UNIVERSITAT DE BARCELONA
SCHOOL OF ECONOMICS

Working Papers

Col·lecció d'Economia E22/436

LIFECYCLE CONSUMPTION AND HOUSEHOLD STRUCTURE: A PSEUDO-PANEL APPROACH

Gemma Abio

Raquel Andrés

Concepció Patxot

Guadalupe Souto

Alexandrina Stoyanova

ISSN 1136-8365

UB Economics *Working Paper No. 436*

Title: Lifecycle consumption and household structure: A pseudo-panel approach

Abstract: This paper aims to identify empirically how household structure and other socioeconomic characteristics affect the lifecycle consumption profile, exploiting data available for Spain for the last 15 years. Given the nature of the data, we build a pseudo-panel of cross-sections. After reviewing the previous literature, adjusting cohort averages to time-invariant and time-variant household characteristics, we propose employing the multilevel modelling approach WBRE (Within-Between Random Effects model). Results confirm the relevance of income as a predictor of household consumption while unveiling some interesting insights. Woman-headed households have a higher consumption profile than their man-headed counterparts after controlling for income. We observe a negative effect for number of toddlers when this characteristic is considered as a lifecycle predictor, while its impact is positive when comparisons are made between cohorts. Changing status from renter to owner has a positive within effect (lifecycle predictor), while the effect of owning on consumption is negative when contextual comparisons are made.

JEL Codes: C20, C21, C23, D12, E21, J11

Keywords: Consumption, lifecycle, age, socioeconomic status

Authors:

Gemma Abio
Universitat de Barcelona and BEAT
Email: abio@ub.edu

Raquel Andrés
Universitat de Barcelona
Email: raquel.andres@ub.edu

Ció Patxot
Universitat de Barcelona and BEAT
Email: cio.patxot@ub.edu

Guadalupe Souto
Universitat Autonoma de Barcelona
Email: guadalupe.souto@uab.es

Alexandrina Stoyanova
Universitat de Barcelona and BEAT
Email: alexandrina.stoyanova@ub.edu

Date: November 2022

Acknowledgements: The research leading to these results received funding from the Spanish Ministry of Science, Innovation and Universities under Grant Agreement PID2020-114040RB-I00; from the WELLCARE project, a project of the Horizon 2020 Joint Program Initiative More Years, Better Live's second joint transnational call, Spain (grant number PCI2021-121913); and the European Union's Horizon Europe Research and Innovation Programme under Grant Agreement number 101095175 (SUSTAINWELL project). Views and opinions expressed are however those of the author(s) only and do not necessarily reflect those of the European Union. We herewith acknowledge data provision for HBS by INE.

1. Introduction

This paper investigates empirically the factors affecting the lifecycle consumption profile, exploiting data available for Spain for the period 2006-2020. Given the nature of the data, we build a pseudo-panel of cross-sections, as first suggested by Browning et al. (1985) and developed by Deaton (1985). After reviewing the previous literature, adjusting cohort averages to time-invariant and time-variant household characteristics, we propose employing the multilevel modelling approach WBRE (Within-Between Random Effects model), also known as the *hybrid model*, introduced by Mundlak (1978), later adapted by Allison (2009) and more recently used by Twisk and Vente (2019).

Our ultimate objective is to obtain an empirical specification of lifecycle consumption that can be implemented in simulation models measuring the impact of ageing. In this respect, our approach has similarities to Fernández-Villaverde and Krueger (2007), who also aim to obtain an empirical specification of the consumption profile to be used in quantitative lifecycle simulation models.[1] Nevertheless, we aim at a different kind of simulation model, more micro-oriented, for the purpose of accounting for the impact of population ageing together with other concomitant trends, like the education transition and changes in household structure.[2] Spielauer et al. (2022) recently approached this issue developing a microsimulation model that distinguishes representative individuals by education and household type. The economic variables are incorporated in a stylized way, using the accounting logic of the National Transfer Accounts (NTA) method. This method transforms National Accounts into age-specific magnitudes, also obtaining an otherwise missing estimation of age-specific private transfers occurring within families.[3] These estimates offer a full account of how consumption needs are financed along the lifecycle, resorting to labor income at active ages, while in inactive periods agents resort to

[1] Following these authors and given our purpose, we refrain from the broad literature obtaining structural tests of the lifecycle model and focus on reduced form estimations.
[2] We also deviate from these authors in that their analysis is performed on US panel data, not on a series of cross-sections. See also Alessie and de Ree (2009), who perform a similar analysis using Dutch data.
[3] See Lee and Mason (2011) for the first NTA comparative estimates and UN (2013) for a thorough methodological description.

government transfers (welfare state), family transfers, or the asset market. As a result, the model allows the impact of welfare state transfers along the lifecycle to be taken into account, interacting with the other two resource allocation mechanisms. The results of this stylized simulation model point at important redistribution issues by gender, education level and parenthood status. This recommends a deeper development at micro level, and estimating lifecycle profiles and the socioeconomic factors affecting them is the starting point to do so.

