



Nursing homes and mortality in Europe: Uncertain causality

Xavier Flawinne, Mathieu Lefebvre, Sergio Perelman, Pierre Pestieau, Jérôme Schoenmaeckers

► To cite this version:

Xavier Flawinne, Mathieu Lefebvre, Sergio Perelman, Pierre Pestieau, Jérôme Schoenmaeckers. Nursing homes and mortality in Europe: Uncertain causality. *Health Economics*, 2023, 32 (1), pp.134-154. 10.1002/hec.4613 . hal-03807685

HAL Id: hal-03807685

<https://amu.hal.science/hal-03807685>

Submitted on 10 Oct 2022

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

Nursing Homes and Mortality in Europe: Uncertain Causality

Xavier Flawinne^{*}, Mathieu Lefebvre[†], Sergio Perelman[‡], Pierre Pestieau[§] and Jerome Schoenmaeckers^{**}

Abstract

The current health crisis has particularly affected the elderly population. Nursing homes have unfortunately experienced a relatively large number of deaths. On the basis of this observation and working with European data (from SHARE), we want to check whether nursing homes were lending themselves to excess mortality even before the pandemic. Controlling for a number of important characteristics of the elderly population in and outside nursing homes, we conjecture that the difference in mortality between those two samples is to be attributed to the way nursing homes are designed and organised. Using matching methods, we observe excess mortality in Belgium, Sweden, Germany, Switzerland, Estonia and Czech Republic but not in France, the Netherlands, Denmark, Austria, Italy and Spain. This raises the question of the organisation and management of these nursing homes, but also of their design and financing.

Keywords: Nursing homes, mortality, propensity score matching, SHARE

JEL codes: C21, I10, J14

^{*} University of Liège.

[†] Aix Marseille University, CNRS, AMSE, Marseille, France.

[‡] University of Liège.

[§] University of Liège; PSE.

^{**} University of Liège; CIRIEC Belgium.

1. Introduction

With the current COVID-19 pandemic, the high mortality of elderly living in nursing homes has been widely debated. According to Canadian Institute for Health Information (CIHI, 2020), the number of COVID-related death in nursing homes during the first wave in spring 2020 represented about 66% of total COVID-related death in Spain, 48% in France, 34% in Germany and only 15% in the Netherlands. Although these numbers do not allow to draw any causal effect, it has been argued that due to the low quality of care and the physical proximity of nursing homes residents, life expectancy was lower than in private homes. This has pointed out that the differences across countries in terms of quality and the institutional features of the nursing homes could be an explaining factor of differential mortality observed in residential housing across European countries.

Using data from the Survey of Health, Ageing and Retirement in Europe (SHARE⁶), the aim of this paper is to determine whether there is excess mortality for nursing homes residents compared to those who live in their homes. Our analysis takes place in years before the COVID 19 pandemic in order to look at the effect of nursing home in normal times. We use propensity score matching in order to construct a sample in which treated (being in a nursing home) and untreated individuals (living at home) have similar characteristics in terms of age, gender, degree of dependence, state of health, availability of informal help (partner and children) and assets. We conjecture that, after controlling for the determinants of entry into a nursing home, the difference in mortality between those two samples is to be attributed to the way the nursing homes are designed and organised or alternatively to the quality of aid and services one finds staying home. This allows us to determine whether there is a causal link of being in a nursing home on the probability of death.

Excess mortality due to age, health or dependency should not be a cause for concern when looking at mortality within nursing homes. It is a societal choice that induces senior citizens to end their life in nursing homes. We could reflect on this choice but it is not the subject of this paper. We argue that, if this excess mortality is due to the structure and organisation of nursing homes, then there is room for reform. It is difficult to distinguish between the two types of factors with SHARE data, because information on the characteristics of nursing homes, housing and staff is not available. We therefore start from the idea that each country has its own conception of what the specific objectives of a nursing home should be: care, entertainment, socialisation, ensuring privacy, etc. These objectives may vary from one country to another. Moreover, their achievement will depend on the resources that are devoted (average expenditure) and the more or less efficient way in which they are used (productive efficiency). We are well aware that we do not necessarily have data to evaluate these factors of "design,

⁶ 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.w8cabeta.001), see Börsch-Supan et al. (2013) for methodological details. (1)

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).

spending and efficiency", but if we have two comparable samples, one from individuals in nursing homes and one from individuals at home, comparing both could give some insights on the *mortality ratio*. If the nursing homes sample has a higher mortality rate, this excess mortality can be attributed to them without necessarily knowing the determinants of this difference.

Although many studies have investigated the determinants of the choice of housing at old age as well as the quality of nursing homes, few studies have provided causal evidence on mortality (see the next section dedicated to a review of the empirical literature on this topic). One important aspect of our study is to compare the situation in several countries and thus to identify institutional and organizational differences that may lead potential reforms. While the choice of housing at old-age is often driven by price and different individual characteristics (health status, daily life limitations, loneliness...), the predictors of nursing home entry may also be diverging across countries because of country-specific institutional alternatives (Angelini and Laferrère, 2012; Laferrère et al., 2013; Schmitz and Stroka-Wetsch, 2020).

Our results show an overall negative impact of being in a nursing home on life expectancy. Looking at European countries separately, we observe excess mortality in Northern, Central and Eastern European countries of our sample. In particular, we find excess mortality in Sweden, Belgium, Germany, Switzerland and Czech Republic. We also find no statistically significant excess mortality in the Southern countries: Italy and Spain. This allows to draw preliminary conclusions about potential differences in terms of the quality of these care facilities and the consideration given to nursing homes. Looking at possible mechanisms behind these results, we identify the role of public spending for long-term care and the overall resources devoted to nursing homes. Furthermore, the share of for-profit nursing homes in a country seems to be correlated with the effect on mortality. In order to confirm our matching empirical strategy and these results, we investigate the effect of the violation of the conditional independence assumption by simulating the effect of relevant unobserved confounders affecting both the treatment and the outcomes of interest. It results that the existence of such confounders is unlikely to affect the results.

Although our findings are obtained for data before the COVID-19 pandemic, they give some insights about the excess mortality observed in nursing homes during the first wave of the pandemic (Cronin and Evans, 2022; Chen et al., 2022; Schultze et al., 2022). However, one striking result is the absence of excess mortality in Spanish and Italian nursing homes. Though evidence during the pandemic shows an excess mortality due to COVID-19 in Italy (Alacevitch et al., 2021) and Spain (Costa-Font et al., 2021). We also discuss the mechanism behind these differences.

After a quick review of the existing empirical literature in Section 2, Section 3 presents the data and the selected sample. Descriptive statistics are also presented. Section 4 presents the propensity score matching method used to overcome selection bias along with a sensitivity analysis performed to test the robustness of our estimation assumptions. Results are presented in Section 5 and Section 6 discusses the mechanisms that can explain the effect of being in a nursing home on mortality. Section 7 concludes.

2. Literature review

This paper is closely related to the important literature that deals with long-term care (LTC) issues. Especially, as pointed by Spasova *et al.* (2018), LTC provision in Europe has been characterised by significant differences between (and within) countries, mainly in the way it is organised (by public, for-profit or non-governmental providers), delivered (home care versus institutional care), financed (cash benefits, in kind benefits or out-of-pocket payments) and how resources are generated (via general taxation, mandatory social security and voluntary private insurance). One important question is how this care is provided. The role of informal care across Europe has been highlighted in extensive research (see Klimaviciute *et al.*, 2017 for a survey) and the substitution between formal and informal care has also been studied (Van Houtven and Norton, 2004; Bolin *et al.*, 2008; Bonsang, 2009). In particular, previous findings suggest that informal care substitutes for nursing home entry (Lo Sasso and Johnson, 2002; Charles and Sevak, 2005). Another issue that is directly related to our paper is that the place of residence is likely to have an impact on the type of care that is received by the elderly. Degavre and Nyssens (2012), by comparing Belgium, England, Germany and Italy, have shown that there are structural differences across countries in terms of the organisation of LTC received at home but also in institutions.

Several studies have tried to determine the factors that induce people to enter nursing homes. Using European data, Angelini and Laferrère (2012) and Laferrère *et al.* (2013) attempted to identify predictors of nursing home entry based on country-specific institutional alternatives. They point to the importance of functional limitations, the absence of partner or the low socio-economic status in the decision (or the need) to enter a nursing home. The impact of low assets on the probability of entering a nursing home has also been pointed out in the US context by Lindrooth *et al.* (1993) who also showed that expectations about nursing home entry were reasonably close to the actual probability of entry. Using German data, Schmitz and Stroka-Wetsch (2020) find that the probability of choosing a nursing home depends on distance from home and price. However, they find no economically significant effect of reported quality on individuals' choice of nursing homes. In a recent study, using longitudinal data from SHARE, Laferrère and Arnault (2021) show that for given levels of health, disability, living arrangements and housing conditions, the more educated and wealthy individuals are less likely to move into a nursing home. This ex-post “objective check” corroborates the idea that the desire to “age in place” is the preferred option, as opposed to living in a community, such as a nursing home. The quality and the organisational structure of the nursing homes, that often diverge across countries and welfare systems, are also determinants in the choice of housing.

