Preferences and matching in elderly care

Joseph-Simon Görlach*

Work in progress: March 2025

Abstract

This paper studies matching between elderly care recipients and private care workers using unique company data which provide information on clients' and workers' ex-ante preferences, matches, wages and durations. Estimating a matching model identifies valuations for different match attributes from both the client and worker perspectives, and informs about the effects of mismatches on payoffs and job stability. Merging with full longitudinal register data suggests elderly population shares increase most strongly among municipalities and demographics more prone to prefer younger and European care workers. Workers' sex, instead, may become less relevant.

Keywords: Aging, discrimination, immigration, long-term care, matching.

*Bocconi University, Department of Economics, CEPR, CReAM, LEAP and IZA; e-mail: josephsimon.goerlach@unibocconi.it. This paper was supported by NextGenerationEU.

1 Introduction

Paid elderly care at home is a rising trend in aging societies, yet it remains largely undocumented. In Italy, many elderly individuals rely on privately employed care workers rather than institutional care, increasingly substituting care by family members.[1]

Matching in this market often is by word of mouth or arranged by agencies specialized on this sector. In both cases, care workers typically are employed directly by care recipients or their families, who do not post explicit vacancies. Vacancy data by employment agencies thus cannot inform about the worker characteristics employers are looking for. Little is known about the attributes that clients seek in care workers, which preferences workers hold over types of clients, and how these preferences shape employment outcomes. However, given the intimacy of care and the fact that care workers often live in care recipients' private homes, preferences and trust are particularly important. Moreover, good matches can raise stability in this sector, which is characterized by low employment protection and high turnover.

This is the first paper to shed light on the preferences of employing families in this important sector. I draw on newly compiled company data that contains unique information about care recipients preferences over care workers' age, gender, origin and other characteristics, as well as workers' ex-ante preferences over care recipients' gender, cognitive and physical health condition. I observe matches including both sides' actual characteristics and how they compare to the ex-ante reported preferences of the respectively other party. In addition, the data contain information on contracts, including pay and work arrangements as well as the duration of formed employment relations.

Using these data, I estimate a matching model that quantifies the value of certain match characteristics and the importance of different preference dimensions from both the employer's and the worker's perspectives. The availability of wage data facilitates the identification of surplus allocation between the two sides. The model moreover leverages information on employment duration and dissolution, allowing also for an assessment of how match quality influences job stability. Information on match durations also helps identify surplus levels relative to each side's outside option.

Finally, I augmenting the company data with full longitudinal population reg-

[1] Appendix Figure 9 uses data from the Survey of Health, Ageing and Retirement in Europe (Börsch-Supan et al., 2013) and shows an increasing trend in paid care at home across many European countries.

ister data. I use these register data to predict growth in elderly population shares across municipalities and demographics until 2040. Growth in the share of individuals aged 80 or older is particularly pronounced in smaller and southern Italian municipalities. Within municipalities, there further is a relatively stronger increase among individuals in their 80s compared to those aged 90 and above. Similarly, the number of men aged 80+ is predicted to grow relatively more than the number of women in that age group. In the company data, these demographics and municipalities are associated with a stronger preference for younger and European care workers. Male care recipients or their families ("clients") moreover are less likely to state a preference over workers' sex, suggesting a potentially weaker future bias towards female care workers.

Overall, about one third of company clients states a preference for European over non-European care workers. Almost 90 percent prefer female carers, while the share of female care recipients is only 70 percent. This highlights the importance of distinguishing matching on preferences from assortative matching on sex. In addition, there is a strong bias towards care workers aged 40 or older. On the worker side, most care worker are ex-ante available for working with both men and women, although 12 percent state a preference for female care recipients. More than 90 percent of care workers are female themselves. My data thus allow for the important differentiation between (typically assortative) matching on observed characteristics and matching on potentially deviating preferences.

This is an important distinction from commonly estimated one-to-one matching models, for instance in the literature on marriage markets. The model I estimate draws on that literature, but with three fundamental differences. First, whereas models of marital matching often assume transferable utility, transfers typically are unobserved. As a result, only total match values are identified, not their allocation across the parties of a match. In the employment relations I analyze, instead, transfers in form of wages are observed and allow a determination of how variation in match characteristics affects each side's payoff separately. This is important, because for instance a given combination of care worker's sex and care recipient's sex may have a different effect on the value of the match for the worker and the value for the client. Second, the availability of ex-ante preference information is an important refinement over matching on observed characteristics only. For instance, while matching between patients and care workers is assortative on sex, some care recipients do prefer care workers of opposite sex. Direct preference information also is important, since for some characteristics there is meaningful variation only on one side of the employment relation. For instance,

close to one half of matched care workers are non-European and more than 80 percent are non-Italian, whereas close to all care recipients are Italian. Conversely, strong cognitive or physical impairment is found only among care recipients, not workers. Third, identification of the level of a match surplus in the marriage market literature relies on the observation of singles. Company data of the kind used in this paper naturally do not cover well the populations of unmatched agents. I thus make employment duration and match dissolution, which are observed in the data, an explicit part of the model that helps identify not only the relative values of different matches, but also the level of the match surplus.

The estimated model reveals the relevance of stated preferences. Clients who are matched with a male care worker despite stating a preference for female care workers suffer a reduction in surplus equivalent to 15 percent of the average wage. Vice versa, workers who are matched with a male despite stating a preference for female care recipients lose the equivalent of 10 percent of the average wage. When clients and workers are indifferent about the other party's sex, gains from same-sex matches instead are higher for workers, illustrating the importance of differentiating between matching on characteristics and preferences. Clients with a preference for European workers suffer a loss of 4 percent of the average wage when matched with a non-European worker.

The effect on match stability in the latter case is modest, since the baseline effect of non-European care workers on the match surplus is positive. Mismatches on either a client's or a worker's sex preference reduce employment duration by between 2 percent (when a worker prefers a male but is matched with a female care recipient) and 6 percent (when a client prefers a female but is matched with a male worker).

This paper contributes to several strands of research in health and labor economics, including the role of immigration in care provision and employer preferences and discrimination. The latter are investigated by Flabbi (2010), who develops a search model that incorporates employer taste discrimination in the labor market and shows that even unprejudiced employers engage in wage discrimination based on gender. Kuhn and Shen (2013) analyze explicit gender preferences in job ads on a Chinese online job board, finding that employers have idiosyncratic preferences for gender-job matches. Such preferences are prevalent for less skilled occupations. Expanding on this, Kuhn and Shen (2023) study the effects of removing gender preferences from job ads and find that this led to more gender-diverse applicant pools, as gender-mismatched applicants applied more frequently and were well received by employers. Card et al. (2023) instead examine

an Austrian ban of explicit gender preferences in job ads and find that this policy increased the hiring of women in male-dominated jobs and vice versa.

In relation to long-term care, the literature examines the choice between different care arrangements. Stöhr (2015) models how social interactions shape siblings' migration decisions in high-emigration countries, showing that the need to care for elderly parents can discourage migration. Mommaerts (2018) studies the impact of nursing home costs and finds that policies reducing out-of-pocket expenses for institutional care lower the prevalence of co-residence with adult children. Achou et al. (2023) examine differences in the marginal utility of financial resources depending on whether care is received at home or in a nursing home, finding that individuals value financial resources more when receiving home-based care. The welfare of care givers is studied by Costa-Font and Vilaplana-Prieto (2024), who show that long-term care provision significantly reduces the subjective well-being of informal caregivers. While their focus is on caring family members, I am not aware of estimates for professional care givers.

Immigration has been linked to elderly care both through its effect on the availability of care services and the downstream effects on labor market outcomes for family members, most notably daughters. Almuhaisen et al. (2023) show that stricter immigration enforcement reduces the supply of household service workers, leading to an increase in institutionalized care for elderly individuals. Similarly, Grabowski et al. (2023) document a direct link between immigration and the supply of labor to nursing homes, showing that higher immigration levels improve staffing and lead to better patient outcomes.

Similar to the provision of other home services studied by Cortes and Tessada (2011) and Cortes and Pan (2013), immigration also influences constraints and decisions by family members through its effects on the availability of professional care. Frimmel et al. (2023) show that declines in parental health negatively affect the labor market outcomes of adult children, but these effects are significantly mitigated when migrant care workers expand the availability of formal care. Also related, Haan and Wnuk (2024) find that increased foreign staffing alleviates labor shortages in the long-term care sector and improves working conditions for native workers.

A recent comprehensive survey of long-term care provision around the world is provided Gruber et al. (2023) and a series of national chapters. The situation in Italy is summarized by Brugiavini et al. (2023).

The remainder of the paper is structured as follows. Section 2 describes the data, outlining client and worker characteristics, stated preferences, and observed

matches, along with summary statistics and an analysis of match quality and contract stability. Section 3 presents the equilibrium matching model, detailing the identification strategy and estimation procedure used to assess the role of preferences and surplus allocation in match formation. The empirical findings follow in Section 4, examining how client and worker preferences influence observed match values, surplus distribution and job durations. Building on this, Section 5 incorporates population register data to project future demographic trends and their potential impact on the care market. Section 6 concludes with a discussion of the policy implications of these findings and potential directions for future research on labor market dynamics in private elderly care.

2 Data and descriptives

The main dataset used in this study was compiled from record from a small Italian company that specializes in matching elderly clients (or their families) with care workers. The data cover all client contacts, client information, service quotations, matches, and contract terminations from January 2020 to June 2023. The company maintains a database of potential care workers, and caseworkers are tasked with forming matches based on client requests. The company charges a fee at contract initiation and smaller monthly fees thereafter.

Client data are extracted from detailed questionnaires (displayed in Appendix B) that collect information on the elderly person's health status, care needs, and specific preferences regarding the care worker's age, sex, and nationality. Similarly, the worker database contains detailed demographic and professional characteristics of potential caregivers, along with their availability for different working conditions, such as live-in arrangements, caring for bedridden or Alzheimer's patients, and willingness to work with male or female clients. This information is collected through an online mask on the company's website.

2.1 Client and worker characteristics

Since company data by construction are non-representative and not a good representation of the full population of potential care recipients, the model in Section 3 will be estimated using matched clients and workers only. Table 1 lists the main characteristics and preferences of matched clients. On average, care recipients are in their mid 80s, 70 percent are female. Regarding health conditions, 11 percent of clients have been diagnosed with Alzheimer's, while 8 percent are bedridden.

A large majority of close to 80 percent, require co-residence with their caregiver. Stated preferences regarding worker characteristics suggest that most clients prefer caregivers aged 40 to 60, while only a small fraction explicitly request younger workers aged 20 to 40. A share of 37 percent of clients are indifferent regarding a worker's age. Gender preferences are strongly biased towards female caregivers. Note that the share of 89 percent stating a preference for female carers exceeds the share of female care recipients by almost 20 percentage points. Only 3 percent explicitly prefer a male worker. More than one third of clients prefer European workers, while 7 percent specifically request a non-European caregiver.

Table 1: Client characteristics and preferences, summary statistics

Variable	**Mean**	**Std. Dev.**	**Obs.**
Age	83.526	8.882	2763
Female	0.699	0.459	2787
Alzheimer	0.108	0.310	2761
Bedridden	0.083	0.276	2743
Demands co-residence	0.791	0.407	2783
Prefers care worker who is			
Aged 20-40	0.052	0.222	2697
Aged 40-60	0.577	0.494	2697
Female	0.886	0.318	2761
Male	0.032	0.175	2761
European	0.340	0.474	2726
Non-European	0.070	0.255	2726

Note: This table shows summary statistics for care recipients' characteristics and preferences regarding care workers. The sample includes clients matched by the firm between January 2020 and June 2023. The bottom panel lists shares of explicitly stated preferences over caregiver attributes, with the remaining shares pertaining to indifference between the categories. Source: Firm administrative data.

Table 2 summarizes the characteristics and preferences of matched workers. Care workers are predominantly female, with an average age of 49. The dataset reflects the diverse backgrounds of workers: 55 percent of workers are European, with Italians, Romanians, and Ukrainians forming the largest groups. Among non-European workers, the most common nationalities are Moroccan, Senegalese,

and Peruvian.

Most workers are available for working with both female and male care recipients, but around 12 percent prefer working with women. Less than 2 percent state a preference for male care recipients. Not all workers declare their availability for working with heavily impaired clients: 42 percent of workers indicate availability to care for Alzheimer's patients, and 43 percent for bedridden patients. About two thirds of care workers are open to co-reside with the care recipient.

Table 2: Worker characteristics and preferences, summary statistics

Variable	**Mean**	**Std. Dev.**	**Obs.**
Age	49.112	10.842	2649
Female	0.924	0.265	2667
European	0.553	0.497	2659
Italian	31.5%		
Romanian	31.3%		
Ukrainian	13.4%		
Non-European	0.446	0.497	2659
Moroccan	15.5%		
Senegalese	13.7%		
Peruvian	11.8%		
Prefers to work with			
Women	0.118	0.322	2655
Men	0.017	0.131	2655
Is available for			
Alzheimer patients	0.417	0.493	2655
Bedridden patients	0.430	0.495	2655
Co-residence	0.680	0.466	2655

Note: This table presents summary statistics for worker characteristics and preferences regarding caregiving roles. The sample includes workers registered in the firm's administrative database. Nationality percentages within European and non-European categories are based on the most frequently reported origins. Preferences refer to explicitly stated choices for client attributes and work conditions. Source: Firm administrative data.