Our analysis is related to the literature testing the lifecycle model, where two main approaches can be found.[4] The most natural approach aims to estimate the parameters behind the structural model which explicitly considers utility maximization. The complexity and limitations of this approach have led in some situations to a more exploratory, or reduced form approach to be taken, estimating cohort averages adjusted by sociodemographic characteristics to build consumption age profiles (Attanasio and Browning, 1994).[5] Given our purpose, we opt for this second approach.[6]

In both research strategies, ideally one would need longitudinal data including the whole lifecycle of individuals. These data are scarcely available. As pointed out by Attanasio and Browning (1994), testing the lifecycle theory requires micro data, while macro data were also extensively used to test some of the predictions of the model. While there is an increasing collection of longitudinal micro datasets, available panel data do not allow the full lifetime of a cohort to be covered. Moreover, panels have a rotatory rule for participating households, which can be followed only for a few years. For this reason, pseudo-panel techniques have been derived to exploit a series of cross-sections. In this paper,

[4] See Attanasio (1999) and Bullard and Feigenbaum (2007) for a survey on previous attempts to test the lifecycle hypothesis.

[5] Besides the need to control for demographic and labor participation variables, several issues arise that affect both the specific formulation of the Euler equation and the possibility of testing it with data: the level of aggregation, the separability of commodities in the utility function (also with respect to leisure) and liquidity constraints are the most fundamental. See Attanasio (1999).

[6] There have been some attempts to implement the structural approach in dynamic microsimulation models. See van de Ven (2017) for a discussion of the difficulties faced by this approach.

we choose this option due to the lack of panel data to analyze consumption in Spain.[7] In this respect, our work improves upon previous studies using different panel data techniques. Our novel contribution is to apply the WBRE model to consumption microdata. This technique allows us to consider the multilevel nature of our data, while taking care of the heterogeneity bias that commonly arises from traditional Random Effects (RE) models.

The rest of the paper is structured as follows. Section 2 presents the data and a first visual characterization of the age pattern of consumption. Section 3 estimates cohort averages adjusted by demographic, family composition and labor market characteristics. Section 4 presents the approach we adopt to estimate consumption profiles, the WBRE model, and shows the results obtained using the created pseudo-panel. The last section is dedicated to the main conclusions.

2. Data and descriptive analysis

The data used to create the pseudo-panel are obtained from the Spanish Household Budget Survey (SHBS), which each year collects the consumption expenditure and sociodemographic information from a representative sample of households that reside in Spain. We include waves since 2006, when the survey started being collected annually, through 2020, the latest wave available (INE, 2006-2020).

Each year a random sample of about 24,000 households is selected and participates for two consecutive years. This implies it is not possible to create a panel of families. In order to use the entire 15-year span of information available, we create what is known as a series of cross-sections or pseudo-panel. This methodology was first developed by Browning et al. (1985) and greatly discussed by Deaton (1985). It was later used by Attanasio and Browning (1994)

[7] In the past, there was a rotatory quarterly panel that covered the period 1985 to 2005. See Cutanda and Labeaga (2001) and Carrasco et al. (2002) for applications of the structural approach based on this data set. See also Alegre and Pou (2008), who focus on reduced form estimations.

and Blundell et al. (1994) to estimate household consumption based on household demographic and labor supply characteristics. In a nutshell, the microdata are aggregated over some defined cohorts that are identifiable in every survey's wave. In our study, we define cohorts using as time-invariant household characteristics the year of birth, sex and education of the head of the household.[8] Then, the averages of these homogeneous groups of households are followed over time. The other explanatory variables are synthesized taking the cohort average, as detailed below. Table 1 describes the birth cohorts, which are defined over a five-year period to avoid having too small cells. The two oldest and two youngest cohorts are observed during a shorter period of time due to lack of enough observations and because the SHBS lumps together any age older than 85.