Our paper is also related to the literature dedicated to identify the factors of mortality or excess mortality, if any, within the nursing homes. Depending on the countries and the methods used, several authors show that co-morbidity is the main marker of mortality, while others point to the primary role of functional limitations. Both definitely play a role on mortality in nursing homes. Braggion *et al.* (2020) pointed out that the first months after admission represent a period at high risk of mortality, especially for patients with a recent hospitalization. The quality of the nursing home institutions has also been pointed out for its impact on mortality. Using a sample of California nursing homes, Antwi and Bowlblis (2018) have shown that the nursing turnover has a positive effect on mortality. Lin (2014) also showed that nurse staffing has a large and significant impact on the quality of care and thus on health outcomes. Recently,

Cronin and Evans (2022) highlighted the role of quality of the nursing home institutions on mortality during the pandemic. Quality does not predict the ability to prevent any resident or staff COVID-19 cases, but higher-quality establishments prevent the spread of resident infections conditional on having one. In the UK, Schultze et al. (2022) pointed out an increase of men mortality in nursing homes due to the first wave of COVID-19 pandemic. Studies on the link between respiratory illnesses and mortality in nursing homes did not originate with COVID-19. Using a population-based national survey in US nursing homes, Beck-Sague et al. (1993) have demonstrated that residents with pneumonia were more likely to die than those with other infections. They also pointed out that lack of mobility (functional status) is a risk factor, as does Sung (2014) who also highlights the role of the dyspnea for elderly Korean people. The role of respiratory diseases on mortality in nursing homes and co-morbidity in general is clarified and debated in various research studies (Sung, 2014; Levy et al., 2015; Falcone et al., 2018).

Our research goes one step further and seeks to determine whether there is a difference in mortality due to the very fact of being in a nursing home. There is relatively limited research on the specific issue of the role of the institutionalization on potential excess mortality. Based on longitudinal French data on people aged 60 and more, Giudici et al. (2019) have estimated that people living in institutions live on average 10 years less than those living in private households. However, they did not take into account the potential issue of endogeneity, as people in nursing home differ from people staying at home. On the contrary, relying on Dutch administrative data, Bakx et al. (2020) found nil impact of being in a nursing home on mortality, although with large heterogeneity of effects among the population. Also, they do not observe a decrease in financial costs if care is provided at home rather than in a nursing home (especially following hospital admissions). Aging-in-place policies would come at the cost of increased curative care, especially hospital admissions, but would not reduce total healthcare spending, suggesting they may not be a win-win after all (when compared to the costs incurred in institutions).

3. Data & descriptive statistics

The Survey of Health, Aging and Retirement in Europe (SHARE) is an international, interdisciplinary and longitudinal survey of people aged 50 and over. About 380,000 in-depth interviews with 140,000 people aged 50 or older from 28 European countries and Israel have been conducted. Every two years since 2004, the questions relate to the health status (objective and subjective) of the respondents and their potential spouses, as well as to the economic and social situation of the household, including relations with their children and close relatives.

In this paper we use the last four waves 4, 5, 6 and 7 of the survey because earlier waves did not include nursing homes residents⁷. The nursing home is clearly defined in SHARE: “*A nursing home provides all of the following services for its residents: dispensing of medication, available 24-hour personal assistance and supervision (not necessarily a nurse), and room &*

⁷ Nursing home residents have always been eligible to participate in SHARE. Interviewers began to register the respondents' place of residence as either a private or a nursing home since wave 2 of the study. However, since the wave 4, it was decided to allow sample (refreshers) also in nursing homes residents to participate, not only after a transition but as a first participation (refresher sample) to the survey. The questionnaire is adapted. Sampling of nursing home residents is not compulsory but and in particular only a few countries recording undercoverage by design.

meals”. Because our research question is to determine whether there is a causal link from being in a nursing home on mortality, our sample is composed of people for whom we have information on the place of residence at time t but also on the status (alive or dead) at time $t+1$. We thus look at the transitions from wave 4 to wave 5, from wave 5 to wave 6 and from wave 6 to wave 7 and in our analysis we pool these transitions together⁸. Because the health status and the choice of housing may be simultaneously determined, we restrict our sample of analysis to individuals aged 65+ who present at least one limitation in activities of daily living (ADL)⁹. The use of ADL limitations for defining our sample has two main advantages. First, these are rather objective indicators of dependency for which the measurement can be controlled by the survey interviewer. Second, by taking a measure of physical limitations, we focus on a factor that may explain the difference in use of a nursing home or not.

We also drop countries for which the time frame “ $t / t+1$ ” is not respected (Hungary for instance) or because there are too few observations in nursing homes. We have made an arbitrary choice to eliminate those countries for which, once the observations had been pooled, fewer than 25 surveyed participants were registered in nursing homes. This is the case for Greece (2), Poland (3), Croatia (4), Portugal (19), Slovenia (22) and Israel (25). The resulting data set contains 13,340 observations among which 863 observations correspond to individuals living in nursing homes. Let us add that the rate of nursing home residency in SHARE is close to what we can observe in the general population¹⁰. Based on the European Census of 2011 (Eurostat, 2016), Schanze and Zins (2019) have shown the share of individuals residing in nursing homes in SHARE is close to the share of institutionalized residents in the general population. Table 1 summarises our sample selection in each country, restricted to people aged 65+ and dependent at time t . The grey areas in the table illustrate the absence of the SHARE survey for the country and the wave concerned.

Insert Table 1: SHARE waves and selected observations

When we look at our sample of elderly with at least one limitation, we observe differences in the rate of nursing home residency: while only 1.8% of Italians reside in a nursing home, more than 16% of Luxembourgiens or Danes are registered in institutions. More than 10% of the Belgians and Swedes in our sample are also in nursing home. On average, 6.5% of the respondents in our sample are in nursing homes.

⁸ The successive waves of SHARE are not taking place every year. There is an average of two years between each interview so that, here, $t+1$ should be understood as wave + 1. The distribution of the sample in terms of transition is as follows: 29.1% from wave 4 to wave 5, 37.0% from wave 5 to wave 6 et 33.9% from wave 6 to wave 7.

⁹ A person is usually considered dependent if two or more difficulties in performing the following daily tasks (ADLs) are met: 1) dressing, including putting on shoes and socks; 2) walking across a room; 3) bathing or showering; 4) eating, such as cutting up your food; 5) getting in or out of bed; 6) using the toilet, including getting up or down. We decide to deviate a little from this threshold and to consider people with at least one limitation.

¹⁰ Appendix A.1. summarizes information about gross data (before restriction to people aged 65+ and having at least one limitation in activities of daily living) and representativeness of nursing homes respondents’ samples. In our data, there is some heterogeneity in European countries in terms of the use of nursing home-type institutions (from 0.4% of the 65+ in Italy to more than 4% in Luxembourg). On average, 1.6% of the 65+ population is a nursing home resident. The reasons for entering in a nursing home (care, entertainment, socialisation, ensuring privacy, etc.) may vary from one country to another. A first glimpse at the people with at least one daily limitation and living in nursing homes shows the variation in the actual design of the nursing home across the different European countries. We observe, for example, that only 42% of Dutch residents in nursing home have at least one limitation while this number climbs to more than 80% in Denmark and Spain (the sample average is 66%).

Table 2 provides a description of our two main variables of interest, i.e., living in a nursing home (as defined previously) or at home at time t and being alive or not at time $t+1$. In total, we observe a probability to die of 20.2% in $t+1$. However, this rate hides very different situations according to the place of living with 41.5% of the elderly living in a nursing home being dead in $t+1$ while it is 18.8% of the elderly living at home. When we look at countries separately, we note a variation both in terms of general mortality (from 11.5% in the Netherlands to 25.4% in Spain and 26.3% in Denmark) but also in terms of mortality conditional on housing type. It is interesting to look carefully at the *mortality ratio*, which is the proportion of people who died while in nursing home (NH) over the proportion of people who died at home (AH). This varies between 1.0 in Italy and almost 3.6 in Switzerland, for an average value of 2.2 in our sample.

Insert Table 2: Mortality rates ratio in SHARE countries (65+ & 1+ ADL)

However, these first descriptive results need to be taken cautiously since they do not take into account the potential selection bias, i.e., that people in nursing home differ from people staying at home. Also, the sample sizes do not allow us to investigate heterogeneity within countries and the fact that the populations using nursing homes in Germany differ fundamentally from the Belgian or Spanish populations. It might be the case that depending on region, education or former occupation, these statistical trends diverge between countries because they are the result of different histories and institutional frameworks. However, the aim of our study is to assess if there is excess mortality in some countries and not in others and suggest reasons for it.

Table 3 summarises the information on the socio-economic, health and family variables according to the type of dwelling. We observe important differences on average between the two subsamples. This also motivates the use of propensity score matching method to control for the differences between the two subsamples of dependent elderly and determine a potential impact of nursing homes on mortality. The table should be read as follows: among the people in nursing homes, 12.4% are between 65 and 74 years old and 65.7 % are low educated. Most of residents of nursing homes are in the 1st tercile of wealth (these terciles have been created by country and by wave on the whole population present in the survey). 69.1% of nursing home respondents are women when they are only 60.4% at home residents. Besides, we note that women are more represented in our sample. They are 61.0%. When we look at the health characteristics, we observe that 38.5% of nursing home residents have 5 to 6 limitations in their daily life activities while 56.8% have more than 6 limitations in the instrumental ones (IADLs¹¹). Finally, differences in housing type according to the number of chronic conditions appear to be less important and even go in an a priori surprising direction (39.8% of people at home have 4 chronic diseases or more when they are 31.4% in nursing homes)¹².

