**Data Appendix**

Article: Lucas Jeanneau, Quitterie Roquebert & Marianne Tenand (2021).

`**No more visits. Informal care in nursing homes prior to the outbreak of Covid-19’,** *Gérontologie et Société*, n°166, December.

This Appendix contains information on the data used in the empirical analyses and data treatment decisions that were made. In addition, it provides a description of the scripts containing the code for outputting the results presented in the article.

Additional details on the survey CARE-Institutions

We first provide additional information regarding the CARE-Institutions survey, in connection to our research question and analyses.

Sampling procedure and questionnaires

The sampling was made in two steps. First, a sample of long-term care institutions was drawn and surveyed in order to and retrieve the list of residents. In a second step, a sample of permanent residents was drawn within each institution.

General information about the institution and some individual information about the residents who were surveyed (e.g. public long-term care transfers received) was collected through a questionnaire at the level of the institution (*Questionnaire Etablissements*). A second questionnaire, containing most of the variables we exploit in our analysis, was administered to the selected residents or to proxy respondents (*Questionnaire Seniors*).

3,262 respondents from 616 institutions participated into the survey. Due to the compulsory nature of the survey the response rate was high (88% at the institution level and 86% at the respondent level).

Scope of the survey

Long-term care institutions include non-medicalized and medicalized nursing homes (respectively EHPA and EHPAD) as well as the long-term care units of hospitals (USLD).

EHPA stands for *Établissement d`hébergement pour personnes âgées*; it hosts individuals with no or limited activity restrictions (GIR 5 and 6, on the disability scale used by the French administration). Professional workers perform tasks relating to room and board, personal hygiene, medication management and social animations. Residents may benefit from paramedical care and a medical surveillance by external professionals. EHPAD stands for *Établissement d`hébergement pour personnes âgées dépendantes*; it hosts individuals with activity restrictions (GIR 1 to 4). Residents are provided a medical surveillance, nursing care and personal care. Finally, USLD stands for *Unités de soins de longue durée*, which are hospital units dedicated to hosting dependent individuals with very high medical care needs. EHPA, EHPAD and USLD welcome only individuals aged 60+.

Residents of intermediate housing facilities (*résidences services* et *residences autonomies,* formerly known as *foyers logements*) were surveyed in CARE-Ménages and not in CARE-Institutions.

Filters on Questions relating to IADL

In the institution-level questionnaire, it was asked whether residents could (i) do the grocery, (ii) do the domestic chores, (iii) prepare meals, (iv) manage medication and (v) move around alone. Within institutions in which one or several activities were typically not allowed for residents, the questions about their ability to perform these tasks were not asked to respondents.

Information regarding data treatment

Version of the survey

The curator of the survey (Drees) may release updates to the survey. Our analysis relies on version v2\_190822.

Statistical software

The analysis was executed using software R, version 4.0.2.

Missing values

We only drop 39 observations for which information on activity restrictions is missing. Our baseline sample counts 3,223 respondents.

When the respondent did not answer to the question ‘Do you have a partner?’, we coded her as not having a partner alive. 4.8% of those who responded they had a partner did not report where she lived. We assumed then she lived in the community. Along the same lines, we considered that those who did not answer to the question ‘Do you have any children (including adopted children)?’ had no children alive.

Declaration of caregivers for ADL and IADL

Discrepancy may occur between the initial declaration of some respondents and the information they provided about the care provided by each of their caregivers (e.g. the respondent declares not being helped for grooming, but later reports that caregiver X helps grooming). The curator of the survey (Drees) used information provided for each caregiver to construct and release corrected versions of the respondent-level dummy variables for informal care receipt for each ADL and IADL (variables RAAIDENT\_R1 to RAAIDENT\_R13). We create a dummy variable indicating help with ADL/IADL if the respondent is helped for at least one ADL or IADL (i.e. at least one of dummies RAAIDENT\_R1 to RAAIDENT\_R13 is equal to 1).

In the same vein, two corrected variables, AIDENTFI\_C and AIDENTSOU\_C were released; they indicate whether the respondent receives any financial support or any moral support respectively, based not only on her initial declaration but also on the detailed information she provided for each of their caregivers. We use these corrected variables to document how common financial support and moral support are.

Relationship between the respondent and their caregivers

Detailed information on the caregivers reported by the (proxy) respondents is contained in a dataset at the caregiver-level. It contains a variable, LIENSENAID, that provides the relationship between the respondent and their informal caregiver (e.g. their partner, their brother etc.). This variable is used to retrieve the proportion of individuals who receive help with ADL/IADL from different categories of relatives, defined as follows:

* Helped by child or partner
* Helped by siblings or parents
* Helped by other family member (grandchild, daughter- or son-in-law, mother- or father-in-law, niece, nephew, aunt, uncle, cousin, other family member)
* Helped by friend or another member of the institution not being a professional caregiver (assumed to be another resident)
* Helped by someone else (another relative not being a family member, or someone whose relationship with the respondent is not known).

The data reveal an inconsistency, which is that a small number of individuals with no children nor partner, as recorded in the survey, receive help from children or a partner (less than 2%). For the categorization of caregivers, we have then assumed that respondents were helped by another family member.

Furthermore, there is a small discrepancy between the number of individuals receiving informal care for ADL or IADL based on the LIENSENAID variable, and that computed based the respondent-level variables provided in the survey (dummies RAAIDENT\_R1 to RAAIDENT\_R13, cf. *supra*). This discrepancy is small, and mainly visible for the subgroup of individuals who have no partner nor children alive (56% of them do not receive informal care with ADL or IADL with our baseline definition, but it is the case of 54% when we compute this proportion using LIENSENAID).

Scripts

The code is split into several scripts.

* ***0.Data.R***

It loads two datasets from the CARE survey that will be used in the analyses:

* carei\_sen\_seniors\_v2\_190822.dta (dataset at the respondent level)
* carei\_sen\_aidants\_v2\_190822.dta (dataset at the informal caregiver level)
* carei\_sen\_enfants\_v2\_190822.dta (dataset at the children level)

From the `senior’ dataset, we proceed to the dropping of the 39 observations with item non-response on activity restrictions.

* ***1.Variables.R***

This code creates the variables used to produce the main descriptive statistics and econometric analysis.

It runs after 0.Data.R.

* ***2.DS.R***

This code outputs the general descriptive statistics displayed in Table 1 of the article, as well as the weighted proportions of individuals who have ADL restrictions (displayed in Figure 1), the weighted proportions of individuals who have IADL restrictions (displayed in Figure 2) and the weighted proportions of individuals who are helped for a range of ADL and IADL (displayed in Figure 4).

It runs after 0.Data.R and 1.Variables.R.

* ***3.Spider\_charts.R***

This code allows to output Figure 1, Figure 2 and Figure 4, based on statistics computed in 2.DS.R.

It runs after 0.Data.R and 1.Variables.R and 2.DS.R.

* ***4.Venn.R***

This code computes the proportion of individuals receiving combinations of different informal care types and to output Figure 3.

It runs after 0.Data.R.

* ***5.Caregivers.R***

This code links respondent-level information with caregiver-level information to retrieve their relationship with the respondent. It allows to compute the probability to be helped by certain types of relatives, depending on the presence of a partner or children.

It runs after 0.Data.R.

* ***6.Extensive\_margin.R***

This code produces the estimates from the econometric analyses (determinants of the probability to receive information care), for the entire sample and by the sub-samples defined by the presence of a partner and/or children. It outputs Table 3.

It runs after 0.Data.R and 1.Variables.R.