A concern may be that workers with less frequently demanded characteristics

(such as workers of non-European origin) falsely overstate their availability to work with difficult care recipients (such as Alzheimer's or bedridden patients). To investigate this, I first restrict the sample to matches with workers who reported an availability to work with care recipients diagnosed with Alzheimer's disease and examine whether non-Europeans exhibit a higher probability of quitting the job if indeed they are matched with Alzheimer's patients. Columns (1) and (2) of Appendix Table 8 show that this is not the case (if anything the opposite is true). Columns (3)-(6) repeat this for workers declaring to be available for working with bedridden patients and with clients demanding co-residence. Similarly, Appendix Tables 9 and 10 consider declared availabilities and quits by young or male workers. None of the overall 18 coefficients pertaining to workers with a less demanded characteristics points towards a false over-reporting of availabilities.

Additional summary figures, including full age distributions of care recipients and workers, client preferences over all worker origins listed in the questionnaire, and wage distributions, are presented in Appendix C.

2.2 Match quality

A key advantage of the dataset is the ability to assess match quality by comparing each side's preferences with actual matches. The data show that the company is able to accommodate preferences to a high degree, though mismatches occur due to high dimensionality of preferences and the limited size of client and worker pools. Appendix Figure 17 provides a comprehensive overview of match quality across different dimensions, illustrating how well stated preferences align with observed matches.

Panel A shows that in 86 percent of matches clients' preference over a worker's age is met. When excluding age-indifferent clients, this share falls to 77 percent. Client preferences over worker origin are more different to meet given the supply of non-European workers. Excluding clients who are indifferent regarding the worker's origin leaves only 65 percent of clients with their preferred origin group (Panel D).

Instead, most matches satisfy clients' sex preferences (Panel B). This is less often the case for workers' preference over the care recipient's sex. Excluding workers who are available to work with both men and women, only 82 percent of care workers are matched with a care recipient corresponding to the worker's preferred sex.

The dataset also allows an assessment of matching outcomes for clients with

specific care needs. Some workers indicate a willingness to care for such clients, while others prefer to avoid these assignments. The data show that matches generally respect these preferences, but cases of mismatches still occur (Panels E and F). A similar pattern emerges for co-residence requirements, where clients requesting live-in arrangements are usually matched with workers open to this condition, though deviations from stated preferences are present (Panel G).

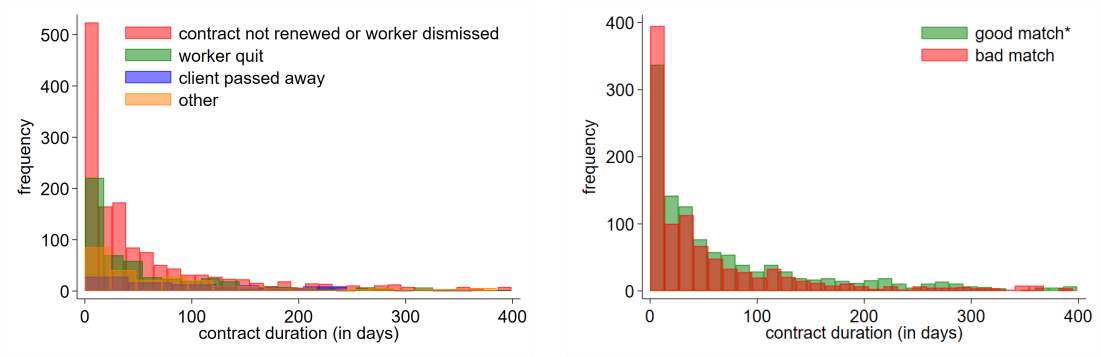
2.3 Contract duration and match quality

Match quality is potentially important for employment stability. The dataset provides detailed information on contract duration, which serves as a useful measure of match stability. The average employment lasts 126 days. Although the vast majority of contracts end within the observation period, this average includes the 11 percent of spells that are censored by the end of the observation window. The distribution is highly skewed, with a median duration of only 41 days. Termination reasons vary: 63 percent of employment relations are discontinued at the client's initiative, either because the worker is dismissed or the contract is not renewed. Worker resignations account for 22 percent of terminations, while client deaths lead to 4 percent of employments ending. The remaining 11 percent of terminations are attributed to other factors, including transfers to institutional care.

Figure 1 illustrates these patterns. The left panel presents the distribution of termination reasons, showing that most contracts end due to client-driven factors. The right panel examines contract duration by match quality, distinguishing matches in which clients' preferences over workers' age, sex and origin and workers' preferences over clients' sex, Alzheimer's and bedridden status as well as over co-residence are matched. These amount to about half of all matches. The right panel of Figure 1 shows that contracts where preferences are unmet tend to end sooner.

Table 3 assesses the relation between match quality and employment duration also conditional on match characteristics. The dependent variable in these regressions is the logarithm of the number of contract days. Panel A regresses log employment duration on the same indicator for matches respecting stated preferences across the different dimensions considered in Figure Figure 1b. Panel B of Table 3 instead counts the number of dimensions matched well. In both cases, the correlation with employment duration is strong and very robust also to conditioning on year and province fixed effects, as well as on client and worker characteristics.

Figure 1: Contract duration: Termination reasons and match quality



(a) Termination reasons

(b) Contract duration by match quality

Note: Panel (a) presents the distribution of contract termination reasons, highlighting the relative frequency of different exit events. Panel (b) examines contract duration by match quality, showing the relationship between how well a worker-client pair aligns and the length of the employment relationship. * matching client preferences on worker's age, sex and origin and worker preferences on client's sex, coresidence, Alzheimer, bedridden status Source: Firm administrative data.

Results for individual preference dimension are reported in Appendix D.

These results indicate that stated preferences may be important for match stability. Mismatches——whether in terms of age, gender, nationality, or work requirements——are associated with shorter employment relationships and higher turnover. The estimated matching model presented below will be able to shed light on the causal effect on match quality on job stability.

3 Model

The model presented in the section draws on the literature on marriage matching (Choo and Siow, 2006; Dupuy and Galichon, 2014; Chiappori, 2020; Ciscato et al., 2020; Galichon and Salanié, 2022), where 1-to-1 matching has been analyzed in great depth. It is augmented to capture key aspects of the matching process in the private care market. A central advantage of the care data is that they record expressed preferences over specific worker characteristics rather than merely observable attributes. Moreover, the observation of contractual transfers allows me to identify how the overall surplus generated by each match is distributed between clients and workers (Dupuy and Galichon, 2022). Finally, the availability of data on separations, which typically are unobservable in matching applications using cross-sectional data, enables the identification of the absolute level of the match surplus.

Table 3: Match quality and contract duration

	(1)	(2)	(3)	(4)
Panel A				
Dependent variable: log contract duration (days)				
All dimensions matched	0.346	0.342	0.384	0.402
	(0.070)	(0.070)	(0.072)	(0.097)
Year effects		X	X	X
Province effects			X	X
Client & worker characteristics				X
Observations	2,380	2,380	2,369	2,348
	(5)	(6)	(7)	(8)
Panel B				
Dependent variable: log contract duration (days)				
# Dimensions matched	0.204	0.201	0.226	0.246
	(0.047)	(0.047)	(0.048)	(0.072)
Year effects		X	X	X
Province effects			X	X
Client & worker characteristics				X
Observations	2,380	2,380	2,369	2,348
Mean (censored) duration: 126 days, Median duration: 41 days				

Note: The table reports conditional correlations between match quality on contract duration, measured as the log number of days a job lasts. Columns (1)-(4) show the relation with and indicator for a perfect match across all stated preferences, while columns (5)-(8) show the relation with the number of matched preference dimensions. Standard errors are reported in parentheses. Source: Firm administrative data.

I model the matching of elderly individuals or their families ("clients") with care workers, as well as potential separations. Each client is characterized by a vector $\boldsymbol{x_i}$ (including e.g. sex or health status) as well as by a vector $\boldsymbol{p_i^y}$ of preferences regarding observed worker characteristics $\boldsymbol{y_j}$ (such as age, origin or sex). Similarly, workers hold preferences $\boldsymbol{p_j^x}$ over the observed characteristics $\boldsymbol{x}$ of clients. Preferences may include indifference.

3.1 Match formation

The values of matches are assumed to be linear as in Dupuy and Galichon (2014) and Ciscato et al. (2020), with the difference that matching can account for ex-

pressed preferences over specific characteristics rather than each side's characteristics only.[2] For any given match with observed characteristics and preference statements $\tilde{\boldsymbol{x}}_{\boldsymbol{i}} \equiv (1 \quad \boldsymbol{x_i} \quad \boldsymbol{p_i^y})'$ for client i and $\tilde{\boldsymbol{y}}_{\boldsymbol{j}} \equiv (1 \quad \boldsymbol{y_j} \quad \boldsymbol{p_j^x})'$ for worker j, the expected value to the client is $\tilde{\boldsymbol{x}}_{\boldsymbol{i}}' \boldsymbol{A} \boldsymbol{y_j}$, whereas the expected value to the worker is $\boldsymbol{x_i}' \boldsymbol{\Gamma} \tilde{\boldsymbol{y}}_{\boldsymbol{j}}$, where $\boldsymbol{A}$ and $\boldsymbol{\Gamma}$ are estimated parameter matrices. The payoffs are adjusted by a transfer $w(\tilde{\boldsymbol{x}}_{\boldsymbol{i}}, \tilde{\boldsymbol{y}}_{\boldsymbol{j}})$ that client i pays to worker j. In addition to these deterministic value components, there are factors ε_{ij} and η_{ij} unobserved to the econometrician, so that the overall payoff to client i from a match with worker j is given by

$$V_{ij}^{client} = \tilde{\boldsymbol{x}}_{\boldsymbol{i}}' \boldsymbol{A} \boldsymbol{y_j} - w(\tilde{\boldsymbol{x}}_{\boldsymbol{i}}, \tilde{\boldsymbol{y}}_{\boldsymbol{j}}) + \varepsilon_{ij}. \tag{1}$$

Similarly, the overall payoff to worker j from a match with client i is given by

$$V_{ij}^{worker} = \boldsymbol{x_i}' \boldsymbol{\Gamma} \tilde{\boldsymbol{y}}_{\boldsymbol{j}} + w(\tilde{\boldsymbol{x}}_{\boldsymbol{i}}, \tilde{\boldsymbol{y}}_{\boldsymbol{j}}) + \eta_{ij}. \tag{2}$$

With independently Gumbel distributed ε_{ij} and η_{ij} and corresponding spread parameters σ_ε and σ_η, we obtain the conditional probability that client i's value of being matched with worker j exceeds the value of being matched with any other worker

$$\pi(j|i) = \frac{\exp\left(\tilde{\boldsymbol{x}}_{\boldsymbol{i}}' \boldsymbol{A} \boldsymbol{y_j} - w(\tilde{\boldsymbol{x}}_{\boldsymbol{i}}, \tilde{\boldsymbol{y}}_{\boldsymbol{j}})\right)^{-\sigma_\varepsilon}}{\sum_{j'} \exp\left(\tilde{\boldsymbol{x}}_{\boldsymbol{i}}' \boldsymbol{A} \boldsymbol{y_{j'}} - w(\tilde{\boldsymbol{x}}_{\boldsymbol{i}}, \tilde{\boldsymbol{y}}_{\boldsymbol{j'}})\right)^{-\sigma_\varepsilon}}. \tag{3}$$

Similarly, the probability that worker j's value of being matched with client i exceeds the value of any other match equals

$$\pi(i|j) = \frac{\exp\left(\boldsymbol{x_i}' \boldsymbol{\Gamma} \tilde{\boldsymbol{y}}_{\boldsymbol{j}} + w(\tilde{\boldsymbol{x}}_{\boldsymbol{i}}, \tilde{\boldsymbol{y}}_{\boldsymbol{j}})\right)^{-\sigma_\eta}}{\sum_{i'} \exp\left(\boldsymbol{x_{i'}}' \boldsymbol{\Gamma} \tilde{\boldsymbol{y}}_{\boldsymbol{j}} + w(\tilde{\boldsymbol{x}}_{\boldsymbol{i'}}, \tilde{\boldsymbol{y}}_{\boldsymbol{j}})\right)^{-\sigma_\eta}}. \tag{4}$$

Note that whereas payoffs $\tilde{\boldsymbol{x}}_{\boldsymbol{i}}' \boldsymbol{A} \boldsymbol{y_j}$ and $\boldsymbol{x_i}' \boldsymbol{\Gamma} \tilde{\boldsymbol{y}}_{\boldsymbol{j}}$ must vary with characteristics of the respectively other party to not cancel out in these choice probabilities. This explains why the vectors $\tilde{\boldsymbol{x}}_{\boldsymbol{i}}$ and $\tilde{\boldsymbol{y}}_{\boldsymbol{j}}$ defined above can include a unit element, whereas vectors $\boldsymbol{x_i}$ and $\boldsymbol{y_j}$ cannot.

[2]Data on partner sex preferences are used in a marriage matching application by Ciscato and Goussé (2024).