Insert Table 3: Summary statistics of covariates (65 + & 1+ ADL)

¹¹ IADL limitations are the following: 1) Using a map to figure out how to get around a strange place; 2) Preparing a hot meal; 3) Shopping for groceries; 4) Making telephone calls; 5) Taking medications; 6) Doing work around the house or garden; 7) Managing money, such as paying bills and keeping track of expenses; 8) Leaving the house independently and accessing transportation services; 9) Doing personal laundry.

¹² Depending on medical needs, it might be the case that nursing homes are not well equipped to welcome patients with severe pathology like cancer. A two-sample t test was performed based on the means of each group (NH: 2.92 vs AH: 3.24) and the difference was statistically significant (0.32, with S.E. = 0.066).

Finally, we notice and highlight the important role of the spouse and children in the LTC issue and their status as informal carers. Indeed, only 7.6% of people in nursing homes have still a partner and they are less likely to still have a living child than those who remain at home (78% vs. 89.7%). They have also fewer daughters that live close to their (9.5% vs. 16.2%). These descriptive results are a good reminder of the role of family carers on help given at home. These carers are therefore essentially the partner, particularly the wife helping her dependent husband, and the daughter, helping her dependent parents and mainly her mother once she is widowed (see Canta et al., 2021).

4. Methods

In order to evaluate the causal effect of nursing homes on mortality, it is important to control for the possible simultaneous determination of health and housing arrangement. This is why we select a sample of individuals aged 65+ who have at least one ADL limitation but, as presented in the previous section, the characteristics of people in nursing homes can differ significantly from those still at home. Therefore, to control for the selection bias due to observables, we use a propensity score matching method. That is, every individual in a nursing home is matched to a set of individuals living at home with similar observable characteristics. It allows us to condition on sufficient observable information to obtain a counterfactual against which we can measure the effect of being in a nursing home¹³.

Following the common terminology of the method, being in a nursing home is assimilated to being treated. We thus need to find individuals who are similar to treated individuals in terms of pre-treatment characteristics and are not in nursing home. The goal is to match similar individuals from the two groups so that the differences in outcomes of these matched pairs can then be attributed to the treatment, i.e. being in a nursing home. Our outcome of interest is the mortality at time $t+1$ while the treated units are residents of nursing home and control units are people still living at home. The main advantage of the method is that it does not require specifying a functional form of the outcome equation and is therefore not susceptible to misspecification bias along that dimension. The matching method goes further since the idea is to compare two individuals who, based on observables, have a very similar probability of being treated, but one of them had received treatment and the other did not.

4.1. Propensity scores and matching

Our analysis relies on the conditional independence assumption (CIA) that the mortality of the individuals in the control group and in the treated group are independent of the residence status once we control for a set of observable characteristics. As shown by Rosenbaum and Rubin (1983, 1985), instead of conditioning the matching on the whole set of individual characteristics, it suffices to concentrate on a summary index, a balancing score. The most prominent balancing score is the conditional probability of selection into treatment, i.e. the propensity score of being into a nursing home.

To obtain propensity scores, we use Probit regressions where the dependent variable is being in a nursing home and covariates influencing simultaneously the fact of being in a nursing home and the probability of dying at time $t+1$, are gender, age, partnership situation, country wealth

¹³ See Imbens (2015) for a detailed presentation of the matching method or Angrist and Pischke (2008) for a popularized explanation.

terciles, educational levels, number of ADLs, IADLs and chronic diseases. We also introduce country dummies as well as the fact of having at least one child and of having a daughter that is not working and lives close. As presented in the literature review, these are strong predictors of entering a nursing home¹⁴. The estimations of propensity scores are done for the total sample, for groups of countries and for each country separately. Table A.2. in the Appendix presents the results, controlling for the different waves. Estimations achieve balance on covariates between treated and controls¹⁵. Probit regression results of residential status on covariates are only reported for the overall sample and for groups of countries¹⁶. They show that being a woman, being wealthy, living with a partner and having at least a child is negatively associated with entering a nursing home. Having a daughter that is not working also decreases the probability of going into a nursing home, as a higher number of chronic diseases (as explained before). Being older or having a high number of ADL or IADL limitations increase the probability of residing in a nursing home. However, there are differences between regions. In terms of socio-economic characteristics, the correlations between respectively gender, age and education are not statistically significant with going into a nursing home in South and East. Age does not seem to play a significant role also in North. Finally, the most interesting results come from family characteristics. While having a partner reduces the likelihood of entering a nursing home in all country groups, the role of children varies greatly across regions. Almost nil in the North and Central regions, they have a considerable impact on nursing home entry in the South and East.

As usual with matching analysis, there is a clear trade-off between bias and efficiency when it comes to choosing a matching algorithm. In our main analysis we match observations using the Kernell matching method with replacement. That is, weighted averages of individuals in the control group are used to construct the counterfactual outcome of the treated individuals. This method has the advantage of reducing the variance that is achieved since more information is used compared with other matching methods. To avoid matches that could bias our analysis, we request that the pairs are perfectly matched according to country, gender, marital status, having at least one child, having a daughter that is not working and having a daughter living at less than 25 kilometers by applying exact matching on these variables¹⁷. However, it possibly uses observations that are not very good matches so in addition, as robustness checks, we also present in the Appendix estimates using another matching method based on Mahalanobis distance¹⁸ (see discussion of King and Nielsen, 2019) but also cross-validation with respect to the means of the dependent variables (X) and the outcome one (Y). Cross-validation is a

¹⁴ Other covariates could not be kept for two reasons. Either they invalidated one of the conditions of validity of the empirical method (i.e., including only covariates that influence simultaneously the treatment (nursing home) and the outcome (mortality) or there were too few observations in the SHARE data (for instance social networks variables are only available for waves 4 and 6 of SHARE or number of house moves during life are only available for SHARELIFE respondents).

¹⁵ If the balancing property is not satisfied, the treatment and comparison groups are unlikely to be sufficiently similar to reduce selection bias in your treatment effect estimate.

¹⁶ The estimations of propensity scores are done by country for the estimations of ATT proposed at country level.

¹⁷ These variables correspond to all the binary variables included in the analysis.

¹⁸ Mahalanobis distance matching (MDM) works by pairing units that are close based on a distance called the Mahalanobis distance, which is like a scale-free Euclidean distance. For two units to have a Mahalanobis distance of 0, they must have identical covariate values. The more different the covariate values, the larger the Mahalanobis distance. The idea is that if you find control units close to the treated units on the Mahalanobis distance, each pair will have similar covariate values, and the distribution of the covariates in the treatment groups in the matched sample will be similar.

convenient approach to selecting the bandwidth value and turns out to work reasonably for some estimators as the Kernell one (Frölich, 2004). The two estimation methods each have their advantages and disadvantages (see discussion of Rippolone et al. (2018) on the trade-off between the two) but lead broadly to the same results.

4.2. Sensitivity analysis

The plausibility of the CIA relies on the possibility to match treated and control individuals on the basis of an informative set of observed variables. In order to assess whether our average treatment effects are robust to possible deviations from the CIA, we implement a simulated sensitivity analysis as proposed by Ichino et al (2008). The idea is to assume that the CIA is not satisfied given the considered observables but would be if one could observe an additional binary variable. This potential confounder can be simulated in the data and used as an additional covariate in combination with the matching estimator. By comparing the results obtained with and without matching on the simulated confounder, we can show to what extent the baseline results are robust to specific sources of failure of the CIA¹⁹.

The assumption of the analysis is that the CIA no longer holds given the set of covariates X but it holds given X and an unobserved binary variable U . This means that as long as U is not observed, the outcome (i.e. the mortality) of the control individuals cannot be used to estimate the counterfactual outcome of the treated individuals. We assume that U may impact both the treatment and the outcome and that the distribution U can be fully characterized by four probabilities p_{ij} given the treatment T (being in a nursing home) and the outcome Y (being alive or not):

$$p_{ij} = \Pr (U = 1|T = 1, Y = j)$$

with $i, j \in \{0, 1\}$, which give the probability that $U = 1$ in each of the four groups defined by the treatment status and the outcome value. Given arbitrary values of the parameters p_{ij} , a value of U is attributed to each individual according to its belonging to one of the four groups defined by the treatment status and the outcome value. U can then be treated as any other observed covariate and is included in the set of variables used to estimate the propensity score and to compute the effect of the treatment (i.e. being in a nursing home).

The difference $d = p_{01} - p_{00}$ can be interpreted as a measure of the effect of U on the untreated outcome, and the difference $s = p_1 - p_0$ ²⁰ as a measure of the effect of U on the selection into treatment. As explained in Ito et al. (2018), one can define the selection effect Λ as the effect of U on the relative probability to be assigned to the treatment and the outcome effect Γ as the effect of U on the relative probability to have a positive outcome in the absence of treatment.

$$\Lambda = \frac{\frac{\Pr (T = 1|U = 1, X)}{\Pr (T = 0|U = 1, X)}}{\frac{\Pr (T = 1|U = 0, X)}{\Pr (T = 0|U = 0, X)}}$$

and

¹⁹ See Nannicini (2007) for a detailed presentation of the simulation method.