Table 1
Birth cohort definition

Cohort	year of birth	observational period	age at start	age at end	average cell size
1	1931-1935	2006-2015	71-75	80-84	220
2	1936-1940	2006-2015	66-70	75-79	234
3	1941-1945	2006-2020	61-65	75-79	257
4	1946-1950	2006-2020	56-60	70-74	311
5	1951-1955	2006-2020	51-55	65-69	337
6	1956-1960	2006-2020	46-50	60-64	397
7	1961-1965	2006-2020	41-45	55-59	425
8	1966-1970	2006-2020	36-40	50-54	414
9	1971-1975	2006-2020	31-35	45-49	371
10	1976-1980	2006-2020	26-30	40-44	280
11	1981-1985	2006-2020	21-25	35-39	148
12	1986-1990	2010-2020	20-24	30-34	73
13	1991-1995	2016-2020	21-25	25-29	36

These birth cohorts are further subdivided by gender and three levels of education of the household head to produce the final average cohort dataset, which is made up of 1026 cohort-year observations. Throughout the paper, we work both with the created pseudo-panel dataset (N=1,026) and with the original

[8] Despite the fact that education is not strictly time-invariant, we expect that the variation will be negligible and we consider it time-invariant for the purposes of this study. The fact that we focus on the household head limits this variation as the education level is fixed from around the age of 20.

series of cross-sections, containing 297,698 household-year observations. We use the former in the descriptive analysis below and the regression analysis using the WBRE method (Section 4), and the latter to generate age profiles adjusted to household characteristics (Section 3).

Our main study outcome is private consumption, which is defined according to the NTA methodology as the sum of household expenditure on durables and nondurables, including private expenditure on health and education. Following Blundell et al. (2016), our consumption measure also includes imputed rents for home-owners. We divide the variables affecting consumption into the following three groups:

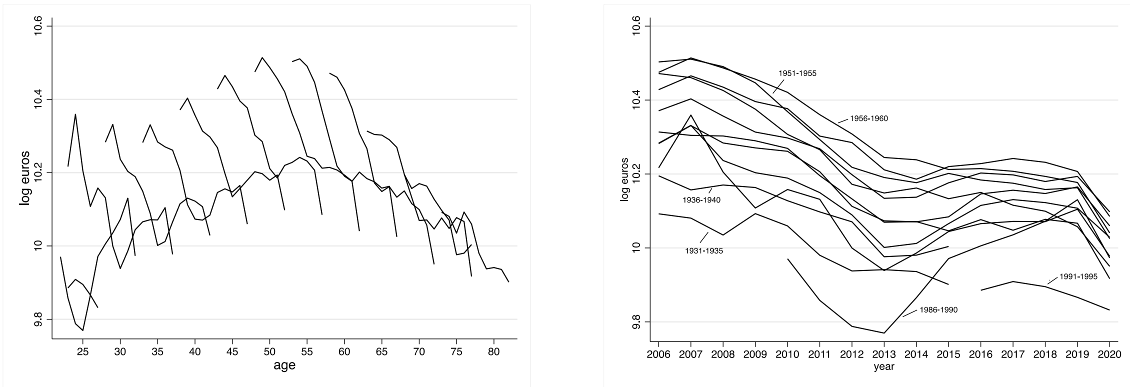
- Cohort characteristics: year of birth (defined over 5-year bands), sex and education of the household head (compulsory, secondary, university or higher). These are assumed to be invariant throughout the lifecycle.
- Time-variant household socio-demographics: total number of children, number of children aged 0 to 4, family composition, represented by categories single person, couple without children, couple with children, adult with children, other family type; ownership status, represented by categories owner with no mortgage, owner with mortgage, rent, and reduced, free or semi-free rent; log of household net earnings.[9]
- Labor participation: head of the household works, partner works.

Below we show the descriptive age profiles and how they are affected at first glance by some of the explanatory variables used later in the regression analysis. Figures 1 to 3 illustrate the average value of consumption by age and birth cohort, the household head being the reference. Each connected segment represents the lifecycle of a birth cohort, which is observed over years 2006 to 2020. For instance, the cohort of households whose head was born between 1971 and 1975 is observed at ages 31-35 to 45-46.

[9] Household earnings and expenditure variables are deflated using the 2010 Spanish aggregate price index.

Figure 1 represents the log of real private consumption for all households in our sample. The lifecycle profile for private consumption (panel a) shows the inverted U, or hump shape, widely documented empirically, as opposed to a smooth consumption profile predicted by the standard lifecycle theory. The age pattern of each longitudinal cohort profile is also affected by the business cycle. Panel (b) clearly shows the fall in consumption during the years of the economic crisis (2007-2013) and later in 2020 due to Covid-19. The resulting lifecycle age profile steadily grows until it reaches its peak at around ages 45-50 and starts to sharply descend afterwards.

