireland, Finland, Belgium, Austria, Germany, France, Switzerland, Luxembourg, and Israel. Column 4, southern countries are Spain, Italy, Greece, Portugal, Cyprus, and Malta. Eastern countries are in Column 5: the Czech Republic, Poland, Hungary, Slovenia, Estonia, Croatia, Lithuania, Bulgaria, Latvia, Romania, and Slovakia. All columns control for constant, female, a set of five-year age dummies, educational attainment and country fixed effects.* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$. Preliminary SHARE wave 9 COVID-19 Survey 2 release 0 data.

Online Appendix

Acknowledgements

This paper uses data from SHARE Waves 1, 2, 3, 4, 5, 6, 7 and 8 (DOIs: 10.6103/SHARE.w1.710, 10.6103/SHARE.w2.710, 10.6103/SHARE.w3.710, 10.6103/SHARE.w4.710, 10.6103/SHARE.w5.710, 10.6103/SHARE.w6.710, 10.6103/SHARE.w7.711, 10.6103/SHARE.w8.100, 10.6103/SHARE.w8ca.100, 10.6103/SHARE.w8caintd.100) and Börsch-Supan (2021b), see Börsch- Supan et al. (2013) and Scherpenzeel et al. (2020) for methodological details. The SHARE data collection has been funded by the European Commission through FP5 (QLK6-CT-2001-00360), FP6 (SHARE-I3: RII-CT-2006-062193, COMPARE: CIT5-CT-2005-028857, SHARELIFE: CIT4-CT-2006-028812), FP7 (SHARE-PREP: GA N°211909, SHARE-LEAP: GA N°227822, SHARE M4: GA N°261982, DASISH: GA N°283646) and Horizon 2020 (SHARE-DEV3: GA N°676536, SHARE-COHESION: GA N°870628, SERISS: GA N°654221, SSHOC: GA N°823782) and by DG Employment, Social Affairs & Inclusion. Additional funding from the German Ministry of Education and Research, the Max Planck Society for the Advancement of Science, the U.S. National Institute on Aging (U01_AG09740-13S2, P01_AG005842, P01_AG08291, P30_AG12815, R21_AG025169, Y1-AG-4553-01, IAG_BSR06-11, OGHA_04-064, HHSN271201300071C) and from various national funding sources is gratefully acknowledged (see www.share-project.org).

A. Parental characteristics

Table A.1: Parental characteristics

	Same-sex		Mixed-sex			
	Mean (1)	SD	Mean (2)	SD	Difference (1)-(2)	(p-value)
Age	71.65	8.78	71.55	8.79	0.10	0.36
Female	0.59	0.49	0.59	0.49	-0.00	0.84
Education:						
Less than high school	0.36	0.48	0.35	0.48	0.01	0.32
High school	0.42	0.49	0.43	0.49	-0.01	0.17
Tertiary education	0.22	0.42	0.22	0.42	0.00	0.63
Place of birth:						
A big city	0.13	0.34	0.13	0.34	-0.00	0.82
The suburbs of a big city	0.04	0.20	0.04	0.21	-0.00	0.48
A large town	0.11	0.32	0.12	0.32	-0.00	0.41
A small town	0.17	0.38	0.18	0.38	-0.01	0.22
A rural area	0.54	0.50	0.53	0.50	0.01	0.06
Age when finished school	17.96	4.57	17.96	4.56	0.00	0.97
A bad student at math at age 10	0.64	0.48	0.64	0.48	-0.00	0.92
A bad student at language at age 10	0.61	0.49	0.61	0.49	0.00	0.62
$<$ 11 books	0.41	0.49	0.42	0.49	-0.00	0.63
11-25 books	0.24	0.43	0.24	0.43	0.00	0.96
$>$ 26 books	0.35	0.48	0.35	0.48	0.00	0.64
Health at age 10:						
No chronic during childhood	0.18	0.38	0.18	0.38	0.00	0.60
1 chronic during childhood	0.57	0.50	0.57	0.50	0.01	0.42
$>$ 2 chronic during childhood	0.25	0.43	0.26	0.44	-0.01	0.17
Mental problem during childhood	0.01	0.08	0.01	0.09	-0.00	0.13
Vaccinated during childhood	0.97	0.18	0.97	0.17	-0.00	0.13
Dwelling at age 10:						
$<$ 2 bedrooms	0.09	0.29	0.09	0.29	-0.00	0.67
2 bedrooms	0.27	0.44	0.27	0.44	0.00	0.44
3 bedrooms	0.28	0.45	0.29	0.45	-0.01	0.34
$>$ 3 bedrooms	0.36	0.48	0.35	0.48	0.00	0.65
Characteristics of the individual's dwelling at age 10:						
Number of services of the individual's dwelling	1.13	0.91	1.15	0.91	-0.02	0.12
Household size at age 10	5.53	2.20	5.56	2.20	-0.03	0.29
Number of individuals per room	2.12	1.46	2.12	1.42	-0.01	0.78
Employment history:						
Work experience:						
Work experience before the first child	6.61	5.02	6.56	4.91	0.05	0.47
Work experience before the second child	9.37	5.96	9.34	5.85	0.02	0.77
The first job:						
Age when started job 1	19.74	5.10	19.68	5.03	0.06	0.38
Fertility history:						
Age of delivery child 1	25.21	4.39	25.11	4.37	0.10	0.08
Age of delivery child 2	28.73	4.81	28.63	4.83	0.10	0.11
Observations	12722		12304			

Note: The list of observed characteristics from SHARELIFE. Column 1 restricts to same-sex firstborn children, and Column 3 - mixed-sex children. The last column reports the p-value of the null hypothesis about the equality of the two means.