3.2 Equilibrium Matching and Separations

Matches are assumed to dissolve with hazard rates proportional to the inverse of the match surplus, with a proportionality factor δ. This reflects the idea that better matches are more stable.[3] Under this assumption, the matchmaker's objective of maximizing aggregate match durations is achieved by maximizing the aggregate surplus. As shown by Galichon and Salanié (2022), the aggregate surplus in turn is maximized by the decentralized matching equilibrium. Focusing on the later, one can apply Bayes' rule to equations (3) and (4) to obtain in both cases the match probability for a pair (i, j):

$$\begin{aligned} \pi(i,j) &= \frac{\exp\left(\tilde{\boldsymbol{x}}_{\boldsymbol{i}}'\boldsymbol{A}\tilde{\boldsymbol{y}}_{\boldsymbol{j}} - w(\tilde{\boldsymbol{x}}_{\boldsymbol{i}}, \tilde{\boldsymbol{y}}_{\boldsymbol{j}})\right)^{-\sigma_\varepsilon}}{\sum_{j'} \exp\left(\tilde{\boldsymbol{x}}_{\boldsymbol{i}}'\boldsymbol{A}\tilde{\boldsymbol{y}}_{\boldsymbol{j'}} - w(\tilde{\boldsymbol{x}}_{\boldsymbol{i}}, \tilde{\boldsymbol{y}}_{\boldsymbol{j'}})\right)^{-\sigma_\varepsilon}} f(i) \\ &= \exp\left(\tilde{\boldsymbol{x}}_{\boldsymbol{i}}'\boldsymbol{A}\tilde{\boldsymbol{y}}_{\boldsymbol{j}} - w(\tilde{\boldsymbol{x}}_{\boldsymbol{i}}, \tilde{\boldsymbol{y}}_{\boldsymbol{j}}) - a(\tilde{\boldsymbol{x}}_i)\right)^{-\sigma_\varepsilon}, \end{aligned} \tag{5}$$

where $a(\tilde{\boldsymbol{x}}_i) = \log \sum_{j'} \exp\left(\tilde{\boldsymbol{x}}_{\boldsymbol{i}}'\boldsymbol{A}\tilde{\boldsymbol{y}}_{\boldsymbol{j'}} - w(\tilde{\boldsymbol{x}}_{\boldsymbol{i}}, \tilde{\boldsymbol{y}}_{\boldsymbol{j'}})\right)^{-\sigma_\varepsilon} - \sigma_\varepsilon \log f(i)$ and $f(i)$ denotes the density over clients. Similarly,

$$\begin{aligned} \pi(i,j) &= \frac{\exp\left(\boldsymbol{x}_{\boldsymbol{i_m}}'\boldsymbol{\Gamma}\tilde{\boldsymbol{y}}_{\boldsymbol{j_m}} + w(\tilde{\boldsymbol{x}}_{\boldsymbol{i}}, \tilde{\boldsymbol{y}}_{\boldsymbol{j}})\right)^{-\sigma_\eta}}{\sum_{i'} \exp\left(\boldsymbol{x}_{\boldsymbol{i'}}'\boldsymbol{\Gamma}\tilde{\boldsymbol{y}}_{\boldsymbol{j}} + w(\tilde{\boldsymbol{x}}_{\boldsymbol{i'}}, \tilde{\boldsymbol{y}}_{\boldsymbol{j}})\right)^{-\sigma_\eta}} g(j) \\ &= \exp\left(\boldsymbol{x}_{\boldsymbol{i_m}}'\boldsymbol{\Gamma}\tilde{\boldsymbol{y}}_{\boldsymbol{j_m}} + w(\tilde{\boldsymbol{x}}_{\boldsymbol{i}}, \tilde{\boldsymbol{y}}_{\boldsymbol{j}}) - b(\tilde{\boldsymbol{y}}_j)\right)^{-\sigma_\eta}, \end{aligned} \tag{6}$$

with $b(\tilde{\boldsymbol{y}}_j) = \log \sum_{i'} \exp\left(\boldsymbol{x}_{\boldsymbol{i_m}}'\boldsymbol{\Gamma}\tilde{\boldsymbol{y}}_{\boldsymbol{j_m}} + w(\tilde{\boldsymbol{x}}_{\boldsymbol{i}}, \tilde{\boldsymbol{y}}_{\boldsymbol{j'}})\right)^{-\sigma_\eta} - \sigma_\eta \log g(j)$ and density $g(j)$ over workers.

Equating (5) and (6) and rearranging yields the equilibrium match probability for client i and worker j

$$\pi(i,j) = \exp\left(\tilde{\boldsymbol{x}}_{\boldsymbol{i}}'\boldsymbol{A}\tilde{\boldsymbol{y}}_{\boldsymbol{j}} + \boldsymbol{x}_{\boldsymbol{i}}'\boldsymbol{\Gamma}\tilde{\boldsymbol{y}}_{\boldsymbol{j}} - a(\tilde{\boldsymbol{x}}_i) - b(\tilde{\boldsymbol{y}}_j)\right)^{-(\sigma_\varepsilon+\sigma_\eta)}, \tag{7}$$

where $a(\tilde{\boldsymbol{x}}_i)$ and $b(\tilde{\boldsymbol{y}}_j)$ capture the outside options of the client and the worker, respectively. These outside options are uniquely determined (up to a constant) by the system

$$\sum_{j=1}^{N} \exp\left(\tilde{\boldsymbol{x}}_{\boldsymbol{i}}'\boldsymbol{A}\tilde{\boldsymbol{y}}_{\boldsymbol{j}} + \boldsymbol{x}_{\boldsymbol{i}}'\boldsymbol{\Gamma}\tilde{\boldsymbol{y}}_{\boldsymbol{j}} - a(\tilde{\boldsymbol{x}}_i) - b(\tilde{\boldsymbol{y}}_j)\right)^{-(\sigma_\varepsilon+\sigma_\eta)} = 1/N, \quad \forall i = 1, ..., N, \tag{8}$$

[3] Appendix Figure 18 shows that "exogenous" separations due to the care recipient's death or other reasons, which would include transfers to institutional care, are not only the least common reason for match dissolution, but that they do not correlate with clients' age or health status in the data.

$$\sum_{i=1}^{N} \exp\left(\tilde{\boldsymbol{x}}_i' \boldsymbol{A} \boldsymbol{y}_j + \boldsymbol{x}_i' \boldsymbol{\Gamma} \tilde{\boldsymbol{y}}_j - a(\tilde{\boldsymbol{x}}_i) - b(\tilde{\boldsymbol{y}}_j)\right)^{-(\sigma_\varepsilon + \sigma_\eta)} = 1/N, \quad \forall j = 1, ..., N, \qquad (9)$$

where N is the number of matches in the data.

3.3 Surplus identification

A challenge in standard cross-sectional matching models is that the outside options for each agent, $a(\tilde{\boldsymbol{x}}_i)$ for clients and $b(\tilde{\boldsymbol{y}}_j)$ for workers, are identified only up to an additive constant. As noted by Dupuy and Galichon (2022), the logit structure of unobserved heterogeneity implies that unmatched agents do not affect the relative probabilities of observed matches. Consequently, without further data, one can only recover the relative match surplus, and an arbitrary normalization (e.g., $a_1 = 0$) must be imposed.

The care data offer a key advantage by including information on match separations. These separations provide additional variation that permits the level of match surplus to be identified without resorting to an arbitrary normalization. Specifically, the unique determination of the value of outside options $a(\tilde{\boldsymbol{x}}_i)$ and $b(\tilde{\boldsymbol{y}}_j)$ allow to recover the full surplus generated by each match.

Note that equations (7)-(9) could be rewritten in terms of $\boldsymbol{A} + \boldsymbol{\Gamma}$, implying that the distribution of matches across types characterizing agents i and j alone cannot separately identify the payoffs to the two sides. Separate identification of the parameter matrices $\boldsymbol{A}$ and $\boldsymbol{\Gamma}$ however is feasible with the availability of transfers between clients and workers. This is important conceptually, because match characteristics may be valued differently by the two sides. For instance, assortative matching on sex may be differently important to care givers and care recipients.

In sum, by jointly exploiting information from observed matches, contractual wages, and separations, the model identifies all the parameters in above expressions, which are crucial for understanding the welfare implications of match quality in the private care market.

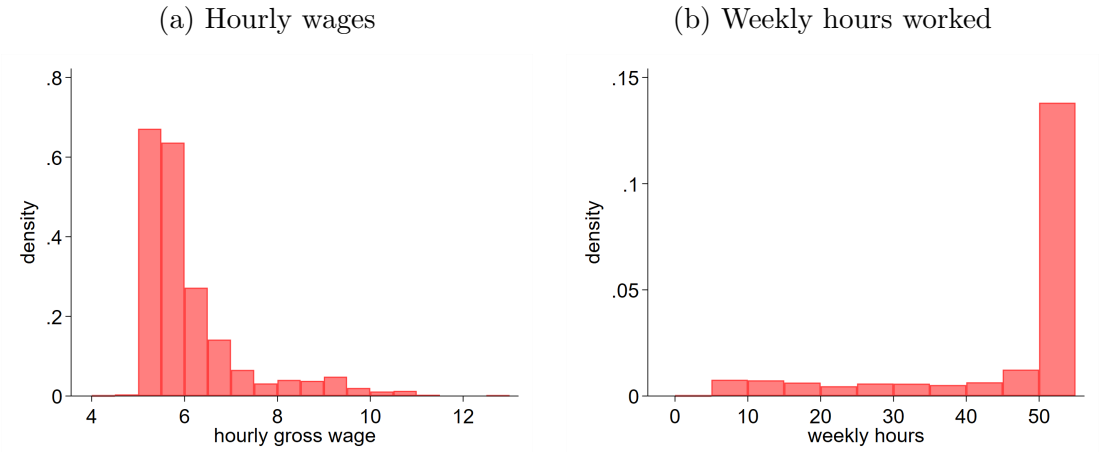
3.4 Transfers

Solving (5)-(6) for $w(\tilde{\boldsymbol{x}}_i, \tilde{\boldsymbol{y}}_j)$ yields the equilibrium transfer paid by client i to worker j:

$$w(\tilde{\boldsymbol{x}}_i, \tilde{\boldsymbol{y}}_j) = \frac{\sigma_\eta}{\sigma_\varepsilon + \sigma_\eta}\left(\tilde{\boldsymbol{x}}_i' \boldsymbol{A} \boldsymbol{y}_j - a(\tilde{\boldsymbol{x}}_i)\right) + \frac{\sigma_\varepsilon}{\sigma_\varepsilon + \sigma_\eta}\left(b(\tilde{\boldsymbol{y}}_j) - \boldsymbol{x}_i' \boldsymbol{\Gamma} \tilde{\boldsymbol{y}}_j\right)$$

I observe contractual wages W_{ij} in the data, displayed in panel (a) of Figure 2. Contractual wages are bounded from below due to collective bargaining agreements. The mean wage in the data is 6.10 euros per hour, with a median of 5.85 euros. Panel (b) presents weekly hours stated in the contract, with bunching at the legal maximum of 54 hours per week.[4]

Figure 2: Distributions of weekly hours worked and hourly wages



Note: This figure illustrates the distribution of hourly wages (panel (a)) and weekly contractual hours (panel (b)). Source: Firm administrative data.

A discrepancy between observed transfers in terms of payment and true transfers may arise if work hours exceed the contractual number of hours used to calculate wages or if part of the compensation is in-kind. I thus allow for observed wages W_{ij} to be an imperfect measure of the true transfers $w(\tilde{\boldsymbol{x}}_{\boldsymbol{i}}, \tilde{\boldsymbol{y}}_{\boldsymbol{j}})$, such that

$$W_{ij} \quad = \quad \alpha + w(\tilde{\boldsymbol{x}}_{\boldsymbol{i}}, \tilde{\boldsymbol{y}}_{\boldsymbol{j}}) + \omega_{ij}, \tag{10}$$

with potential shifter α and log-normally distributed measurement error $\log(\omega_{ij}) \sim \mathcal{N}(0, \sigma_\omega^2)$.

3.5 Estimation

The model is estimated iteratively using match characteristics, wage observations, and job separations. The parameters to be estimated include preference parameters $\{\boldsymbol{A}, \boldsymbol{\Gamma}, \sigma_\varepsilon, \sigma_\eta\}$, wage function parameters $\{\alpha, \sigma_\omega\}$ and the hazard factor δ that governs match dissolution. All parameters are estimated jointly as follows.

[4] Appendix Figure 16 shows the distribution of monthly earnings, both in full-time equivalent and actual gross wages.

First, conditional on the parameters $\boldsymbol{A}$, $\boldsymbol{\Gamma}$, σ_ε and σ_η, the outside options $a(\tilde{\boldsymbol{x}}_i)$ and $b(\tilde{\boldsymbol{y}}_j)$ are retrieved by iterating over equations (8) and (9).

Second, to recover the wage function parameters, observed contractual wages W_{ij} are regressed on $(\tilde{\boldsymbol{x}}_{\boldsymbol{i}}'\boldsymbol{A}\boldsymbol{y}_{\boldsymbol{j}} - a(\tilde{\boldsymbol{x}}_i))$ and $(b(\tilde{\boldsymbol{y}}_j) - \boldsymbol{x}_{\boldsymbol{i}}'\boldsymbol{\Gamma}\tilde{\boldsymbol{y}}_{\boldsymbol{j}})$. Based on (10), this regression yields estimates for the shift parameter α and for the ratios $\sigma_\eta/(\sigma_\varepsilon+\sigma_\eta)$ and $\sigma_\varepsilon/(\sigma_\varepsilon+\sigma_\eta)$, while its residual variance provides an estimate of σ_ω^2.