Fig. 1



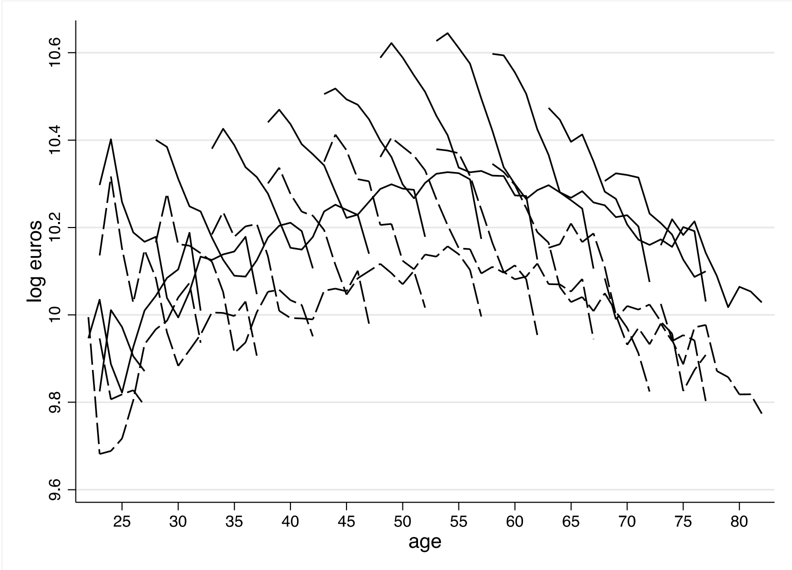
(a) (b)

Log of private consumption over the lifecycle (a) and the business cycle (b)

As expected, consumption in male-headed households is slightly above that of female-headed households (see Figure 2). In particular, this difference seems to become larger after the age of 50-55, probably reflecting cohort effects for older female cohorts participating less in the labor market.

Similarly, Figure 3 (panel a) shows that households whose male head has university or higher education exhibit more consumption over the lifecycle compared with those whose head only has compulsory studies. This gap also becomes wider after ages 50-55. Panel (b) looks at these differences for female-headed households. It is interesting to see that in this case education seems to have a bigger impact on the consumption profile than it does for male-headed households. This probably reflects the fact that we are comparing highly educated married women who earn enough to become household heads against single mothers. In fact, in 2006 35.4% of households with a low-educated female head were single-person and 33.6% were single-moms, compared with 23.7% and 20.8%, respectively, in high-educated female-headed households. On the contrary, the family composition of male-headed households is pretty similar regardless of the education of the head.

Fig. 2



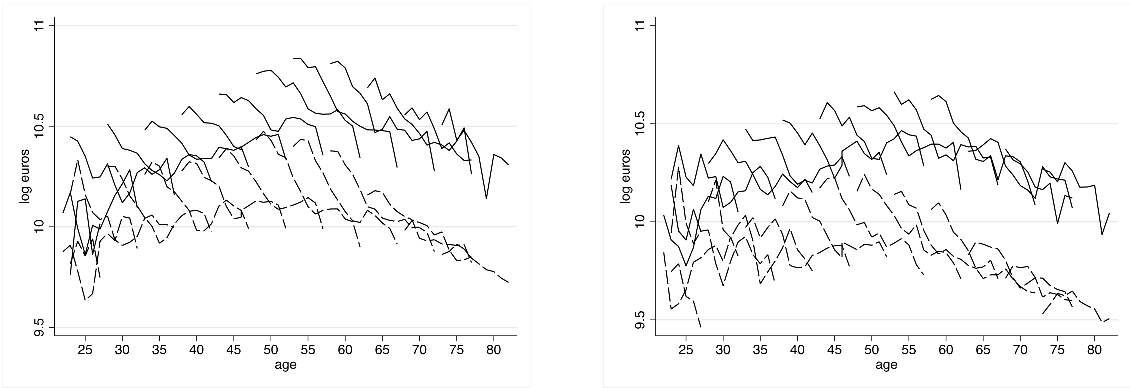
Log of private consumption for males (connected lines) vs. females (dashed lines)

One of the reasons posed in the literature to explain the hump shape of the lifecycle consumption profile is the change in household needs associated with changing family structure, including raising children which leads to an increase in consumption in middle ages. Similarly, retirement leads to a decline in consumption at older ages.[10] The path of the number of children over the lifecycle is presented in panel (a), Figure 4. We can see that the peak occurs during ages 40-45. This contrasts with Blundell et al. (1994), where the peak was seen between 35 and 40 years of age. This simply reflects the fact that, in recent decades, women in developed countries have delayed the age of having their first child. In Spain, the average age of first-time mothers has increased to 31 in 2020, up from 25 in 1980 (Eurostat, 2022).