[a] the number of services includes hot running water supply, having a toilet inside the house and others.

B. Bivariate model

The bivariate model is written in the following way:

$$\begin{aligned} y_i &= \mathbb{1}[\lambda C_i + \eta X_i + u_i > 0] \\ C_i &= \mathbb{1}[\kappa z_i + \zeta X_i + v_i > 0], \end{aligned} \tag{1}$$

where y_i is the probability to get vaccinated. C_i is the dummy for having more than two children. z_i equals one if the two firstborns are either boys or girls, it is equal to 0 if the two firstborns of different sex. The set of controls, X_i, includes constant, a set of five-year age dummies, country fixed effects, gender and educational attainment. We allow for the common unobservables in both decisions and assume that u_i and v_i are distributed as bivariate normal, with $\mathbb{E}(u_i|\cdot) = \mathbb{E}(v_i|\cdot) = 0$, $\text{var}(u_i|\cdot) = \text{var}(v_i|\cdot) = 1$ and $\text{Cov}(u_i, v_i|\cdot) = \rho$.

Since the coefficients from Equation 1 are not directly interpretable about the magnitude of the impact, we report the average difference between the probability of being vaccinated against Covid if she/he has more than two children and the probability - if she/he has two children and equals to $1/n \sum_i [\Phi(\lambda + \eta X_i) - \Phi(\eta X_i)]$.

Table B.1: Family size and probability to get vaccinated against COVID-19

	Individual probits		Bivariate model	
	Endog	Exclusion	Got vaccine	>2
	(1)	(2)	(3)	(4)
Dummy more than two	-0.101***		0.545***	
	(0.0220)		(0.184)	
Two firstborns same-sex		0.170***		0.171***
		(0.0166)		(0.0164)
rho			-0.397***	
Loglikelihood	-9509.62	-15576.14	-25082.57	
Average partial effect of having more than two children:				
Average partial effect			0.113***	
s.e.			0.0389	

Note: * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$. Number of individuals is 25026. The list of controls includes constant, a set of five-year age dummies, female, educational attainment, and country fixed effects. Column 1 is a probit regression. Columns 3 and 4 correspond with Equation 1. Preliminary SHARE wave 9 COVID-19 Survey 2 release 0 data.

C. Data construction

Table C.1: Different ways to construct the number of children

	Main	Different methods			
	(1)	(2)	(3)	(4)	(5)
Dummy more than two	0.146**	0.135*	0.157**	0.160**	
	(0.0738)	(0.0785)	(0.0740)	(0.0764)	
Dummy more than two					0.144*
					(0.0761)
Individual controls	Yes	Yes	Yes	Yes	Yes
F-test (IV)	105.852	90.9378	103.5316	97.6455	99.3927
Hausman test (p-value)	.0188	.	.0122	.0135	.0238
N	25026	27822	26696	26873	25170

Note: The list of individual controls includes constant, a set of five-year age dummies, female, educational attainment, and country fixed effects. Column 2 includes respondents who reported different composition or number of children across SHARE waves. Column 3 includes non-natural children. Column 4 includes respondents who report inconsistently having more than two children between the number of children and summing up children with gender. Column 5 uses the first reported information about the number of children and first two same sex. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$. Preliminary SHARE wave 9 COVID-19 Survey 2 release 0 data.

D. COVID-19 vaccine outcome variable

Table D.1: Family size and stated preferences about COVID-19 vaccine

	Got/Want	Do not want	Undecided
	(1)	(2)	(3)
Dummy more than two	0.122*	-0.0759	-0.0196
	(0.0687)	(0.0547)	(0.0495)
Individual controls	Yes	Yes	Yes
F-test (IV)	105.852	105.852	105.852
Hausman test (p-value)	.035	.0834	.6471
N	25026	25026	25026

Note: The list of individual controls includes constant, a set of five-year age dummies, female, educational attainment, and country fixed effects. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$. Preliminary SHARE wave 9 COVID-19 Survey 2 release 0 data.

E. Does family size impact preventive behavior?

Table E.1: Family size and preventive behavior

	PCA in Summer 2021 Component 1	PCA in Summer 2021 Component 2	PCA in Summer 2021 Component 3	PCA in Summer 2020 Component 1	PCA in Summer 2020 Component 2
	(1)	(2)	(3)	(4)	(5)
Dummy more than two	0.257	0.0778	-0.0423	0.145	0.190
	(0.303)	(0.177)	(0.168)	(0.327)	(0.172)
Individual controls	Yes	Yes	Yes	Yes	Yes
F-test (IV)	100.653	100.653	100.653	105.605	105.605
Hausman test (p-value)	.623	.901	.894	.862	.213
N	23732	23732	23732	24524	24524

Note: All columns control for constant, a set of five-year age dummies, female, educational attainment, and country fixed effects.* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$. Preliminary SHARE wave 9 COVID-19 Survey 2 release 0 data.

F. Intake of other vaccines

Table F.1: Family size and intake of flu and pneumonia vaccines

	Got flu vaccine	Got pneumonia vaccine
	(1)	(2)
Dummy more than two	-0.0884	0.0694
	(0.0921)	(0.0655)
Individual controls	Yes	Yes
F-test	105.852	105.852
Hausman test (p-value)	.4782	.2093
N	25026	25026

Note: The list of individual controls includes constant, a set of five-year age dummies, female, educational attainment, and country fixed effects. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$. Preliminary SHARE wave 9 COVID-19 Survey 2 release 0 data.