Finally, I construct the joint likelihood function that integrates the likelihood contributions from match formation, wage observation, and match separation:

$$\begin{aligned} LL \;=\; & \sum_{m=1}^{N}\left(\frac{\tilde{\boldsymbol{x}}_{\boldsymbol{i_m}}'\boldsymbol{A}\boldsymbol{y}_{\boldsymbol{j_m}} + \boldsymbol{x}_{\boldsymbol{i_m}}'\boldsymbol{\Gamma}\tilde{\boldsymbol{y}}_{\boldsymbol{j_m}} - a_{i_m} - b_{j_m}}{\sigma_\varepsilon+\sigma_\eta}\right) \\ & -\sum_{m=1}^{N}\left(\frac{W_{i_m j_m} - \alpha - w(\tilde{x}_{i_m},\tilde{y}_{j_m})}{2\sigma_\omega^2}\right) - \frac{N}{2}\log\sigma_\omega^2 \\ & +\sum_{m=1}^{N}\Big(separated_m\left(-duration_m\Big/(\delta\bar{V}_{i_m,j_m}) - \log(\delta\bar{V}_{i_m,j_m})\right) \\ & +(1-separated_m)\left(-duration_m\Big/(\delta\bar{V}_{i_m,j_m})\right)\Big), \end{aligned} \tag{11}$$

where the term

$$\bar{V}_{i_m,j_m} = \tilde{x}_{i_m}'A\,y_{j_m} + x_{i_m}'\Gamma\tilde{y}_{j_m} - (\sigma_\varepsilon+\sigma_\eta)\log\big(\pi(i_m,j_m)\big) \tag{12}$$

captures the expected match surplus for each observed match m.

The joint likelihood is maximized with respect to the parameters $\boldsymbol{A}$, $\boldsymbol{\Gamma}$ and δ, and the three-step procedure is iterated until convergence is achieved.

4 Results

The estimated parameters provide insight into how clients and workers value different attributes in the private care market. Table 4 reports the estimated payoff matrix $\boldsymbol{A}$ for clients, while Table 5 shows the corresponding parameters in $\boldsymbol{\Gamma}$ for workers. To focus on preferences and limit the number of parameters that need estimation, $\boldsymbol{A}$ and $\boldsymbol{\Gamma}$ are constrained such that an agent's payoff depends on whether their own preferences are met by the other party's corresponding characteristic. In addition, both matrices include full interactions between the two sides' characteristics described in Appendix Figure 17.

The estimated client valuations in Table 4 confirm that on average clients derive a higher surplus from matching with workers who align with their stated

Table 4: Client valuations (Payoff parameters in $\boldsymbol{A}$)

	Young Worker	Male Worker	Non-European Worker
Client Preferences			
Prefers younger worker	0.166		
	(0.000)		
Prefers older worker	-0.260		
	(0.002)		
Prefers female worker		-0.930	
		(0.064)	
Prefers male worker		0.825	
		(0.068)	
Prefers European worker			-0.249
			(0.050)
Prefers non-European worker			0.120
			(0.058)
Client Characteristics			
Client is male	-0.132	0.282	-0.066
	(0.034)	(0.057)	(0.059)
Client demands co-residence	-0.105	0.198	-0.034
	(0.034)	(0.060)	(0.053)
Client is bedridden	-0.080	-0.629	0.000
	(0.056)	(0.056)	(0.039)
Client has Alzheimer	0.096	0.428	-0.053
	(0.051)	(0.053)	(0.059)
Baseline effect	0.122	-0.082	0.173
	(0.061)	(0.057)	(0.057)

Note: This table presents the estimated payoff matrix $\boldsymbol{A}$, capturing client valuations for different match characteristics, denoted in euros/hour. Indifference in preferences over workers' age, sex and origin are omitted categories, as are workers being aged 40-60, female or European. Asymptotic standard errors are reported in parentheses. Source: Firm administrative data.

preferences regarding age, sex, and origin. The largest penalty, 93 cents/hour, occurs when clients express a preference for female caregivers are matched with a male worker. This is sizable, in particular compared to a median wage of 5.85 euros/hour.

While the estimated baseline coefficients indicate that for clients who are in-

different, being matched with young or non-European workers raises the surplus. These gains, however are eroded for clients who explicitly prefer older or European workers. These estimates confirm that preferences significantly shape the subjective quality of a match.

The bottom part of Table 4 captures the full interaction between care recipient and worker characteristics. These interactions largely act as controls. But the interaction between a client's and a worker's sex shows that for clients who report indifference over workers' sex, there is value to assortative matching on sex. Yet, the value of 28 cents/hour is only a third the magnitude estimated for the effects of stated preferences in the upper part of the table.

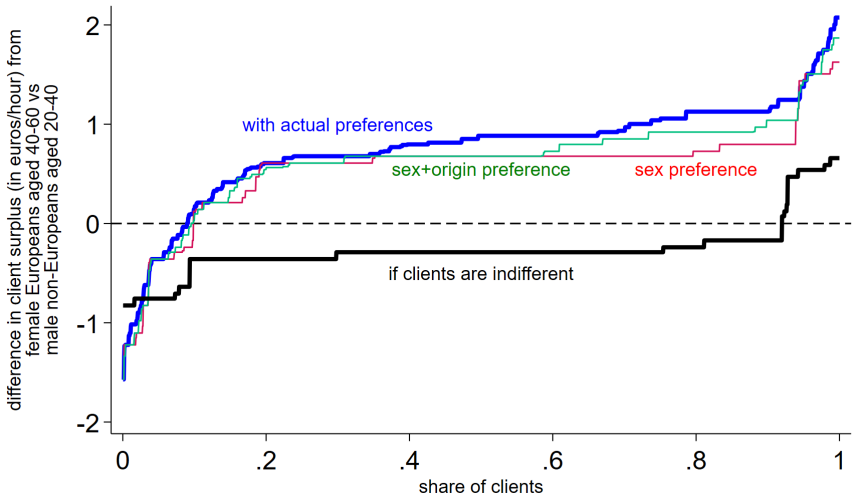
To illustrate the importance of preferences, Figure 3 shows variation in client valuations depending on worker characteristics and different preference scenarios. The figure displays the difference between the payoff to clients when matched with a female European worker aged 40–60 (the most frequently demanded set of attributes) versus a male non-European worker aged 20–40. The blue line depicts this difference for the actual distribution of client preferences in the data. It shows that for the vast majority of clients, the former type of worker generates a higher payoff, equal to about 1 euro/hour for the median client. For comparison, the black line represents the hypothetical case in which clients are indifferent to worker attributes. In that case, younger male non-European workers actually generate a slightly higher payoff to clients than their older female European counterparts. The difference between the blue and the black lines can be attributed to client preferences. The red line re-introduces clients preferences over workers' sex, showing that these bridge the larges part of the gap. The green line adds origin preferences, with the remaining gap deriving from clients' preferences over workers' age.

For the worker side, Table 5 displays the estimated payoff matrix $\mathbf{\Gamma}$. The estimates show that caregivers also place weight on matching with clients whose characteristics conform to their preferences.

Workers who prefer female clients experience a penalty of 58 cents/hour when assigned to men, although smaller than the corresponding loss to clients demanding a female carer (reported in Table 4). The payoff to co-residence is strong and positive, whereas being matched with a bedridden patient implies a loss of 61 cents/hour. The latter does not actually vary with caregivers' stated availability. These results underscore that preferences operate on both sides of the market.

As in Table 4, the bottom part of Table 5 captures the full interaction between care recipient and worker characteristics. sorting on own sex is more important

Figure 3: Client payoffs by worker type and preference scenario



Note: The figure shows the distribution of the difference in client payoffs (in euros/hour) when matched with a female European worker aged 40-60 versus a male non-European worker aged 20-40. The blue line represents client payoffs under different preference scenarios: a counterfactual scenario where clients are indifferent to all three attributes (black); preferences over workers' sex only (red); preferences workers' sex and origin (green); and actual preferences over workers' age, sex and origin (blue). Positive values indicate higher surplus from the female European worker. Source: Firm administrative data.

to workers than to clients, with male caregivers who do not report an explicit preference enjoying a 86 percent higher payoff when being matched with a male patient. The positive values of co-residence to workers is lower for male workers, but higher for non-European workers. All other estimated model parameters are listed in Appendix Table G.

Taken together, these results confirm that preferences significantly affect match quality and surplus. They also invite two broader questions. First, what are the effects of mismatch on separations? The estimates suggest that mismatches over worker sex, origin, or health conditions could drive higher turnover, as further illustrated in Figure 4. In particular, the figure shows that mismatches on either caregiver or care recipient sex shortens employment durations by between 2 and 6 percent.

Second, which elderly populations will grow in the near future? Since preferences are unevenly distributed across demographics, changes in the composition of elderly populations may shifts future client preferences. The next section exploits full register data to examine this question.

Table 5: Worker valuations (Payoff parameters in $\mathbf{\Gamma}$)

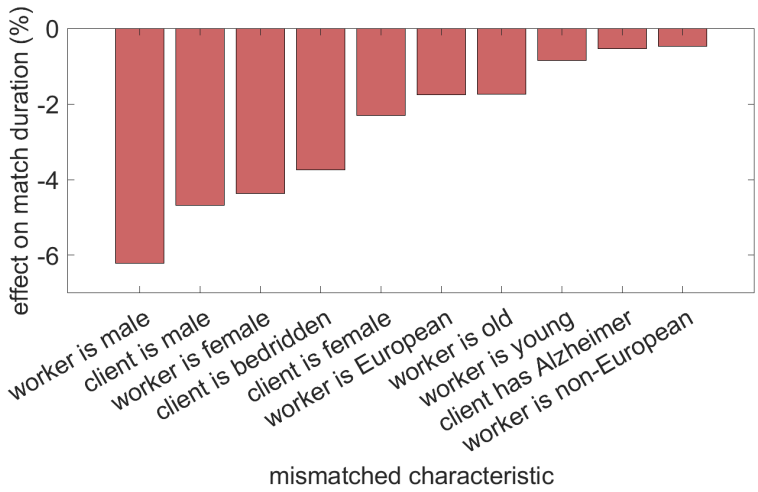
	Male Client	Co-residence	Bedridden Client	Alzheimer Client
Worker Preferences				
Prefers female patients	-0.584			
	(0.064)			
Prefers male patients	0.564			
	(0.055)			
Available for co-residence		1.707		
		(0.060)		
Available for bedridden patients			-0.033	
			(0.052)	
Available for Alzheimer patients				0.181
				(0.052)
Worker Characteristics				
Worker is young	-0.063	0.095	-0.012	-0.098
	(0.049)	(0.058)	(0.051)	(0.055)
Worker is male	0.864	-0.323	0.089	-0.164
	(0.055)	(0.059)	(0.051)	(0.053)
Worker is non-European	0.023	0.295	-0.022	0.155
	(0.066)	(0.060)	(0.061)	(0.061)
Baseline effect	-0.180	0.591	-0.611	-0.089
	(0.057)	(0.053)	(0.052)	(0.050)

Note: The table presents the estimated payoff matrix $\mathbf{\Gamma}$ capturing worker valuations for different match characteristics, denoted in euros/hour. Indifference in preferences over clients' sex is an omitted category, as is workers' inavailability for co-residence, patients with Alzheimer's disease or bedridden patients, and clients being female, not demanding co-residence, not having Alzheimer's and not being bedridden. Asymptotic standard errors are reported in parentheses. Source: Firm administrative data.

5 Outlook from current population registers

To assess future demand for elderly care, I use full register panel data by municipality, sex, and age. I project future elderly populations as follows. First, I compute the geometric mean of survival rates during the last five years prior to the pandemic (2015–2019) for each municipality-sex-age cell (approximately 1.5 million non-empty cells). Note that "survival" rates can be larger than one, accounting for population growth for instance due to migration. Second, I apply

Figure 4: Effect of mismatch on match duration (%)



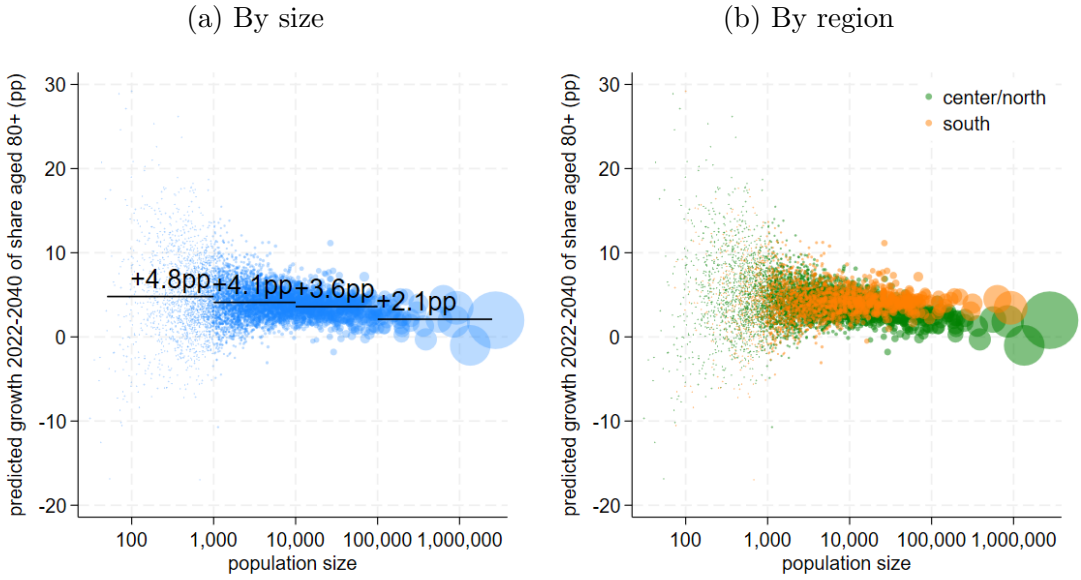
Note: The figure shows the estimated percentage change in match duration associated with various types of mismatches between client and worker characteristics. Negative values indicate a reduction in match duration. Source: Firm administrative data.

these survival rates to predict population changes for each cell from 2022 to 2040. I choose to not project further into future, since an implicit assumption in this projection is that life expectancy is as during the 2015-2019 period. For most of this section, I focus on the evolution of the adult population share aged 80 and older, as this group drives the demand for long-term care.