²⁰ The expression p_1 and p_0 correspond to the probability of being treated (in a nursing home) given the value of U and controlling for the set of covariates W : $p_1 = \Pr (T = 1|U = 1, W)$ and $p_0 = \Pr (T = 1|U = 0, W)$.

$$\Gamma = \frac{\frac{\Pr(Y = 1|T = 0, U = 1, X)}{\Pr(Y = 0|T = 0, U = 1, X)}}{\frac{\Pr(Y = 1|T = 0, U = 0, X)}{\Pr(Y = 0|T = 0, U = 0, X)}}$$

By measuring the two effects Γ and Λ , one can characterize the simulated confounder U . If $\Gamma > 1 (< 1)$, it means that the unobserved U positively (negatively) affects mortality. Similarly, if $\Lambda > 1 (< 1)$, it means that the unobserved U increases (decreases) the probability of being in a nursing home.

Ichino et al. (2008) propose two approaches to pick the parameters p_{ij} . One approach makes the assumption that the distribution of the unobserved variable U is similar to the empirical distribution of important binary covariates. We thus fix p_{ij} according to their values for a set of covariates used in the propensity score model. A second approach aims at searching for values for p_{ij} such that if U were observed, the estimated average treatment effect would be driven to zero. If the parameters leading to such a result can be considered unlikely, the exercise support the robustness of the estimates derived under the CIA.

5. Results

5.1. Effects of being in a nursing home

Table 4 presents the Average Treatment of the Treated (ATT) estimates for the total sample, by group of countries and for each country separately. Group of countries allows us to compare group of similar characteristics but avoiding the small number of observations we obtain when looking at a single country. We present the results according to the Kernell matching method but results obtained with other matching methods are presented in Table A.3. in the Appendix. Estimates are qualitatively similar when looking at the sample of 13 countries altogether. The results at the level of the regions are also stable, with the exception of the results for the North²¹. We will discuss this particular case later and use the methods proposed by Ichino et al. (2008) to "adjudicate" this issue.

Overall, the effect of being in a nursing home on excess mortality is positive and significant at 1% level. We find that for the 13 countries consolidated sample, elderly in a nursing home are almost 10.7 percentage points more likely to die in the next wave of the survey than those staying at home. The full sample mean is 20.2 implying an increase of about 53%. Although we test the validity of the CIA in the next subsection, we also looked at the effect of our various covariates on the ATT. The Figure 1 shows how much each of the existing observables impact the treatment effect. We present the evolution of the ATT after the cumulative addition of socioeconomic status, then health, and finally informal help availability (in relation to the presence and type of children) variables.

Insert Figure 1: Cumulative addition of covariates and estimated ATTs

²¹ It is normal for the number of people in the treatment or control sample to differ depending on the estimation method (Kernell PSM or MDM) or the bandwidths chosen (by default 1.5 times the 90% quantile of the (non-zero) distances in pair matching with replacement for the MDM method) or with cross-validation w.r.t the means of X or w.r.t. Y .

We observe a U-shaped curve for the effect on all countries. The effects always remain statistically significant but decrease to a minimum (0.045*) with the addition of the chronic diseases. The final global effect rises to 0.107*** once we take into account the binaries of presence of children and particularly of daughters. However, these results may be misleading because they are potentially comparing very different population, e.g. an Italian and a Danish woman, who are known to live in different institutional situations. In the rest of Table 4, we present estimates for each country separately.

Insert Table 4: ATT estimation with the Propensity-Score Kernel Matching (exact matching with replacement)

With the exceptions of southern European countries, in all other regions, the individuals who live in a nursing home are more likely to die than those who are in their own home. The size of the effect is similar in the northern, central and eastern countries. The small number of observations makes country-specific analysis uneasy but keeping in mind this caveat we see that with the exceptions of Italy (and Spain), in all countries, the individuals who live in a nursing home are more likely to die than those who are in their own home. However, these positive differences are only significantly different from zero in Belgium, Sweden and Estonia at 10% level and in Czech Republic at 5%. It is highly significant in Germany and Switzerland. The highest effect is observed in Germany and the lowest in Belgium. In the Netherlands, Denmark, Austria, France and Luxembourg but also in Spain and Italy, the estimated ATT is not statistically significant.

These differences between countries will be further investigated in Section 6 but sample sizes do not allow us to investigate the heterogeneity of the effects within the countries. We do this at the regional level in Appendix A.4. where we perform the same analysis as in Figure 1, adding the covariates step by step. In the central European countries, there is continuity in the significance of the positive effect. A U-shaped curve also appears but is less pronounced. The addition of child-related variables has little impact on the ATT coefficient. For the South, the significant effect disappears as soon as the wealth is taken into account in our different matchings. The absence of nursing home effect on mortality seems quite robust. For the North and East, the effect is no longer significant (it can be positive or negative) once the instrumental limitations variable (IADLs) is added. The positive and significant effect reappears once child-related characteristics are taken into account. For the North, the effect appears with the addition of the last variable, i.e. having a daughter living within 25 km. This result for the North is therefore rather unstable and the inclusion of another variable (omitted until now) could make the observed positive effect disappear. Therefore, in order to confirm these results, we conduct a sensitivity analysis such as presented in Section 4 and we remove the uncertainty about the ambivalent outcome obtained for the North.

5.2. Sensitivity analysis

We conduct the sensitivity analysis suggested by Ichino et al. (2008) to test whether our results are robust to the violation of the CIA. Indeed, one may think of one unobserved variable that would simultaneously influence the decision to go into a nursing home (selection effect, Λ) and the probability to die (outcome effect, Γ). For example, actual informal care may influence simultaneously the decision to enter a nursing home and the health of the elderly. The former would be related to a selection effect and the later would have an outcome effect.

In order to investigate how sensitive our estimates are with respect to the possible existence of this unobservable variable; we perform two simulation exercises. In a first step, we simulate an unobserved variable which would have a distribution similar to the empirical distribution of important binary covariates. Table 5 presents the results for four binary covariates giving important selection and outcome effects: being a woman, having at least one child, having a daughter not working or having a daughter living at less than 25 kilometers. This does not confound our results and the ATTs for the total sample are very close to the ones presented in Table 4. The selection effect and the outcome effects differ according to the simulations and do not deviate too much from 1.

We also perform the same analysis for each group of countries (region) separately by simulating an unobserved variable which would have a distribution similar to binary covariates. Table A.5. in the Appendix presents the results for cofounder like being a woman, having one child, having a daughter not working or having a daughter living at less than 25 kilometers. The ATT is close to the baseline and overall, we find that any unobserved variable with similar treatment and selection effects as those four covariates already introduced in the propensity score matching do not confound our results, except for the North where the positive effects disappear and become negative but very close to zero.

Insert Table 5. Sensitivity analysis: confounder-like and killer confounder

Table 5 presents also the results of the second simulation exercise where we look at a set of distribution parameters such that the size of the outcome and selection effect of the unobserved variable would kill our results. That is, we want to find parameters such that if the confounder were observed the estimated ATT would be driven to zero²². To reduce the dimensionality of the problem, we fix $p_{11} = p_{10} = 0.9$. Since these quantities are not expected to represent a real threat for the baseline estimate, they can be held fixed and the simulated confounder can be fully described by the differences d and s . We make d increase by 0.1 and let s being automatically estimated with respect to d . The values of s and d are associated with the estimated values of Λ and Γ , respectively. Table 5 displays some examples of outcome and selection effects for which our main result would disappear²³. Results show that in order to find an effect of being in a nursing home on the probability to die that tends to zero, the potential confounder should have an outcome effect and a selection effect that are much higher than what we observe in the covariates distribution. In order to “kill” our results, the outcome and selection effects should be from 6 to 15 times bigger which is very implausible.

6. Mechanisms

The difference in terms of causal effect of being in a nursing home on mortality across countries is an important result. One may wonder what the mechanisms at work behind these effects are. Can we find country-specific channels or characteristics that might explain these results?

There are potentially several mechanisms behind those results. One may think of differences across countries in terms of level of general health but we already control for health level and individual limitations. Other explanations may consist of differences across countries as to long-term care such as for example the presence of informal care, the quality of the nursing

²² For the sake of simplicity, we do not present the second simulation by group of countries but results by region are available upon request and give similar conclusions.

²³ For $d = 0.1$ and until $s = 0.48$, for $d = 0.2$ and $s = 0.36$ or for $d = 0.3$ and $s = 0.24$, we still observe a significant effect of being in nursing home on the probability of dying next wave.

homes or their access. Indeed, European countries are rather different in the way they organize the care of the elderly. Unfortunately, we do not have micro data on nursing homes that would allow looking in details for differences across countries. However, Table 6 reports aggregated indicators which shed light on differences across countries on formal and informal long-term care. Being careful about issues of reverse causation and thus without concluding of any causal effects, we present some empirical evidence that may be interpreted as supporting mechanisms behind the cross-country mortality difference in nursing homes.

Table 6 shows that countries from Central Europe in our sample, where the effect on mortality of being in a nursing home is the highest, present some specific features compared to the other countries.