---

[10] See Blundell et al. (1994), Attanasio and Weber (1995), Attanasio (1999), Attanasio et al. (1999), Gourinchas and Parker (2002). Interestingly, Fernández-Villaverde and Krueger (2007) find that accounting for equivalence scales explains half of the hump.

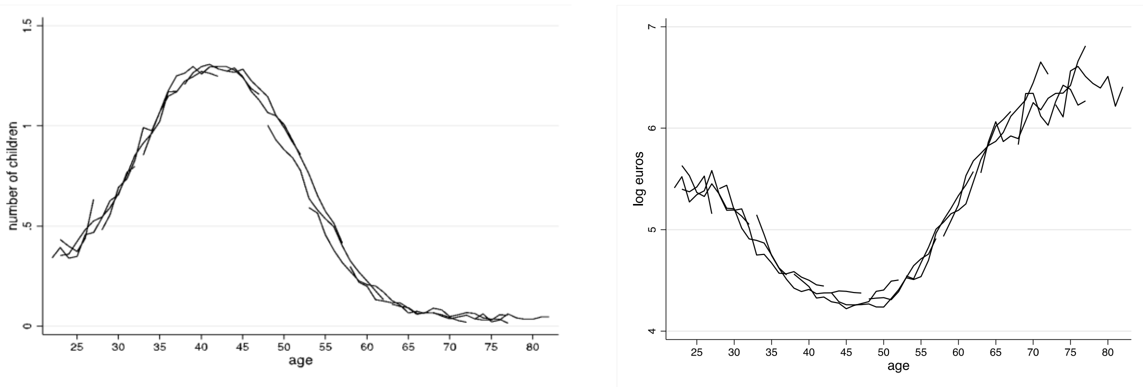
Fig. 3



(a) males (b) females

Log of private consumption for the highly educated (connected lines) vs the low educated (dashed lines) by gender

Fig. 4



(a) (b)

Number of children in the household and log of private consumption per equivalent adult over the lifecycle

Figure 4 (panel b) presents private consumption per equivalent adult[11], exhibiting a marked U-shape. Compared to Blundell et al. (1994), who found that deflating by household size totally removed the hump from the consumption profile, in our data the consumption shape is basically inverted when household consumption is divided by household size. The difference in shape is not due to the use of slightly different equivalence scales but to the difference in family compositions between Spain and the UK. In particular, our U-shape reflects larger family sizes, due to children staying at home with their parents for longer and due to having other adults co-habiting with the nuclear family, such as

[11] We used the OECD scale which assigns 1 to the first adult in the household, 0.7 to every additional adult, and 0.5 to household members younger than 14. The difference with Blundell et al. is that they assigned 1 to every adult and 0.4 to every child.

grandparents or an uncle co-residing with other generations in extended families.

3. Adjusting cohort averages

The cohort averages shown in the previous section are unadjusted in the sense that they are just cohort averages, not properly controlled for other socio-demographic characteristics of the household. The previous descriptive analysis shows the importance of household characteristics and also the need to carefully take into account the interaction among them. In this section, we revise previous attempts to estimate lifecycle consumption profiles controlling for other characteristics, leaving a thorough econometric analysis for Section 4.

Attanasio and Browning (1994) run a series of OLS regressions that allow adjusting time-variant outcomes of interest, such as consumption or income, on time-invariant characteristics, namely gender and education of the head of the household. To illustrate this method, y_t^{ci} is defined as total consumption from household *i*, belonging to cohort *c*, at time *t* and we consider the following equation

$$y_t^{ci} = \delta_t^c + \gamma' w^{ci} + u_t^{ci} \qquad [1]$$

where w^{ci} are observable household characteristics that do not change over time, also known as time-invariant, and u_t^{ci} is the error term which is assumed to be uncorrelated with w^{ci}. The parameters δ_t^c represent the average consumption of households that belong to cohort *c*, at time *t*. In practice, we estimate the delta parameters by adding birth cohort as year dummies to a regression model that also contains the observable time-invariant household characteristics, w^{ci}, which in our case are gender and education of the head. The so-called adjusted cohort averages are the delta coefficients estimated using ordinary least squares (OLS). It is worth noting that the coefficients δ_t^c in equation [1] are estimated using all the SHBS cross-sections combined, i.e. the original microdata as opposed to the pseudo-panel.