Figure 5 plots for each municipality the percentage point increase in the share of 80+ old resident against population sizes (on a log scale), with size also indicated by dot sizes. Panel (a) shows that smaller and midsized municipalities are expected to experience the largest increases in their elderly population shares. In the smallest municipalities (fewer than 1,000 residents), the share of 80+ individuals is projected to grow by 4.8 percentage points, compared to just 2.1 percentage points in large cities. The second panel adds a regional dimension, revealing that these demographic shifts are particularly pronounced in the south, where municipalities are expected to see a disproportionate rise in the elderly population. This suggests that future demand for elderly care services will increasingly be concentrated in these areas. Appendix Figure 19 further visualizes predicte changes by age and sex.

To understand how these demographic changes interact with client preferences, I merge these projections with clients in the company data based on their municipality, sex and age. This allows me to assess correlations between predicted elderly population growth and preferences. If preferences remain stable over time, shifts

Figure 5: Projected growth in the 80+ population by municipality size (top) and by municipality size and region (bottom).



Note: Panel (a) shows the projected growth in the adult population share aged 80 and older between 2022 and 2040 across municipalities of different sizes. Panel (b) further differentiates these projections by region, highlighting differences between southern and central/northern municipalities. Larger dots indicate municipalities with a greater total population. Source: Own computation based on ISTAT municipal population register data 2015-2022.

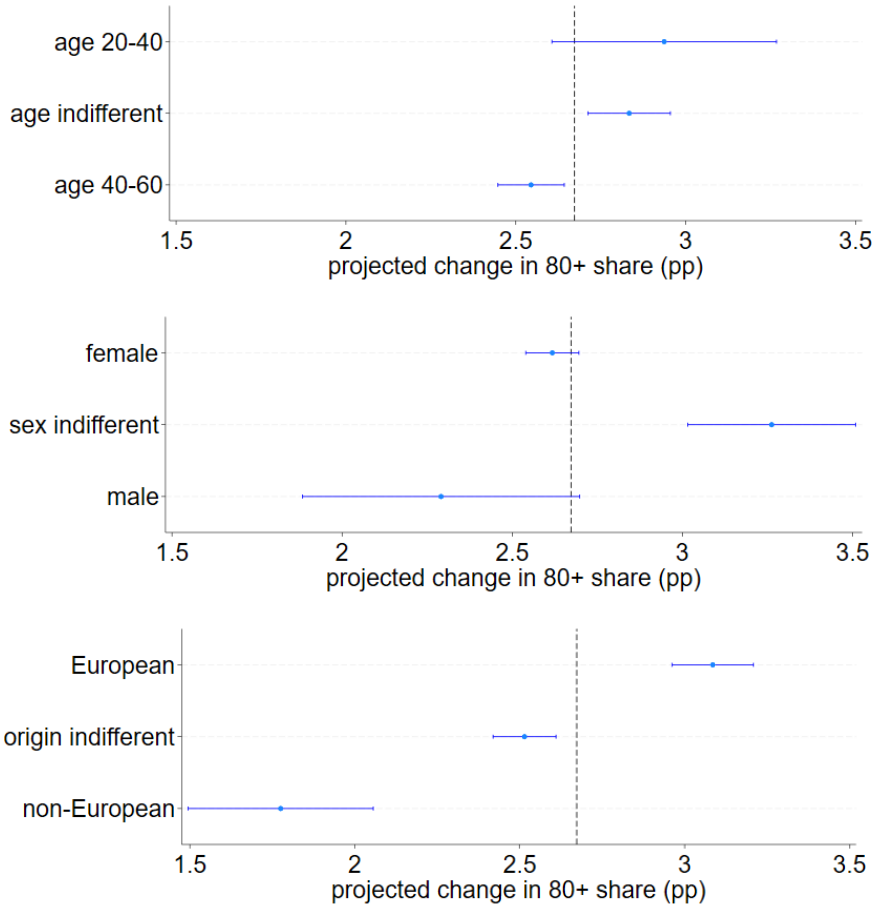
in the population structure will influence the types of workers that are in higher demand.

Figure 6 illustrates predicted 80+ population shares are related to client preferences for worker characteristics. Although the company has clients all over the country, northern Italian municipalities are overrepresented. Mean growth in the adult population share aged 80 and older for the demographics in the company data, indicated by the vertical dashed lines, hence is lower than the country average.

The top panel shows how demand for caregivers of different ages is projected to shift, while the middle and bottom panels depict projected changes in preferences for worker sex and origin, respectively. These projections indicate an erosion of the bias towards middle-aged care workers and more flexibility regarding care workers' sex. These shifts derive from a relatively stronger increase in men and in individuals in their low 80s among all individuals aged 80 and older, see also Appendix Figure 19. Appendix Figure 22 indicates a correlation between care recipients' and preferred caregivers' age, whereas Appendix Figure 23 shows that

male care recipients are more likely indifferent about a care worker's sex.

Figure 6: Projected change in 80+ share by client preference for worker age (top), origin (middle), and sex (bottom).

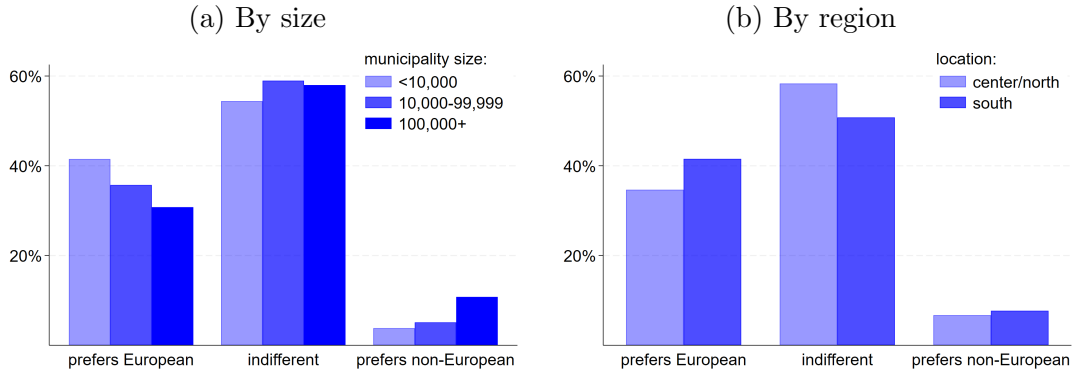


Note: The figure presents projected changes in the share of the population aged 80 and older between 2022 and 2040, disaggregated by client preferences for worker age (top), worker origin (middle), and worker sex (bottom). These projections are based on demographic shifts in the elderly population and assume persistence in the distribution of preferences over time. Source: ISTAT municipal population register data, firm data on client preferences.

The strongest trend is apparent for origin preferences, with growth in elderly population shares be stronger among demographics more likely to state a preference for European caregivers and less likely to prefer non-Europeans. Figure 7 shows that this is driven by clients in smaller and southern municipalities more often expressing a preference for European care workers. Adult population shares aged 80 and older are predicted to grow particularly strongly in these municipalities (see Figure 5).

These findings raise important welfare questions. If the growing demand for elderly care services can only be met by non-European workers, the observed

Figure 7: Client preferences for worker origin by municipality size (top) and by region (bottom).



Note: The figure shows client preferences for caregivers' origin for different locations. Panel (a) shows the distribution of preferences by municipality size, panel (b) presents preferences by region. Source: ISTAT municipal population register data, firm administrative data.

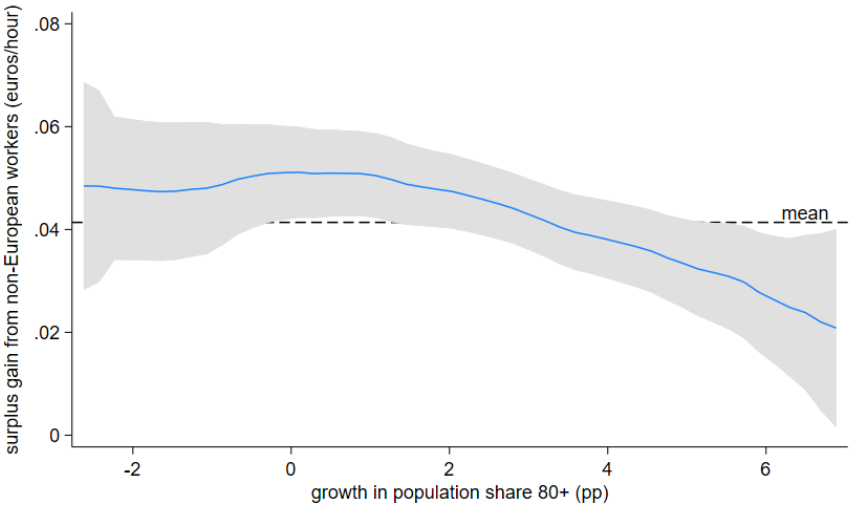
preferences in certain regions and municipalities may not be matched, with ensuing surplus losses.

To further assess the implications of this, Figure 8 examines the relationship between the projected growth in the 80+ population share and the difference in surplus from matching with non-European compared to European care workers. Recall from Table 4 that non-European workers generate a higher baseline value, which however turns negative for clients with a preference for European caregivers. Pooling clients of all kinds of preferences yields a rather small average surplus difference. The dashed horizontal line in Figure 8 indicates an average of only 4 cents per hour. The smoothed downward-sloping schedule illustrates a further erosion in communities with a high predicted growth in the share of elderly residents. Yet, the estimated values do not support the notion of a negative aggregate effect of non-European carers on care recipients' welfare even in municipalities with very high predicted growth rates.

6 Conclusion

This paper draws on novel company data that provide unique insights into preferences and matching between elderly care recipients and private care workers, offering a detailed view of both stated on each side of the market. Through the lens of a matching model, I estimate match values as functions of clients' (care

Figure 8: Surplus gain from non-European workers by projected growth in the 80+ share.



Note: The figure depicts the relationship between the projected growth in the 80+ population share and the surplus gain from non-European workers. The blue line represents the difference in estimated surplus from matches with a non-European compared to a European worker, with the shaded area indicating the 95% confidence interval. The mean difference is indicated by the dashed line. Source: Own elaboration based on firm data and demographic projections from ISTAT.

recipients or their family) and workers' preference over characteristics of the respectively other side, such as age, sex, origin, and health conditions. The results reveal substantial penalties for mismatches, including a reduction in surplus of up to 15 percent of the average wage when clients who prefer a female worker are assigned a male caregiver. Such a mismatch reduces expected contract duration by 6 percent. Compared to clients who are indifferent about a worker's origin, clients preferring European workers experience a loss equal to 4 percent the average wage when matched with a non-European worker. Non-European workers are estimated to generate higher baseline surpluses for clients who are indifferent about workers' origin. This surplus turns negative for clients with a preference for European workers, although the net loss translates into only a limited effect on contract duration.

Using municipal-level population registers, I project the evolution of the elderly populations and highlight that the share of individuals aged 80 and above is expected to increase most rapidly in smaller and southern municipalities. These areas also exhibit a stronger preference for European care workers. Nonetheless, the predicted impact on surplus from an incrased non-European workforce in the private care sector remains modest. Overall, this study underscores the importance

of accounting for explicit preferences in both policy design and market analyses, as demographic trends continue to reshape the private care sector. Future research might build on these findings by investigating how policy interventions, labor mobility, and technological changes could mitigate potential mismatches and improve outcomes for both clients and care workers.

References

Achou, B., Donder, P. D., Glenzer, F., Lee, M., and Leroux, M.-L. (2023). At Home Versus in a Nursing Home: Long-Term Care Settings and Marginal Utility. *CESifo Working Paper*, (10482).

Almuhaisen, A., Amuedo-Dorantes, C., and Furtado, D. (2023). Immigration Enforcement and the Institutionalization of Elderly Americans. *IZA Discussion Paper*, (16357).

Börsch-Supan, A., Brandt, M., Hunkler, C., Kneip, T., Korbmacher, J., Malter, F., Schaan, B., Stuck, S., and Zuber, S. (2013). Data Resource Profile: The Survey of Health, Ageing and Retirement in Europe (SHARE). *International Journal of Epidemiology*.