Insert Table 6: Information about Formal and Informal Care by country

In particular, they show a mix of low average public spending (from 1.6% of GDP in Germany to 2.4% in Switzerland), a low number of LTC workers per 100 individuals (particularly in France) and a large proportion of for-profit nursing homes (from 22% in France to 40% in Germany or Switzerland). These figures tend to show a lower concern for the elderly, which could be associated with lower quality. On the contrary, countries of the North of Europe (Denmark, the Netherlands and Sweden) devote more resources to long-term care than the other countries. The Netherlands and Sweden have also the highest number of workers per 100 people over 65 years old and the highest number of beds per 100,000 inhabitants. These countries are also characterized by a small proportion of private nursing homes (less than 20% of the available supply). Without concluding as to any causal effect, these countries do not present excess mortality pertaining to being in a nursing home. Also, data on informal care in the Northern countries shows a high proportion of the population providing care but for a rather small number of hours; actually the smallest number of our sample. In these countries, help to the elderly is widespread but is rather a complement to formal care largely provided by public services (Klimaviciute et al., 2017).

In Italy and Austria, public spending (as a share of GDP) dedicated to formal long-term care is similar to what we observe in Central Europe but with much more assistance in terms of cash benefits. Also, intensive informal support is widespread. The situation seems to be more heterogeneous in the East with the majority (80%) of nursing homes in Estonia being for profit, this associated with low state intervention (0.4% of GDP). Informal care is more intensive but less frequent in the Czech Republic than in Estonia.

One interesting finding from this exploratory analysis is the association between health outcome and the for-profit status of the nursing homes. ENCO (2021) identified 14 private companies in the elderly care sector in Europe and provided an overview, country by country, of the institutional statuses of nursing home beds. Behind the considerable variation from country to country, they observe that the share of the for-profit sector is constantly growing. They show that in France, particularly in the field of nursing homes, the growth and internationalisation of national champions has been very rapid. For Germany, Europe's largest market, they explain that there are powerful groups that are mainly active at the national level and whose profits are exploding (ENCO, 2021). They underline the evils encountered in terms of practices in commercial care facilities: institutional abuse, lack of equipment and personal; exorbitant prices and prioritisation of profit.

Future research should address this for profit/not for profit distinction on the impact of well-being or mortality, in line with Comondore et al. (2009) who suggested that not-for-profit facilities delivered higher quality care than did for-profit facilities. Unfortunately, SHARE data does not allow to identify for each resident the nature of the nursing home. But other data from SHARE gives us additional information on the sample we studied. Table 7 summarizes information (unfortunately not available for our entire sample) and illustrates that the Spanish and Italian populations over 65 years with 1 ADL differ greatly from those of the other countries of the sample considered. In this table, we do not distinguish between people living in a nursing home or at home, but we can see that their social network is rather limited and very close to their home, that their social activities²⁴ are reduced (66% of the Italian and 63% of the Spanish respondents declare not to have any), that they have a frail mental health, fairly low cognitive performance and a reduced geographical mobility throughout their life. These populations therefore seem to be more fragile, and it is possible that their relatives, and in particular their children, play an important role (as shown from microdata on informal care presented in Table 6) by coming to help them even in nursing homes where they would compensate for the deficient services provided. Empirically, Klimaviciute et al. (2017) were able to show that the social norm was a motive for intergenerational relationships in southern European countries, implying that children helping their parents regardless of the situation.

Insert Table 7: Other SHARE interest variables

Although our analysis takes place before COVID-19, our results confirm the excess mortality observed during the pandemic. The pandemic has been particularly deadly for residents of nursing homes and other long-term care facilities in the US and in European countries (Shen, 2022). Recent studies have pointed out the heterogeneity of the quality of the nursing homes (Cronin and Evans, 2022), the organization structure of the staff (Dean et al., 2022) and the private ownership of the nursing home (Arnedo et al., 2022) as explaining factors of the excess mortality due to COVID-19. These findings support our analysis of the potential mechanisms behind the excess mortality in nursing homes.

One striking result of our study is the absence of excess mortality in Southern Europe nursing homes. We have proposed a hypothesis to answer this result, which does not contradict findings of Alacevitch et al. (2021) showing, for the Lombardy region in Italy, an excess mortality due to COVID-19 in municipality where there were nursing homes compared to municipalities without nursing homes. As Trabucchi and De Leo (2020) state, nursing homes of northern Italy were like isolated citadels during the first wave, with very little contact with the external environment. Similar results have been documented for Spain by Costa-Font et al. (2021) who emphasize the underfunding of nursing homes, reflected by understaffing, larger nursing homes, and higher occupancy rates. Policy choices may also be responsible for the explosion of deaths in nursing homes during the pandemic. Practices of discharge into nursing homes have been reported. Also, the lack of coordination between health and long-term care have been pointed out as major problem in several countries, including Italy and Spain (Costa-Font et al., 2017). During the pandemic, Costa-Font et al. (2021) report that, in Spain, older patients were refused emergency healthcare by hospitals, with clinical guidelines explicitly saying not to

²⁴ The social activities proposed by SHARE are the following: 1) Done voluntary or charity work; 2) Attended an educational or training course; 3) Gone to a sport, social or other kind of club; 4) Taken part in a political or community-related organization; 4) Read books, magazines or newspaper; 5) Did word or number games such as crossword puzzles or Sudoku; 6) Played cards or games such as chess.

admit older patients residing in nursing homes. These elements are likely to have worsened the situation for nursing homes residents compared to the treatments or the help from the relatives they have access to before the pandemic.

7. Conclusion

The current health crisis has highlighted the vulnerability of the society's most elderly people, whose numbers are expected to double by 2050 (OECD, 2019). One important finding from the COVID pandemic is the high number of COVID-related death in nursing homes and in particular the differences observed between countries. During the COVID pandemic, the structural shortcomings of the long-term care sector have become even more visible: care workers are under enormous pressure in often very difficult conditions and with limited support (OECD, 2020). This has had an impact on the management of the pandemic by exposing older people to additional risks (OECD, 2020; Chen *et al.*, 2021). On the basis of these observations of excess mortality in homes during the pandemic period, we wanted to verify the pre-existence of this impression/trend.

Using the propensity score matching method, we constructed a sample in which treated (being in a nursing home) and untreated individuals staying at home have similar characteristics in terms of age, gender, degree of dependence, state of health, availability of informal help and assets. The goal was to determine whether there is a causal link in being in a nursing home on the probability of death. Our results indicate that overall, residing in nursing homes increases the probability to die earlier than staying at home. This hides important differences across countries with Germany, Switzerland and Belgium showing “deadlier” nursing homes than the other countries. One should note that this finding can be interpreted alternatively as indicating that staying home in these countries is a safer option than moving to a nursing home.

These results can be related to country-specific features of the long-term care. It appears that countries in which the mortality in nursing homes is higher are also the countries in which the public spending and the resources devoted to long term care are low. This does not allow to draw a causal conclusion but gives some hints on the mechanisms that would explain the mortality differential. Future research on the specific role of nursing homes should investigate the cross-country differences. It might be the case that some countries give more weight to the possibility of long-term care (and dying) at home surrounded by family. Where this is the case, one may expect that there will be fewer nursing homes and less dying in nursing homes than where less value is attributed to long-term care and dying at home.

The Covid-19 crisis has shown the extent to which older people staying in institutions were exposed when the virus managed to infect their nursing homes. This would only confirm, on a much larger (and dramatic) scale, the *mortality ratio* already observed previously in several countries between dependent people with similar characteristics being cared for in institutions or staying at home.

Acknowledgements: *We are grateful to the editor and two anonymous referees. Their suggestions helped us to greatly improve the paper. This research participates in the SHARE-COVID19 Project funded by the European Commission. This work was also supported by the French National Research Agency Grant ANR-17-EURE-0020, and by the Excellence Initiative*

of Aix-Marseille University - A*MIDEX. The authors would like to thank the editor as well as two anonymous referees for their insightful comments. They thank Anne Laferrère, Matthias Kredler and Pedro Mira as well as the participants of the SCOR - TSE Workshop on Long Term Care and Aging which was held on January 28th 2022, the conference of the Public Economic Theory Association in Marseille on June 8 and the Max Planck/NBER workshop on aging and health (Spring 2022 in Munich), especially Kathleen McGarry and Alessandro Previtero for many useful comments.