Since part of the lifecycle movements in δ_t^c will reflect changes in time-variant variables, it would make sense to further decompose δ_t^c into a part that can be explained by changes in a vector of observable time-variant characteristics z_t^{ci} and the remainder. Therefore, we can rewrite [1] as follows:

$$y_t^{ci} = \tilde{\delta}_t^c + \gamma' w^{ci} + \theta' z_t^{ci} + \tilde{u}_t^{ci} \qquad [2]$$

where z could be family composition (e.g. number of children) and labor participation variables. Now, we consider that time-variant characteristics (z) might depend on time-invariant ones and assume that we can model the z variables in the same fashion as y was previously modeled.

$$z_t^{ci} = \alpha_t^c + \beta' w^{ci} + \eta_t^{ci} \qquad [3]$$

Then, it can be proved that

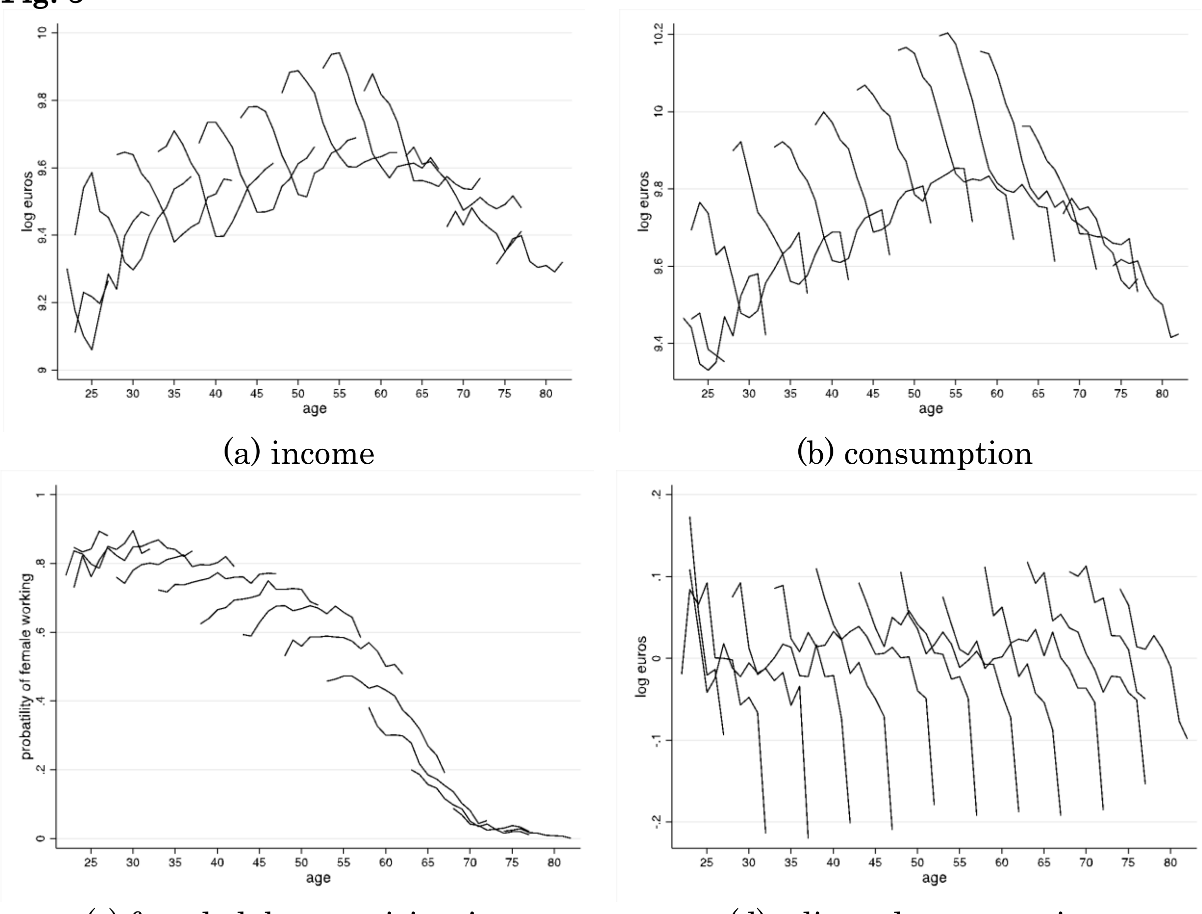
$$\delta_t^c = \tilde{\delta}_t^c + \theta' \alpha_t^c \ , \qquad [4]$$

Therefore, $\tilde{\delta}_t^c$ can be interpreted as the average level of y_t^{ci} after further removing the effect of the z characteristics. In practice, the adjusted delta's $\tilde{\delta}_t^c$ are estimated as the residuals of the regression of the estimated δ_t^c on the estimated α_t^c.