Brugiavini, A., Carrino, L., and Pasini, G. (2023). Long-term care in italy. *National Bureau of Economic Research Working Paper*, (31861).

Card, D., Colella, F., and Lalive, R. (2023). Gender Preferences in Job Vacancies and Workplace Gender Diversity. *National Bureau of Economic Research Working Paper*, (29350).

Chiappori, P.-A. (2020). The Theory and Empirics of the Marriage Market. *Annual Review of Economics*, 12:547–578.

Choo, E. and Siow, A. (2006). Who Marries Whom and Why. *Journal of Political Economy*, 114(1).

Ciscato, E., Galichon, A., and Goussé, M. (2020). Like attract like? a structural comparison of homogamy across same-sex and different-sex households. *Journal of Political Economy*, 128(2):740–781.

Ciscato, E. and Goussé, M. (2024). Matching on gender and sexual orientation. *IZA Discussion Paper*, (17420).

Cortes, P. and Pan, J. (2013). Outsourcing household production: Foreign domestic workers and native labor supply in hong kong. *Journal of Labor Economics*, 31(2):327–371.

Cortes, P. and Tessada, J. (2011). Low-skilled immigration and the labor supply of highly skilled women. *American Economic Journal: Applied Economics*, 3(3):88–123.

Costa-Font, J. and Vilaplana-Prieto, C. (2024). The Hidden Value of Adult Informal Care in Europe. *CESifo Working Paper*, (11535).

Dupuy, A. and Galichon, A. (2014). Personality Traits and the Marriage Market. *Journal of Political Economy*, 122(6).

Dupuy, A. and Galichon, A. (2022). A Note on the Estimation of Job Amenities and Labor Productivity. *Quantitative Economics*, 13:153–177.

Flabbi, L. (2010). Gender Discrimination Estimation in a Search Model with Matching and Bargaining. *International Economic Review*, 51(3).

Frimmel, W., Halla, M., Paetzold, J., and Schmieder, J. (2023). Health of Parents, their Children's Labor Supply, and the Role of Migrant Care Workers. *Journal of Labor Economics*. Accepted manuscript, DOI: `https://doi.org/10.1086/729102`.

Galichon, A. and Salanié, B. (2022). Cupid's Invisible Hand: Social Surplus and Identification in Matching Models. *Review of Economic Studies*, 89:2600–2629.

Grabowski, D. C., Gruber, J., and McGarry, B. (2023). Immigration, The Long-Term Care Workforce, and Elder Outcomes in the U.S. *National Bureau of Economic Research Working Paper*, (30960).

Gruber, J., McGarry, K. M., and Hanzel, C. (2023). Long-term care around the world. *National Bureau of Economic Research Working Paper*, (31882).

Haan, P. and Wnuk, I. (2024). The Effect of Migration on Careers of Natives: Evidence from Long-Term Care. *IZA Discussion Paper*, (16749).

Kuhn, P. and Shen, K. (2013). Gender Discrimination in Job Ads: Evidence from China. *Quarterly Journal of Economics*, 128(1):287–336.

Kuhn, P. and Shen, K. (2023). What Happens When Employers Can No Longer Discriminate in Job Ads? *American Economic Review*, 113(4):1013–1048.

Mommaerts, C. (2018). Are Coresidence and Nursing Homes Substitutes? Evidence from Medicaid Spend-Down Provisions. *Journal of Health Economics*, 59:125–138.

Stöhr, T. (2015). Siblings' Interaction in Migration Decisions: Who Provides for the Elderly Left Behind? *Journal of Population Economics*, 28:593–629.

A Elderly care in Europe

Figure 9: Share of 70+ year old care recipients by mode and country



Note: This figure shows the share of individuals aged 70 and above receiving institutional care, paid home care, or unpaid home care across selected European countries and Israel. The data are drawn from selected waves of the SHARE survey, including only those waves where long-term care information was collected. Source: SHARE data.

B Client questionnaires

This appendix contains the questionnaire used to collect information on firm clients. It includes details on the elderly person's health status, living arrangements, and preferences for care worker characteristics.

Figure 10: Client Questionnaire - Part 1

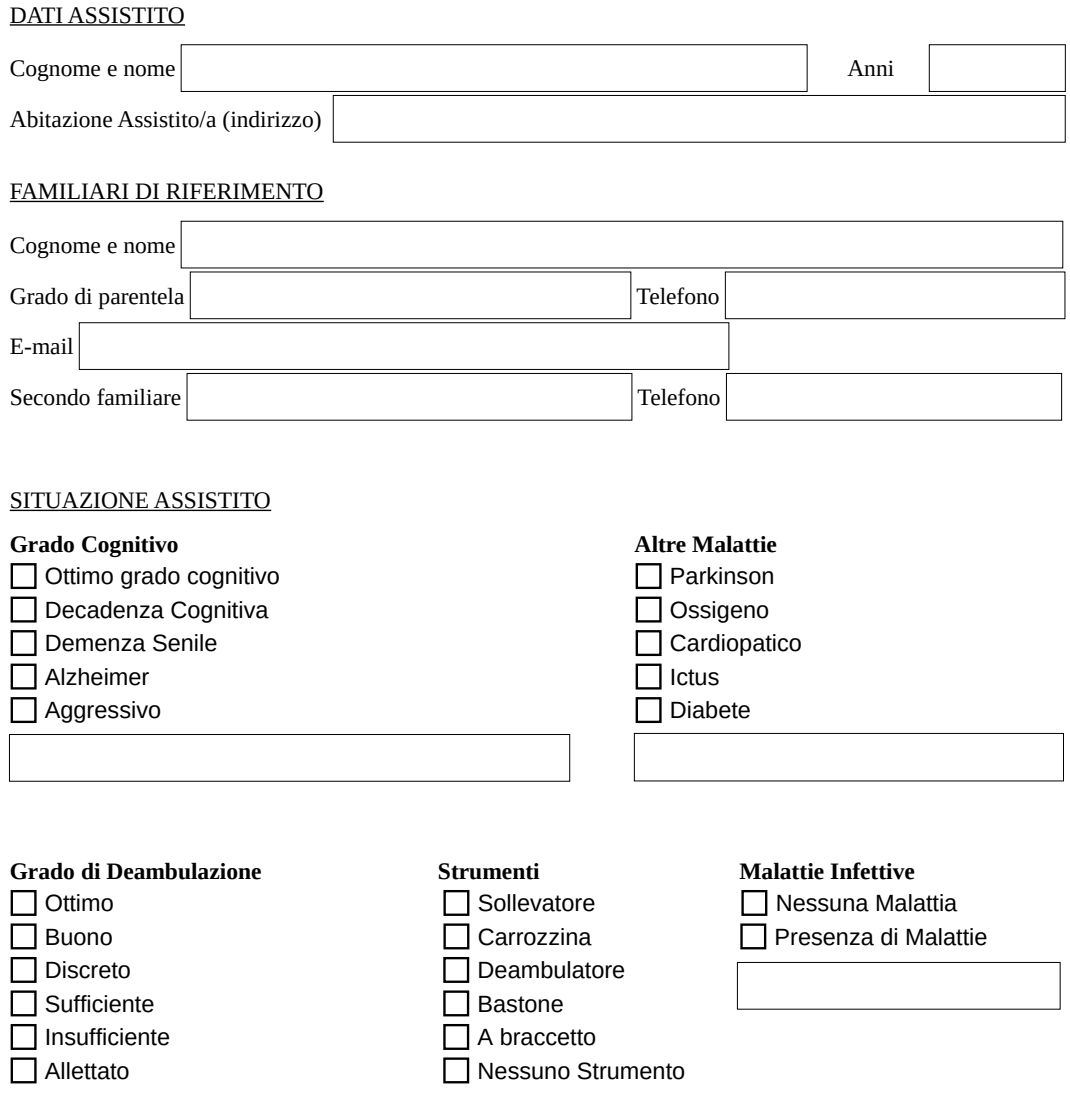


Figure 11: Client Questionnaire - Part 2

Igiene Personale
- ☐ Ottimo / totale autonomia
- ☐ Buono / accompagnare in bagno
- ☐ Sufficiente / parziale autonomia
- ☐ Insufficiente

Strumenti
- ☐ Catetere / Stomia
- ☐ Pappagallo
- ☐ Comoda
- ☐ Pannolone
- ☐ Nessuno Strumento

Notte
- ☐ Dorme tutta la notte
- ☐ Si sveglia raramente
- ☐ 2/3 volte alla settimana
- ☐ 1 volta tutte le notti
- ☐ Dalle 2 volte in sù tutte le notti (si necessita assistenza notturna con ulteriore collaboratore)

Assistenza Notturna
- ☐ Sì, è necessaria l'assistenza
- ☐ No, l'assistito/a è autonomo/a

Pasti
- ☐ Necessità di preparazione pasti
- ☐ Mangia in autonomia
- ☐ Ha bisogno di aiuto nel mangiare
- ☐ Deve essere imboccato

Corporatura
- ☐ Esile
- ☐ Normale
- ☐ Robusta

Peso _____ **Altezza** _____

Medicine
- ☐ La badante dà le medicine
- ☐ L'assistito è autonomo
- ☐ Ci pensano i familiari

Spesa
- ☐ La badante fa la spesa
- ☐ I familiari fanno la spesa

Altre patologie _____

Figure 12: Client Questionnaire - Part 3

FAMILIARE CONVIVENTE

Presenza di familiari conviventi

☐ Si

☐ No

Grado di parentela

Servizio COLF per familiare convivente (Stiro, preparazione pasti, lavaggio biancheria, ecc...)

☐ Si

☐ No

Presenza di animali domestici

☐ Si

☐ No

APPARTAMENTO

Grandezza appartamento in mq

Stanza privata per la badante ☐ Si ☐ No

Bagno privato per la badante ☐ Si ☐ No

Se NON è presente la stanza privata

☐ Divano letto in soggiorno

☐ Letto separato in camera con l'assistito/a

Figure 13: Client Questionnaire - Part 4

CARATTERISTICHE BADANTE

(Le caratteristiche indicate sono da ritenersi indicative e non vincolanti)

Sesso

☐ Donna
☐ Uomo
☐ Indifferente

Nazionalità

☐ Europa
☐ Sud America
☐ Africa
☐ Asia
☐ Indifferente

Fascia d'età

☐ 20 - 40
☐ 40 - 60
☐ Indifferente

Preferenza Caratteriale

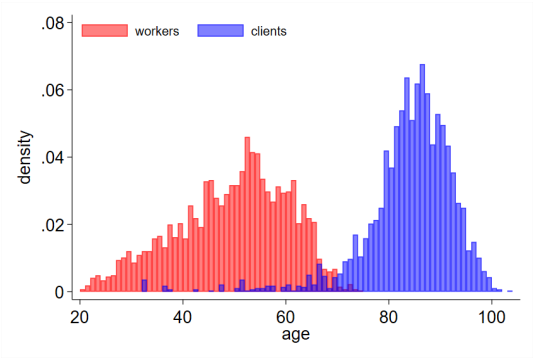
Preferenza data di inizio Servizio della Lavoratrice

(data indicativa e non vincolante per la ricerca)

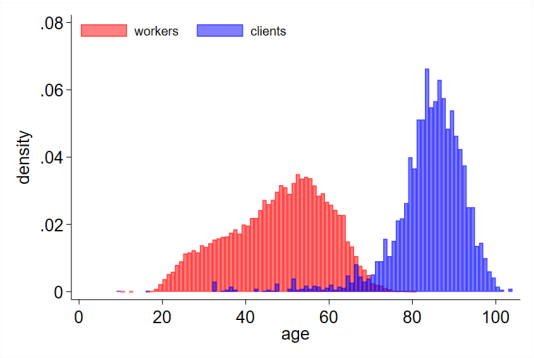
C Additional Summary Figures

Figure 14: Age distribution of clients and matched/all workers

(a) Matched sample

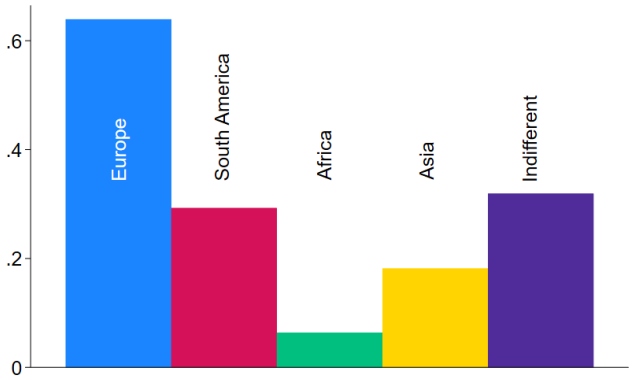


(b) Full data



Note: This figure presents the age distribution of clients and workers. Panel (a) shows the distribution for matched workers, while Panel (b) includes all clients and workers in the data. Source: Firm administrative data.