References

- Alacevich, C., Cavalli, N., Giuntella, O., *et al.* (2021). The presence of care homes and excess deaths during the COVID-19 pandemic: Evidence from Italy, *Health Economics*, 30(7), 1703-1710.
- Angelini, V. and A. Laferrère (2012). Residential mobility of the European elderly. *CESifo Economic Studies*, 58(3), 544-569. DOI:10.1093/cesifo/ifr017.
- Angrist, J. D. and J. Pischke (2008). Mostly harmless econometrics: An empiricist's companion. *Princeton University Press*.
- Antwi Y. and J.R. Bowblis (2018). The Impact of Nurse Turnover on Quality of Care and Mortality in Nursing Homes: Evidence from the Great Recession, *American Journal of Health Economics*, MIT Press, 4(2), 131-163.
- Arnedo-Pena, A., Romeu-Garcia, M.A., Gascó-Laborda, J.C., *et al.* (2022). Incidence, Mortality, and Risk Factors of COVID-19 in Nursing Homes. *Epidemiologia*, 3, 179–190.
- Bakx, P., Wouterse, B., van Doorslaer, E. and A. Wong (2020). Better off at home? Effects of nursing home eligibility on costs, hospitalizations and survival. *Journal of Health Economics*, 73, 102354.
- Beck-Sague, C., Banerjee, S. and R. Jarvis (1993). Infectious Diseases and Mortality among US Nursing Home Residents. *American Journal of Public Health*, vol.93 (12), 1739-1742.
- Bolin, K., Lindgren, B. and P. Lundborg (2008). Informal and formal care among single-living elderly in Europe. *Health Economics*, 17 (3), 393-409.
- Bonsang, E. (2009). Does informal care from children to their elderly parents substitute for formal care in Europe? *Journal of Health Economics*, vol.28 (1), 143-154.
- Börsch-Supan, A., Brandt, M., Hunkler, C., *et al.* (2013). Data Resource Profile: the Survey of Health, Ageing and Retirement in Europe (SHARE). *International Journal of Epidemiology*, 42(4), 992-1001.
- Braggion, M., Pellizzari, M., Basso, C., *et al.* (2020). Overall mortality and causes of death in newly admitted nursing home residents. *Aging Clinical and Experimental Research*, 32, 275-280.
- Canadian Institute for Health Information (CIHI, 2020). Pandemic Experience in the Long-Term Care Sector: How Does Canada Compare With Other Countries?. Ottawa, ON: CIHI.

Canta, C., Pestieau, P. and J. Schoenmaeckers (2021). Blood and Gender Bias in Informal Care within the Family? LIDAM Discussion Paper CORE ; 2021/15, 30p.

Charles, K. and P. Sevak (2005). Can family caregiving substitute for nursing home care? *Journal of Health Economics*, 24, 1174-1190.

Chen, K., Chevalier, J. and E. Long (2021). Nursing home staff networks and Covid-19, *Proceedings of the National Academy of Sciences*, 118(1).

Comondore, V., Devereaux, P., Zhou, Q., *et al.* (2009). Quality of Care in For-Profit and Not-for-Profit Nursing Homes: Systematic Review and Meta-Analysis. *BMJ: British Medical Journal*, 339: 7717: 381–384.

Costa-Font, J., E. Norton and L. Siciliani (2017). The Challenges of Public Financing and Organisation of Long-Term Care, *Fiscal Studies*, 38(2): 365-368.

Costa-Font, J., Jiménez Martin S., Viola A. (2021), Fatal Underfunding? Explaining COVID-19 Mortality in Spanish Nursing Homes. *Journal of Aging and Health*. 33(7-8), 607-617.

Cronin, C. and W. Evans (2022). Nursing home quality, COVID-19 deaths, and excess mortality. *Journal of Health Economics*, 82, 102592.

Dean, A., McCallum, J., Kimmel, S. and A; Venkataramani (2022), Resident Mortality And Worker Infection Rates From COVID-19 Lower In Union Than Nonunion US Nursing Homes, 2020–21, *Health Affairs*, 41(5), 751-759.

Degavre, F. and M. Nyssens (2012). Care regimes on the move: Comparing home care for dependent older people in Belgium, England, Germany and Italy, 281 p.

ENCO (2021), Caring for profit. <https://corpwatchers.eu/en/investigations/caring-for-profit-en/mapping-the-privatisation-of-healthcare-in-europe?lang=en>

EUROSTAT (2016). European Statistical System, Census Hub. URL: <https://ec.europa.eu/CensusHub2>.

Falcone, M., Russo, A., Gentiloni Silverj, F., *et al.* (2018). Predictors of mortality in nursing-home residents with pneumonia: a multicentre study. *Clinical Microbiology and Infection*, 24(1), 72-77.

Frölich, M. (2004). Finite-Sample Properties of Propensity-Score Matching and Weighting Estimators. *The Review of Economics and Statistics*, 86(1), 77-90.

Giudici, C., Poletti, S., de Rose, A. and N. Brouard (2019). Which aspects of Elderly Living Conditions are Important to predict Mortality? The complex role of Family Ties at Home and in Institutions. *Social Indicators Research*, 142, 1255-1283.

Ichino, A., Mealli, F. and T. Nannicini (2008). From Temporary Help Jobs to Permanent Employment: What Can We Learn from Matching Estimators and their Sensitivity? *Journal of Applied Econometrics*, 23(3), 305-327.

Imbens, G. (2015). Matching Methods in Practice. Three examples. *Journal of Human Resources*, 50(2), 373-419.

- Ito, S., Lépine A. and C. Treibich (2018). The effect of sex work regulation and well-being of sex workers: Evidence from Senegal. *Health Economics*, 27, 1627-1652.
- King, G. and R. Nielsen (2017). Why Propensity Scores Should Not Be Used for Matching. *Political Analysis*, 27(4), 435-454.
- Klimaviciute, J., Perelman, S., Pestieau, P. and J. Schoenmaeckers (2017). Caring for elderly parents: is it altruism, exchange or family norm? *Journal of Population Economics*, 30(3), 835-873.
- LaFerrère, A. and L. Arnault (2021). Choix résidentiels des plus de 65 ans en Europe: logement privé ou collectivité ?, *Gérontologie et Société*, 43 (165), 125-149.
- LaFerrère, A., Van den Heede, A., Van den Bosch, K. and J. Geerts (2013). Entry into institutional care: predictors and alternatives. In Börsch-Supan, A.; Brandt, M.; Litwin, H.; Weber, G (Eds), *Active Ageing and Solidarity between generations in Europe* (pp. 187-198). Berlin, Germany: De Gruyter.
- Lo Sasso, A.T. and R.W. Johnson (2002). Does informal care from adult children reduce nursing Home Admissions for the elderly? *Inquiry*, 39 (3) (2002), 279-297.
- Levy, C., Kheirbek, R., Alemi, F., *et al.* (2015). Predictors of 6-month mortality among nursing home residents: Diagnoses maybe more predictive than functional disability. *Journal of Palliative Medicine*, 18(2), 100-106.
- Lin H. (2014). Revisiting the relationship between nurse staffing and quality of care in nursing homes: an instrumental variables approach. *Journal of Health Economics*, 37:13-24.
- Lindrooth, R., Hoerger, T. and E. Norton (1993). Expectations among the Elderly about Nursing Home Entry. *HSR: Health Services Research*, 35 (5 Pt 2), 1181–1202.
- Nannicini, T. (2007). Simulation-based sensitivity analysis for matching estimators. *Stata Journal*, 7(3), 334-350.
- OECD (2019), Health at a Glance 2019: OECD Indicators, OECD Publishing, Paris, <https://doi.org/10.1787/4dd50c09-en> .
- OECD (2020), Who Cares? Attracting and Retaining Care Workers for the Elderly, OECD Health Policy Studies, OECD Publishing, Paris, <https://doi.org/10.1787/92c0ef68-en> .
- Ripollone, J., Huybrechts, K., Rothman, K. *et al.* (2018). Implications of the Propensity Score Matching Paradox in Pharmacoepidemiology. *American Journal of Epidemiology*, 187(9), 1951-1961.
- Rosenbaum, P. and Rubin, D. (1983). The central role of the propensity score in observational studies for causal effects. *Biometrika*, Volume 70, Issue 1, Pages 41-55.
- Rosenbaum, P. and Rubin, D. (1985). Constructing a Control Group Using Multivariate Matched Sampling Methods that Incorporate the Propensity Score. *The American Statistician*, 39, 33-39.

Schanze, J-L. and S. Zins (2019). Undercoverage of the elderly institutionalized population: The risk of biased estimates and the potentials of weighting. *Survey Methods: Insights from the field*.

Schmitz, H. and M. Stroka-Wetsch (2020). Determinants of nursing home choice: does reported quality matter? *Health Economics*, 29, 766-777.

Schultze, A., Nightingale, E., Evans, D., *et al.* (2022), Mortality among Care Home Residents in England during the first and second waves of the COVID-19 pandemic: an observational study of 4.3 million adults over the age of 65, *The Lancet Regional Health - Europe*, 14.

Shen, K. (2022), Relationship between nursing home COVID-19 outbreaks and staff neighborhood characteristics, *PLoS ONE*, 17(4): e0267377.

Spasova, S., Baeten, R., Coster, S., *et al.* (2018). Challenges in long-term care in Europe. A study of national policies, *European Social Policy Network (ESPN)*, Brussels: European Commission.

Sung, K. (2014). Predictive Factors associated with Death of Elderly in Nursing Homes. *Asian Nursing Research*, 8, 143-149.

Trabucchi, M. and D. De Leo (2020). Nursing home or besieged castles: COVID-19 in northern Italy. *Lancet Psychiatry*, 7(5), 387-388.