Hence, through this set of regressions, this technique further allows us to remove the effect of time-variant household characteristics, such as family composition, from consumption. This approach is described elsewhere in detail (Attanasio and Browning, 1994).

Figure 5 shows the results of estimating equations [1] to [4] by the OLS regression analysis just discussed. Panels (a) and (b) represent the cohort averages for the log of income and the log of consumption, respectively, having removed the effect of time-invariant characteristics (sex and education of the household head) using the regression in equation [1]. The plotted connected segments represent the regression coefficients corresponding to each combination of birth cohort and year. As expected, consumption varies less than income along the lifecycle just due to pure age effects, although they still follow each other quite closely.[12]

Fig. 5



(a) income
(b) consumption
(c) female labor participation
(d) adjusted consumption

(a)Log of net household earnings adjusted by sex and education of head, (b) log of household private consumption adjusted by sex and education of head, (c) female labor participation, (d) log of household private consumption adjusted by sex and education of head, number of children, number of toddlers and household earnings

[12] As pointed out by Attanasio and Browning (1995), once demographic variables are controlled for, the excess sensitivity of consumption to labor income is reduced.

Panel (c) in Figure 5 shows the results of estimating an adapted version of equation [3][13] to model dichotomous time-variant characteristics as a function of time-variant ones. Specifically, it shows the estimation of the path of female labor participation in couple households where the head of the household is the male. In Blundell et al. (1994), female employment shows a drop only during the child-bearing years. However, we cannot replicate the same pattern with the Spanish data. Our data show a steady decline after the ages of 30-35, which coincides with the current average age for having children in Spain. Unlike Blundell's, our data do not show a recovery of female employment when children are old enough. This could reflect either women's choice to stay at home or the difficulty in returning to the job market.

Finally, the age and education adjusted cohort averages for private consumption can be further regressed as shown in equation [4] to obtain the average level of consumption after removing the effect of time-variant characteristics. More specifically, panel (d) in Figure 5 shows the log of consumption from panel (b) after having removed the effect of family composition (number of children, number of children 0-4) and household income. As expected, it shows a flatter consumption profile.

4. Estimating the impact of demographic and labor market variables in consumption profiles

In this section, we aim to better capture the relevant predictors of household consumption over the lifecycle. To that end we fit a series of regression models using the WBRE model. We start by describing the method and then we present the results.

[13] Since participation in the labor market is a dichotomous variable, the regression model used is a logistic regression and the estimated intercepts or (α^c_c) are back-transformed to estimate the corresponding probabilities.

4.1. Method: Within-Between Random Effects

The nature of the data used in this study, in which we have occasions (level one) nested within individuals, households or countries (level two), requires what is known as multilevel analysis, usually applied to panel data or time series of cross-sections. More specifically, in our average cohort data, the occasions are the years or survey waves and the level-two entities are the built-in cohorts. By structure, panel data violate the classical assumption of independence of the error term, rendering OLS inefficient. Multilevel modeling improves upon OLS by allowing for data dependencies (Rice and Jones, 1997). However, there is an ongoing debate about which is the most appropriate multilevel modeling technique, fixed effects or random effects. While fixed effects (FE) have been the gold standard in economics and political science (Moffitt, 1993; Blundell and Windmeijer, 1997), random effects (RE) have been favored in other disciplines such as education, epidemiology and biomedicine (Stram and Lee, 1994; Liu et al., 2007). A clear attraction of the FE model is that it treats the individual heterogeneity as a nuisance by controlling out all higher-level sources of variability. This is achieved by including dummies (D in equation 5) for each level-2 entity (e.g. cohort) and then subtracting the higher-level entity means on both sides of the equation. Let us illustrate this with a model with a single predictor:

$$y_{it} = \sum_{i=1}^{N} \alpha_i D_i + \beta_1 x_{it} + e_{it} \qquad [5]$$

where $t = 1$ to T_i represents occasions (in turn depending on the birth cohort) and $i = 1$ to N, represents higher level or level-2 entities (cohorts)

$$y_{it} - \bar{y}_i = \beta_1 (x_{it} - \bar{x}_i) + e_{it} \qquad [6]$$

By subtracting the group means as shown in equation [6], only the within-effects are estimated and, therefore, this type of model cannot suffer from the so-called heterogeneity bias or higher-level omitted variables bias (Bafumi and Gelman, 2007). However, by trying to avoid the heterogeneity bias, the FE models suffer from a great limitation, which is that they cannot produce estimates for time-invariant variables. To put it in context with our average cohort data, this would mean not being able to obtain estimates for the effect of gender and education of the household head. Furthermore, as we just mentioned, since any between effects are removed, we are left with only the within effects of the time-varying predictors.