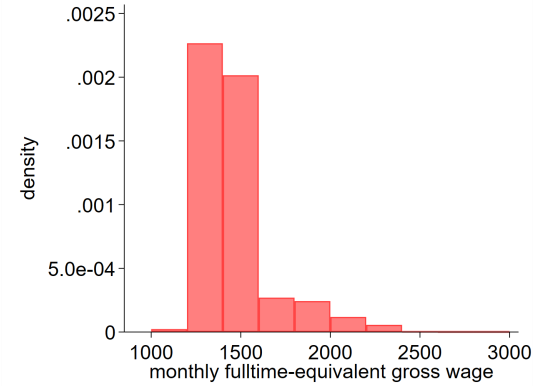
Figure 15: Client preferences over worker origin



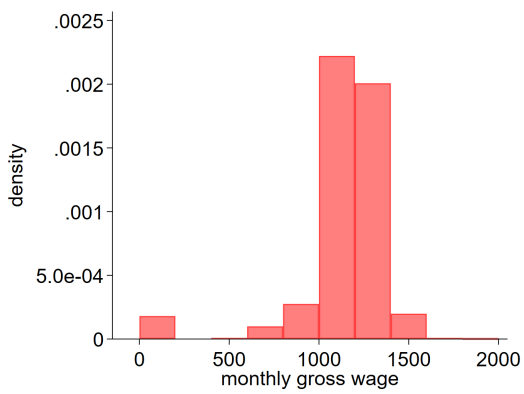
Note: This figure displays the distribution of client preferences regarding worker origin. Preferences are based on stated choices recorded at the time of match formation. Source: Firm administrative data.

Figure 16: Distribution of monthly wages

(a) Full-time equivalent monthly gross wages



(b) Monthly gross wages

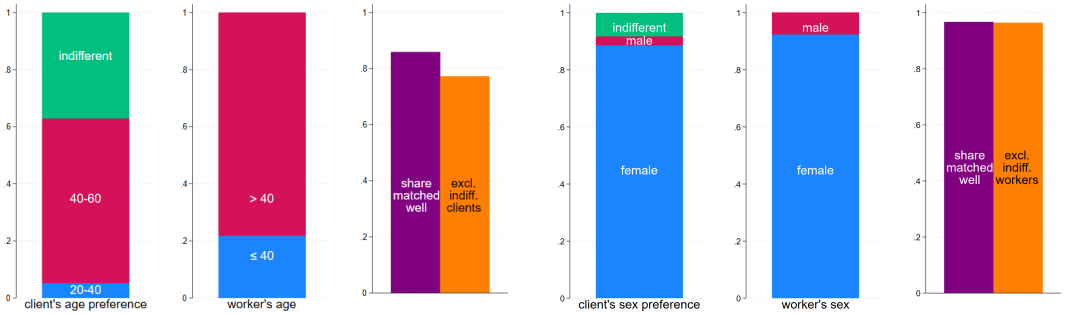


Note: This figure shows the distribution of monthly wages for care workers. Panel (a) presents full-time equivalent monthly gross wages, adjusting for differences in weekly hours worked, while Panel (b) displays the distribution of observed monthly gross wages. Source: Firm administrative data.

Figure 17: Match quality across different dimensions

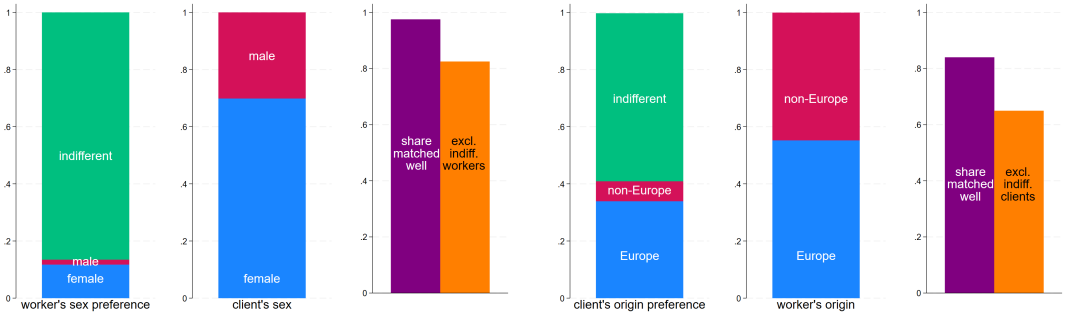
(a) Match quality on worker's age

(b) Match quality on worker's sex



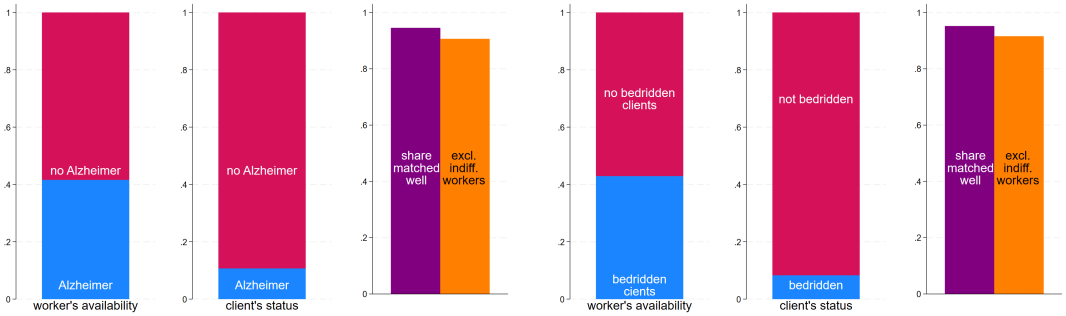
(c) Match quality on client's sex

(d) Match quality on worker's origin

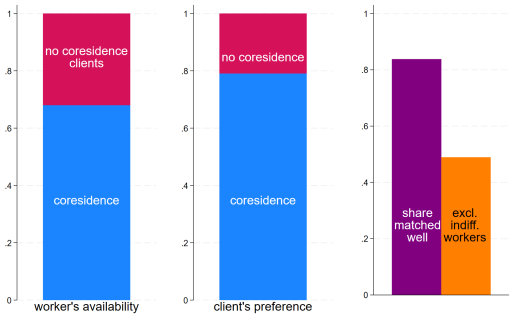


(e) Match quality on client's Alzheimer status

(f) Match quality on client's mobility



(g) Match quality on client's co-residence need



Note: The figure illustrates match quality across different dimensions, including worker and client characteristics such as age, sex, origin, health conditions, and living arrangements. Purple bars indicate the share of agents whose match does not violated their preference. Orange bars show this share excluding indifferent agents. Where indifference is not an elicited category (Panels (e)-(g)), availability is interpreted as indifference. Source: Firm administrative data.

D Contract Duration and Match Quality

This appendix presents additional analyses on the relationship between contract duration and match quality. Tables 6 and 7 report conditional correlations between contract length and match quality from clients' and workers' perspectives, respectively.

Table 6: Correlations between match quality (age, origin, sex) and contract duration

	(1)	(2)	(3)
Dependent variable: Log contract duration (days)			
Good match	0.174	0.257	0.280
	(0.141)	(0.108)	(0.102)
Match on	Age	Age + Origin	Age + Origin + Worker's Sex
Year effects	X	X	X
Province effects	X	X	X
Characteristics & preferences	X	X	X
# Matches	2,437	2,402	2,400
Mean (censored) duration: 126 days, Median duration: 41 days			

Note: The table reports conditional correlations between match quality and contract duration. Column (1) considers matches based only on workers' age, column (2) extends the match definition to include workers' age and origin, while column (3) further includes workers' sex. The dependent variable is the log of contract duration in days. Estimates control for year and province fixed effects, as well as client and worker characteristics and preferences. Source: Firm administrative data.

Table 7: Correlations between match quality (health conditions, co-residence, sex) and contract duration

	(1)	(2)	(3)	(4)
Dependent variable: Log contract duration (days)				
Good match	0.123	0.137	0.359	0.340
	(0.228)	(0.177)	(0.139)	(0.129)
Match on	Alzheimer	Alzheimer + Bedridden	Alzheimer + Bedridden + Co-residence	Alzheimer + Bedridden + Co-residence + Client's Sex
Year effects	X	X	X	X
Province effects	X	X	X	X
Characteristics & preferences	X	X	X	X
# Matches	2,499	2,471	2,461	2,458
Mean (censored) duration: 126 days, Median duration: 41 days				

Note: This table reports conditional correlations between match quality and contract duration, considering health conditions, co-residence, and client's sex. Column (1) accounts for match quality based on Alzheimer's status, column (2) adds bedridden status, column (3) further includes co-residence availability, and column (4) incorporates clients' sex. The dependent variable is the log of contract duration in days. All specifications control for year and province fixed effects, as well as client and worker characteristics and preferences. Source: Firm administrative data.

E Potential misreporting of availabilities

This appendix investigates whether workers with less frequently demanded characteristics overstate their availability. Table **??** finds no evidence for a tendency by non-European workers to overstate availability and to later quit if the match does not correspond to their true preference. For Alzheimer's preferences the coefficients actually point in the opposite direction: when workers who report an availability to worker with Alzheimer's patients and they end up being matched with an Alzheimer's patient, Europeans are more likely to quit, whereas non-Europeans are not.

Table 8: Non-European workers' ex-ante availabilities and ex-post quits

	(1)	(2)	(3)	(4)	(5)	(6)
Dependent variable: Worker quits						
Alzheimer client	0.111	0.098				
	(0.056)	(0.059)				
Alzheimer client	-0.163	-0.137				
× non-European worker	(0.077)	(0.079)				
Bedridden client			-0.018	-0.026		
			(0.061)	(0.062)		
Bedridden client			0.030	0.071		
× non-European worker			(0.091)	(0.092)		
Client demands Co-residence					0.067	0.077
					(0.050)	(0.051)
Client demands Co-residence					-0.057	-0.064
× non-European worker					(0.072)	(0.074)
Year effects		X		X		X
Province effects		X		X		X
Characteristics		X		X		X
Sample: Worker available for						
	Alzheimer	Alzheimer	Bedridden	Bedridden	Co-residence	Co-residence
# Matches	1,080	1,078	1,107	1,104	1,789	1,772

Note: This table reports conditional correlations between match characteristics and the propensity for non-European workers to quit. Columns (1) and (2) include matches where workers declared availability for working with clients diagnosed with Alzheimer's. Columns (3) and (4) include matches where workers declared availability for working with bedridden clients. Columns (5) and (6) include matches where workers declared availability for co-residence. The dependent variable is an indicator for a worker quitting the job. Estimations in columns (2), (4) and (6) control for year and province fixed effects, as well as for the worker's origin, age group and sex. Source: Firm administrative data.

Table 9: Young workers' ex-ante availabilities and ex-post quits

	(1)	(2)	(3)	(4)	(5)	(6)
Dependent variable: Worker quits						
Alzheimer client	0.026	0.020				
	(0.043)	(0.045)				
Alzheimer client	-0.023	0.018				
× young worker	(0.094)	(0.095)				
Bedridden client			-0.016	-0.005		
			(0.050)	(0.052)		
Bedridden client			0.058	0.054		
× young worker			(0.115)	(0.116)		
Client demands Co-residence					0.023	0.039
					(0.040)	(0.042)
Client demands Co-residence					0.051	0.047
× young worker					(0.092)	(0.094)
Year effects		X		X		X
Province effects		X		X		X
Characteristics		X		X		X
Sample: Worker available for						
	Alzheimer	Alzheimer	Bedridden	Bedridden	Co-residence	Co-residence
# Matches	1,079	1,078	1,106	1,104	1,784	1,772

Note: This table reports conditional correlations between match characteristics and the propensity for workers aged 20-40 to quit. Columns (1) and (2) include matches where workers declared availability for working with clients diagnosed with Alzheimer's. Columns (3) and (4) include matches where workers declared availability for working with bedridden clients. Columns (5) and (6) include matches where workers declared availability for co-residence. The dependent variable is an indicator for a worker quitting the job. Estimations in columns (2), (4) and (6) control for year and province fixed effects, as well as for the worker's origin, age group and sex. Source: Firm administrative data.

Table 10: Male workers' ex-ante availabilities and ex-post quits

	(1)	(2)	(3)	(4)	(5)	(6)
Dependent variable: Worker quits						
Alzheimer client	0.005	0.011				
	(0.041)	(0.043)				
Alzheimer client	0.097	0.098				
× male worker	(0.112)	(0.113)				
Bedridden client			0.006	0.020		
			(0.048)	(0.050)		
Bedridden client			-0.103	-0.110		
× male worker			(0.138)	(0.140)		
Client demands Co-residence					0.030	0.026
					(0.040)	(0.042)
Client demands Co-residence					0.084	0.109
× male worker					(0.091)	(0.093)
Year effects		X		X		X
Province effects		X		X		X
Characteristics		X		X		X
Sample: Worker available for						
	Alzheimer	Alzheimer	Bedridden	Bedridden	Co-residence	Co-residence
# Matches	1,080	1,078	1,107	1,104	1,790	1,772

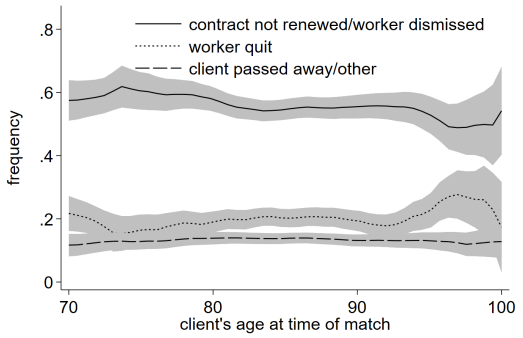
Note: This table reports conditional correlations between match characteristics and the propensity for male workers to quit. Columns (1) and (2) include matches where workers declared availability for working with clients diagnosed with Alzheimer's. Columns (3) and (4) include matches where workers declared availability for working with bedridden clients. Columns (5) and (6) include matches where workers declared availability for co-residence. The dependent variable is an indicator for a worker quitting the job. Estimations in columns (2), (4) and (6) control for year and province fixed effects, as well as for the worker's origin, age group and sex. Source: Firm administrative data.