Van Houtven, C.H. and E.C. Norton (2004). Informal care and health care use of older adults. *Journal of Health Economics*, 23, 1159-1180

Tables

Table 1. SHARE waves and selected observations

<i>Groups of countries/Countries</i>		W4	W5	W6	W7	Pooled Obs. of 65+ & 1 ADL at time t (#)	NH if 65+ & 1 ADL at time t (%)
North	Denmark	x	x	x	x	566	16.4
	Netherlands	x	x			200	8.5
	Sweden	x	x	x	x	760	10.1
Central	Austria	x	x	x	x	1018	6.7
	Belgium	x	x	x	x	1708	11.6
	France	x	x	x	x	1236	6.7
	Germany	x	x	x	x	823	6.1
	Luxembourg		x	x	x	164	16.5
	Switzerland	x	x	x	x	417	9.1
South	Italy	x	x	x	x	1134	1.8
	Spain	x	x	x	x	1761	4.4
East	Czech. Rep.	x	x	x	x	1331	5.4
	Estonia	x	x	x	x	2222	1.9
All						13340	6.5

Table 2. Mortality rates ratio in SHARE countries (65+ & 1+ ADL)

<i>Groups of countries/Countries</i>		Deceased at time $t+1$ (%)			
		Both NH & AH at time t	Nursing Home at time t	At Home at time t	Mortality ratio (NH/AH)
All		20.2	41.5	18.8	2.2
North	Denmark	26.3	46.2	22.4	2.1
	Netherlands	11.5	29.4	9.8	3.0
	Sweden	19.6	48.0	16.4	2.9
Central	Austria	19.4	33.8	18.4	1.8
	Belgium	19.1	42.4	16.0	2.6
	France	17.3	42.2	15.5	2.7
	Germany	19.7	44.0	18.1	2.4
	Luxembourg	20.1	40.7	16.1	2.5
	Switzerland	15.3	44.7	12.4	3.6
South	Italy	19.6	20.0	19.6	1.0
	Spain	25.4	39.7	24.7	1.6
East	Czech Rep.	21.9	41.7	20.8	2.0
	Estonia	18.9	38.1	18.6	2.0

Note: NH = Nursing Home resident; AH = At Home

Table 3. Summary statistics of covariates (65 + & 1+ ADL)

<i>Covariates</i>		Both NH & AH at time <i>t</i> (%)	Nursing Home at time <i>t</i> (%)	At Home at time <i>t</i> (%)
<i>Socio-economic characteristics</i>				
Sex	Men	39.0	30.9	39.6
	Women	61.0	69.1	60.4
Age	65-74	34.0	12.4	35.5
	75-84	42.3	32.9	42.9
	85+	23.7	54.7	21.6
Education	Low	60.6	65.7	60.3
	Medium	24.1	21.8	24.2
	High	15.3	12.5	15.5
Wealth	1st tercile	50.0	89.6	47.3
	2nd tercile	29.7	7.5	31.2
	3rd tercile	20.3	2.9	21.5
<i>Health characteristics</i>				
ADLs	1 or 2	67.4	41.1	69.2
	3 or 4	17.6	20.4	17.4
	5 or 6	15.0	38.5	13.4
IADLs	0	23.9	5.5	25.1
	1 to 3	40.4	18.5	42.0
	4 or 5	14.9	19.2	14.6
	6 or +	20.8	56.8	18.3
Chronic diseases	0 or 1	18.6	24.9	18.2
	2 or 3	42.2	43.7	42.0
	4 or +	39.2	31.4	39.8
<i>Family characteristics</i>				
In couple	Yes	47.8	7.6	55.2
	No	52.2	92.4	44.8
At least one child	Yes	89.0	78.0	89.7
	No	11.0	22.0	10.3
At least one daughter at less than 25 kms	Yes	15.8	9.5	16.2
	No	84.2	90.5	83.8
At least one daughter not working	Yes	21.0	23.1	20.8
	No	79.0	76.9	79.2
		100%	100%	100%
Observations		13340	863	12477

Note: NH = Nursing Home resident; AH = At Home

Table 4. ATT estimation with the Propensity-Score Kernel Matching (exact matching with replacement)

	# treated	# control	ATT	S.E.
All	803	6301	0.107***	0.023
North	172	746	0.113**	0.052
Denmark	87	309	0.056	0.077
Netherlands	14	37	0.133	0.186
Sweden	70	406	0.145*	0.082
Central	434	2822	0.105***	0.032
Austria	66	634	0.062	0.079
Belgium	190	1011	0.092*	0.049
France	76	495	0.039	0.084
Germany	47	326	0.277***	0.080
Luxembourg	26	35	0.205	0.126
Switzerland	37	226	0.270***	0.099
South	91	1113	0.001	0.062
Italy	20	343	-0.084	0.110
Spain	74	785	0.001	0.991
East	106	1558	0.146***	0.052
Czech Rep.	67	779	0.165**	0.070
Estonia	39	336	0.147*	0.088

Note: ***, ** and * stand for statistically significant at the 1%, 5% and 10% levels respectively.

Table 5. Sensitivity analysis: confounder-like and killer confounder

	Outcome Effect Γ	Selection Effect Λ	ATT
<i>PSM (Kernell)</i>	<i>1</i>	<i>1</i>	<i>0.107</i>
<i>Confounder-like</i>			
Being a woman	0.659	1.613	0.077
Having at least one child	0.963	0.410	0.068
Having one daughter not working	1.407	1.089	0.069
Having one daughter at less than 25 kms	1.057	0.537	0.070
<i>Killer confounder</i>			
U' ($d = 0.1$ & $s = 0.58$)	1.568	18.215	0.011
U'' ($d = 0.2$ & $s = 0.46$)	2.235	10.320	0.001
U''' ($d = 0.3$ & $s = 0.34$)	4.039	6.053	0.001

Table 6: Information about Formal and Informal Care by country

	Formal LTC							Informal LTC	
	Public spending in LTC				Number of LTC workers per 100 individuals 65+	LTC beds per 100,000 inhabitants	Share of private NH for profit (%)	Share of population providing informal care (%)	Share of informal carers providing more than 20h care per week (%)
	% of GDP	Institutional care	Home care	Cash benefits					
Denmark	3.5	62.0	38.0	0.0	8.1	750	6.5 ²	15.2	8.1
Netherlands	3.7	51.0	16.4	32.6	8.0	1371	20.0	36.7	3.3
Sweden	3.3	52.6	44.7	2.6	12.4	1388	15.0	22.0	5.4
Austria	1.8	49.1	9.9	41.0	4.1	865	21.0	8.1	19.0
Belgium	2.2	62.5	26.8	10.7	4.8	1276	33.0*	11.6	15.0
France	1.9	69.6	24.8	5.6	2.3	981	22.0	14.1	10.5
Germany	1.6	35.7	23.5	40.8	5.1	1152	40.0	6.8	15.0
Luxembourg	1.0	63.8	35.6	0.6	7.9	1168	9.6 ³	6.2	17.8
Switzerland	2.4	82.9	17.1	n.a.	8.3	1170	40.0	n.a.	
Italy	1.7	28.2	19.5	52.3	1.9	416	22.0	5.8	40.5
Spain	0.7	50.2	25.9	23.9	4.5	830	53.0	11.5	52.9
Czech Rep.	1.5	57.0	15.4	27.5	2.3	687	3.0	4.6	33.3
Estonia	0.4	52.7	42.7	4.6	5.3	871	80.0	13.4	17.3

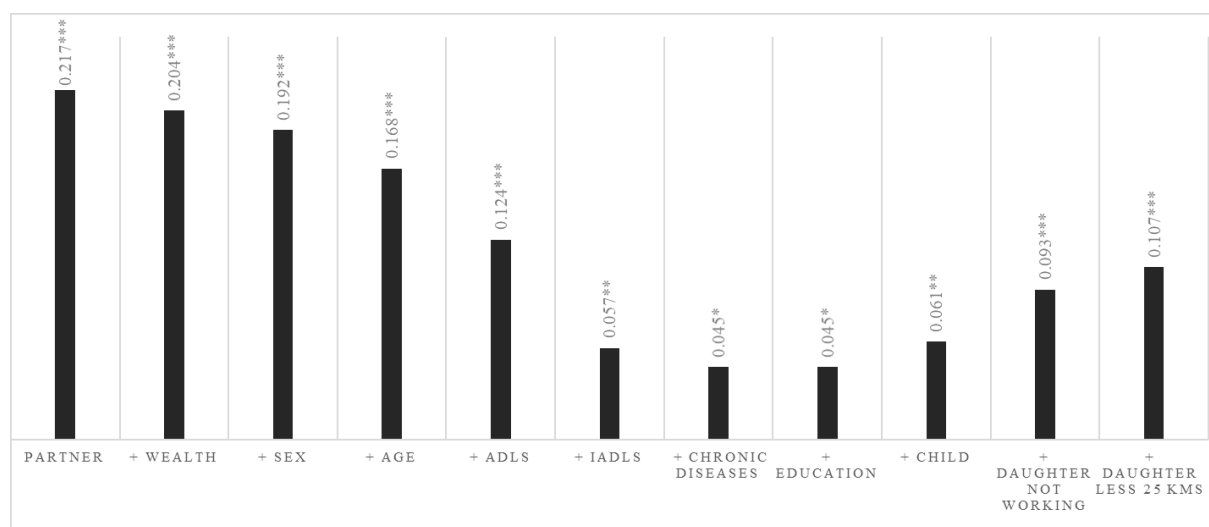
Note: Figures for the column "Private Nursing Home for profit" come from the European Network of Corporate Observatories (2021). When there is missing data, we use firstly STATISTA information (<https://www.statista.com/statistics/1239811/distribution-of-nursing-home-care-beds-by-public-or-private-ownership/>), indicated by the symbol "2", this is the case of Denmark. For Luxembourg, information is not available on STATISTA and data then comes from SPC and DG EMPL (2021), indicated by the symbol "3". For the other variables, data come from SPC and DG EMPL (2021) for countries from EU and from OECD (2021) or Office fédéral de la Statistique (<https://www.bfs.admin.ch/bfs>) for Switzerland. The data correspond to data collected between 2016 and 2019, prior to COVID.