F Reasons for match dissolution

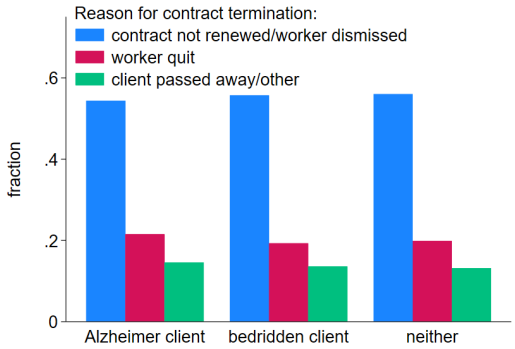
This appendix show that client characteristics are not a good predictor for matches being dissolved due to death or other reasons, which would include moving to institutional care.

Figure 18: Reasons for match dissolution

(a) By client's age



(b) By client's health



Note: This figure shows reasons for match separations by the client's age when the match is made (Panel (a)) and by whether the client has been diagnosed with Alzheimer's disease or is bedridden. Source: Firm administrative data.

G Additional Parameter Estimates and Surplus Valuations

Table 11: Other estimated parameters

Parameter	Estimate
Baseline surplus level (V_0)	16.288
Hazard factor (δ)	3.569
Idiosyncratic client tastes (σ_ε)	0.000
Idiosyncratic worker tastes (σ_η)	0.603
Wage intercept (α)	6.339
Wage standard deviation (σ_ω)	0.714

Note: This table presents additional estimated parameters from the matching model. The baseline surplus level (V_0) represents the baseline match surplus, while the hazard factor (δ) scales inverse match surpluses to the hazard rates. The parameters σ_ε and σ_η denote the spread of unobserved client and worker preferences, respectively, influencing match probabilities. The wage intercept (α) reflects the baseline level of observed wages, while σ_ω captures the variation in observed wage data relative to true transfers. Source: Firm administrative data.

Table 12: Client valuations without explicit preferences

	Young Worker	Male Worker	Non-European Worker
Client Characteristics			
Client is male	-0.106	0.561	-0.062
	(0.000)	(0.092)	(0.107)
Client demands co-residence	-0.124	-0.200	0.030
	(0.028)	(0.096)	(0.046)
Client is bedridden	-0.101	-0.414	0.000
	(0.044)	(0.102)	(0.087)
Client has Alzheimer	0.109	0.437	-0.040
	(0.054)	(0.162)	(0.054)
Baseline effect	-0.003	-0.179	0.056
	(0.079)	(0.043)	(0.043)

Note: This table presents client valuations for different match characteristics estimated from a model constraining $\boldsymbol{A}$ and $\boldsymbol{\Gamma}$ to exclude stated preferences. Asymptotic standard errors are reported in parentheses. Source: Firm administrative data.

Table 13: Worker valuations without explicit preferences

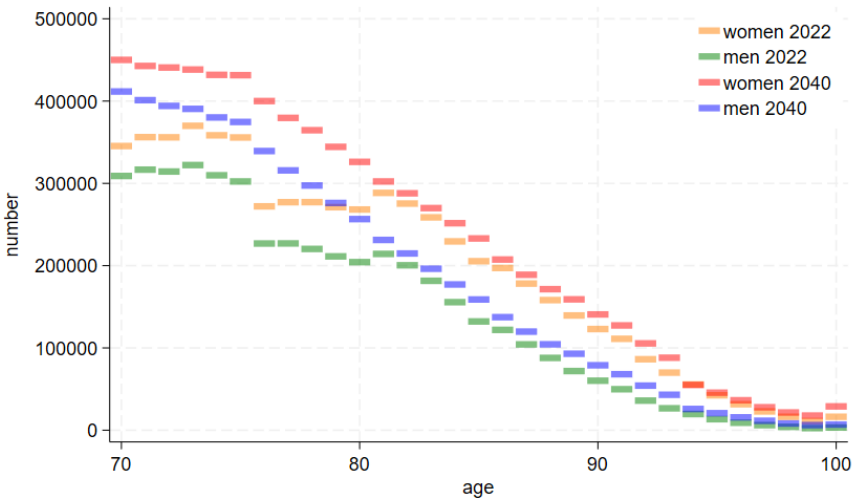
	Male Client	Co-residence	Bedridden Client	Alzheimer Client
Worker Characteristics				
Worker is young	-0.089	-0.245	0.075	-0.119
	(0.065)	(0.060)	(0.065)	(0.065)
Worker is male	1.321	0.314	0.293	-0.136
	(0.065)	(0.097)	(0.079)	(0.146)
Worker is non-European	0.019	0.271	-0.033	0.202
	(0.116)	(0.049)	(0.082)	(0.052)
Baseline effect	-0.255	1.695	-0.665	-0.005
	(0.077)	(0.065)	(0.065)	(0.071)

Note: This table presents worker valuations for different match characteristics estimated from a model constraining $\boldsymbol{A}$ and $\boldsymbol{\Gamma}$ to exclude stated preferences. Asymptotic standard errors are reported in parentheses. Source: Firm administrative data.

H Additional figures on client preferences and demographic projections

This appendix provides additional figures on the projected demographic composition of the elderly population and client preferences for caregiver attributes.

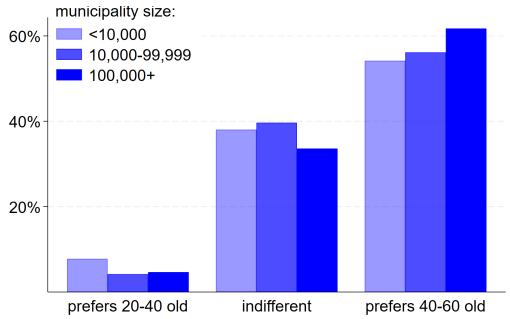
Figure 19: Projected Growth in the Elderly Population by Sex



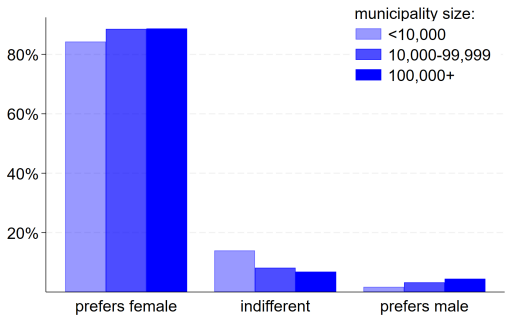
Note: This figure presents the number of individuals aged 70 and older by age and sex in 2022 and projected to 2040. The projections are based on municipality-level survival rates observed between 2015 and 2019 as explained in the text. Source: ISTAT, own elaboration.

Figure 20: Client Preferences by Municipality Size

(a) Workers' age



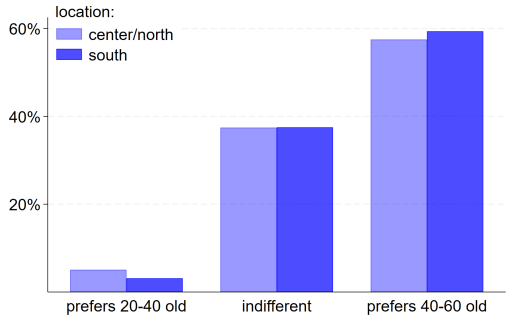
(b) Workers' sex



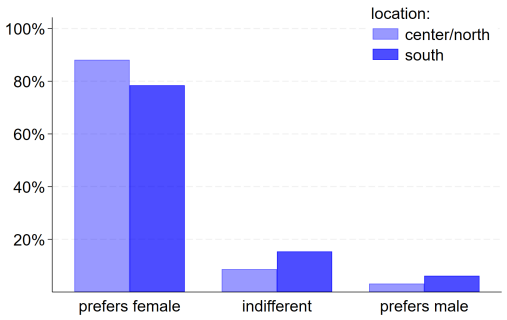
Note: The figure shows client preferences for caregiver age (Panel (a)) and sex (Panel (b)) across different municipality sizes. Preferences for caregiver origin by municipality size are presented in the main text. Source: Firm administrative data.

Figure 21: Client Preferences by Region

(a) Workers' age



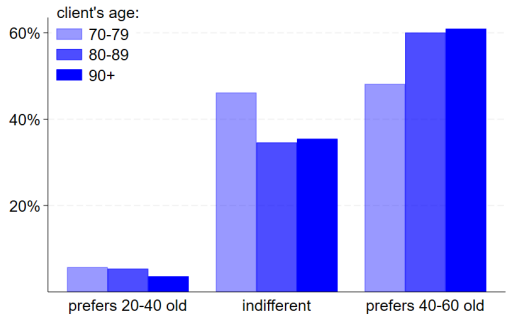
(b) Workers' sex



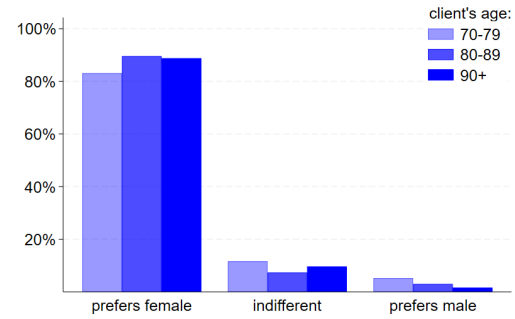
Note: The figure shows client preferences for caregiver age (Panel (a)) and sex (Panel (b)) across different regions. Preferences for caregiver origin by region are presented in the main text. Source: Firm administrative data.

Figure 22: Client Preferences by Client Age

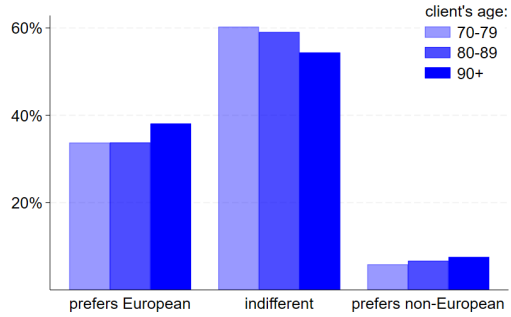
(a) Workers' age



(b) Workers' sex

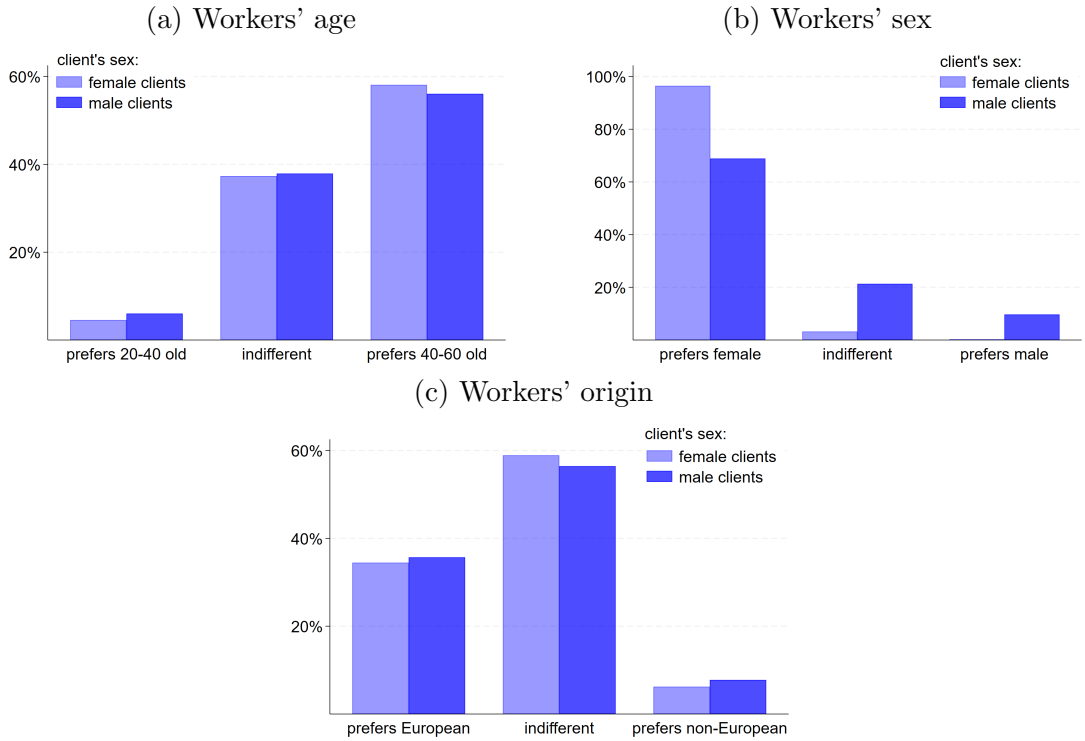


(c) Workers' origin



Note: The figure shows client preferences for caregiver age (Panel (a)), sex (Panel (b)), and origin (Panel (c)) based on the age of the client. Source: Firm administrative data.

Figure 23: Client Preferences by Client Sex



Note: The figure shows client preferences for caregiver age (Panel (a)), sex (Panel (b)), and origin (Panel (c)) by the sex of the client. Source: Firm administrative data.