Table 7: Other SHARE interest variables

	Social networks				No Social activities	Mental Health	Cognitive skills	Number of moves
	Size (0 to 7)	# contacts	Proximity less 1 km (#)	Proximity less 5 km (#)	0 or 1	0 or 1	0 to 20	0 to 19
All	2.34	1.99	1.14	1.56	0.27	0.58	6.25	3.57
Denmark	2.22	2.28	0.82	1.30	0.11	0.41	7.00	7.27
Netherlands	2.31	2.34	0.97	1.5	0.05	0.39	7.63	5.21
Sweden	2.55	2.22	0.78	1.30	0.10	0.41	7.31	5.79
Austria	2.70	1.96	1.24	1.69	0.18	0.56	6.90	2.61
Belgium	2.51	2.31	0.96	1.57	0.17	0.52	6.72	4.15
France	2.34	2.20	1.08	1.48	0.23	0.59	6.01	4.27
Germany	2.59	2.09	1.23	1.64	0.15	0.44	7.21	3.32
Luxembourg	2.03	1.93	0.91	1.24	0.15	0.58	5.91	3.20
Switzerland	2.84	2.25	1.10	1.60	0.10	0.44	7.29	4.94
Italy	1.86	1.37	1.39	1.68	0.66	0.73	4.69	2.46
Spain	2.18	1.49	1.59	1.96	0.63	0.73	3.79	2.35
Czech Rep.	2.15	1.93	1.17	1.51	0.20	0.56	6.73	2.12
Estonia	2.21	1.95	1.03	1.40	0.21	0.68	6.42	3.35

Note: the questions about social networks the questions were only asked in waves 4 and 6 of SHARE, hence the impossibility to include them in the matching analysis (loss of one third of the sample). The questions on social activities, mental health or cognitive performance were not asked to all the respondents (and particularly to the persons who are in a nursing home), hence again the impossibility to include them in the matching estimates. For mental health binary variable, 1 means you have at least 4 of the following mental issues: depression, pessimism, suicidality, guilt, sleep, interest, irritability, appetite, fatigue, concentration, enjoyment or tearfulness. For the cognitive skills continuous variable, this is the number of words (out of 20) that the respondent is able to remember. Finally, the number of lifetime moves comes from the retrospective SHARELIFE surveys (waves 3 and 7). The available information concerns 9,124 respondents of our final sample of 13,340 persons aged + 65 years and having 1+ ADL.

Figure 1: Cumulative addition of covariates and estimated ATTs



Appendixes

A.1. Gross data and representativeness of nursing homes respondents' samples

		Pooled Obs. of 65+ (#) at time t	NH if 65+ (%) at time t	ADL if NH (%) at time t
All		80670	1.6	66.0
North	Denmark	4743	2.5	81.2
	Netherlands	2785	1.5	42.9
	Sweden	6849	1.5	76.7
Central	Austria	6721	1.8	55.3
	Belgium	7550	3.8	70.2
	France	6663	1.7	73.7
	Germany	5413	1.3	73.5
	Luxembourg	1098	4.4	56.2
	Switzerland	4820	1.7	47.5
South	Italy	6824	0.4	74.1
	Spain	8993	1.2	81.7
East	Czech Rep.	8009	2.0	44.2
	Estonia	10202	0.6	73.8

A.2. Probit estimations of nursing home choice

<i>Covariates</i>	All	North	Central	South	East
Socio-economic characteristics					
Female	-0.198***	-0.409***	-0.164**	-0.184	-0.067
Age	0.239***	0.105	0.360***	0.079	0.112
Education	0.099***	-0.080	0.169***	0.066	0.046
Wealth	-0.658***	-0.772***	-0.537***	-0.806***	-1.288***
Health characteristics					
ADLs	0.084***	0.052	0.092***	0.060	0.130***
IADLs	0.164***	0.280***	0.171***	0.109***	0.112***
Chronic diseases	-0.064***	-0.081**	-0.077***	-0.090***	-0.025
Family characteristics					
In couple	-1.100***	-1.175***	-1.338***	-1.046***	-0.644***
At least one child	-0.247***	0.059	-0.101	-0.707***	-0.307**
At least one daughter at less than 25 kms	-0.049	0.154	0.008	0.151	-0.612***
At least one daughter not working	-0.207***	0.072	-0.218***	-0.356**	-0.359**
Pseudo-R ²	0.370	0.443	0.363	0.339	0.340

Notes: Note: ***, ** and * stand for statistically significant at the 1%, 5% and 10% levels respectively. If it is not significant, it means that there is no correlation established between the variable and being in a nursing home. The “family characteristics” and the “female” variables are dummy. The others are continuous.

A.3. Robustness tests of Average Treatment Effects of the Treated with other algorithms

<i>Methods / Groups of countries</i>	ATT estimation with Propensity-Score Kernel Matching (Baseline)				ATT estimation with Propensity-Score Kernel Matching + Cross validation with respect to X				ATT estimation with Propensity-Score Kernel Matching + Cross validation with respect to Y (Frölich 2004, 2005)			
	# treated	# control	ATT	S.E.	# treated	# control	ATT	S.E.	# treated	# control	ATT	S.E.
All	803	6301	0.107***	0.023	700	5573	0.047*	0.026	802	6268	0.105***	0.023
North	172	746	0.113**	0.052	110	449	-0.046	0.074	164	729	0.077	0.056
Central	434	2822	0.105***	0.032	380	2358	0.061*	0.035	455	2923	0.173***	0.028
South	91	1113	0.001	0.062	88	1076	-0.018	0.063	91	1113	0.005	0.062
East	106	1558	0.146***	0.052	95	1242	0.120**	0.058	112	1577	0.171***	0.051
<i>Methods / Groups of countries</i>	ATT estimation with Mahalanobis-Distance Kernel Matching (King and Nielsen, 2019)				ATT estimation with Mahalanobis-Distance Kernel Matching (King and Nielsen, 2019) : Cross validation with respect to X				ATT estimation with Mahalanobis-Distance Kernel Matching (King and Nielsen, 2019) : Cross validation with respect to Y (Frölich 2004, 2005)			
	# treated	# control	ATT	S.E.	# treated	# control	ATT	S.E.	# treated	# control	ATT	S.E.
All	842	6141	0.192***	0.019	652	2227	0.093***	0.026	841	5980	0.188***	0.019
North	178	770	0.206***	0.045	135	335	0.057	0.060	173	633	0.178***	0.047
Central	459	2760	0.217***	0.026	352	1040	0.121***	0.035	459	2748	0.216***	0.025
South	94	922	0.032	0.055	84	390	-0.023	0.066	94	1169	0.072	0.052
East	112	1484	0.186***	0.050	97	599	0.129**	0.059	111	1335	0.168***	0.050

A.4. Addition of covariates and ATT estimated by Propensity-Score Kernel Matching (exact matching with replacement)

Addition of Covariates	<i>Partner</i>	+ <i>Wealth</i>	+ <i>Sex</i>	+ <i>Age</i>	+ <i>ADLs</i>	+ <i>IADLs</i>	+ <i>Chronic diseases</i>	+ <i>Education</i>	+ <i>Child</i>	+ <i>Daughter not working</i>	+ <i>Daughter less 25 kms</i>
# treated	863	863	863	863	855	837	817	808	803	806	803
# control	9681	6064	9229	8860	9188	8859	8687	8612	7705	6882	6301
ATT All	0.217***	0.204***	0.192***	0.168***	0.124***	0.057**	0.045*	0.045*	0.061**	0.093***	0.107***
ATT North	0.229***	0.213***	0.181***	0.180***	0.086*	-0.028	-0.040	-0.031	0.006	0.042	0.113**
ATT Center	0.241***	0.233***	0.181***	0.220***	0.144***	0.091***	0.093***	0.083**	0.082**	0.092***	0.105***
ATT South	0.098*	0.080	0.064	0.060	-0.017	-0.062	-0.044	-0.054	-0.006	-0.006	0.001
ATT East	0.204***	0.176***	0.146***	0.134***	0.137***	0.090	0.058	0.078	0.119**	0.156***	0.146***

A.5. Sensitivity analysis by country: confounder-like

	U = distribution woman			U = distribution at least one child			U = distribution daughter not working			U = distribution daughter at less than 25 kms		
	ATT	Out. Eff.	Sel. Eff.	ATT	Out. Eff.	Sel. Eff.	ATT	Out. Eff.	Sel. Eff.	ATT	Out. Eff.	Sel. Eff.
North	-0.008	0.773	1.428	-0.012	0.940	0.616	-0.012	1.137	1.310	-0.010	1.405	0.716
Central	0.102	0.659	1.974	0.091	0.952	0.493	0.091	1.470	1.274	0.092	0.902	0.652
South	0.016	0.647	1.451	0.019	1.026	0.163	0.014	1.286	0.518	0.010	0.970	0.535
East	0.124	0.591	1.856	0.115	1.031	0.401	0.116	1.527	0.906	0.117	1.249	0.200