The RE models have a completely different approach when dealing with the multilevel nature of the data. The unexplained residual variance is split into a higher-level or level-2 variance (a.k.a. between-subjects) and a lower-level or level-1 variance (a.k.a. within-subjects or occasions). The two-stage formulation for a *random intercept model*[14] is written as:

$$y_{it} = \beta_{0i} + \beta_1 x_{it} + e_{it} \tag{7}$$

$$\beta_{0i} = \beta_0 + \varsigma_{0i} \tag{8}$$

where [7] is the within-subject model and [8] the between-subject model. Replacing the latter with the former we obtain the *reduced form* formulation:

$$y_{it} = \beta_0 + \beta_1 x_{it} + \varsigma_{0i} + e_{it} \tag{9}$$

where ς_{0i} is subject *i*'s random intercept or subject *i*'s specific contribution to the residual variance, also known as random effects, and e_{it} represents the occasion-level residual for occasion t of subject i. The error terms ς_{0i} and e_{it} are sometimes referred to as permanent and transitory components as ς_{0i} represents the time-invariant characteristics of the individuals and e_{it} the moment's random deviation. Both error terms are assumed to follow a normal distribution, although RE models are shown to be quite robust to violations of the normality

[14] Note that this formulation can be expanded to the random coefficient model by adding a random slope to equation [7].

assumption (Beck and Katz, 2007). The RE models are built upon the assumption that the residual terms, both ς_{0i} and e_{it}, are independent from the covariates x_{it} (exogeneity assumption). This is a very strong assumption that is rarely met and is a reason why many researchers prefer to work with FE models. Endogeneity can be better addressed if we realize that the time-varying covariates in the model can be divided into two parts: one that only varies between subjects or higher-level entities, and one that represents between occasions or within subjects' variation:

$$x_{it} = x_i^B + x_{it}^W \qquad [10]$$

Therefore, a given time-varying covariate will have a total effect that is the addition of the between and the within effects. The limiting assumption that RE models make is that in equation [9] the within and between effects are equal. When this assumption does not hold, which happens more often than not, we have an omitted variable bias, coined heterogeneity bias by Li (2011), which will lead to correlation between the covariate and the error terms. A solution to deal with the problem of heterogeneity bias consists of (1) adding in the model the group mean corresponding to each time-varying covariate, $\bar{x}_i$, which accounts for the between effect and (2) group mean center the time-varying covariates, which breaks the correlation between them and their group mean, $\bar{x}_i$, and the level-2 error term, ς_{0i}. After these two steps, equation [9] is transformed into:

$$y_{it} = \beta_0 + \beta_1(x_{it} - \bar{x}_i) + \beta_2\bar{x}_i + \varsigma_{0i} + e_{it} \qquad [11]$$

where β_1 estimates the within effect and β_2 the between effect of variable x_{it}. This reformulation of the RE model is known as the WBRE model (Mundlak, 1978; Allison, 2009).

The well-known Hausman specification test (Hausman, 1978) is commonly used to choose between fixed and random effects. In fact, what the Hausman test does is to compare the estimates of the FE and the RE models. In other words, it tests whether the RE assumption that both within and between effects are equal holds (Bell and Jones, 2015). If it does not hold, then according to what we have argued so far, the WBRE model should be used unless we are only interested in the time-varying variables and their within effects, in which case the FE model should be good enough.

Finally, another advantage of using the WBRE model, in the context of estimating lifecycle consumption, is that it allows us to separate the lifecycle effect (the within effect) from the contextual effect (the between minus the within effect).

4.2 Results of the WBRE regression analysis

Models estimated with aggregated data tend to have higher correlations between variables that would not necessarily be found if working with the original microdata. This phenomenon, known as the ecological fallacy, was first discovered by Robinson (1950). High correlations can lead to multicollinearity that can affect the standard errors and even the signs of the coefficients of the variables affected by such multicollinearity. Therefore, in order to select the model covariates, an exploration of their correlations was performed. We considered any Pearson correlation between predictors higher than 0.7 to be of concern (Yu et al., 2015). The correlations between level-1 variables and cross-level correlations are found in Table 4 in the appendix. Following this correlation analysis, it was decided to collapse the categories *couple with children* and *couple without children* into *couple* to avoid the high correlations found between *couple with children* and *number of children* and between *couple with children* and *head works*. Likewise, we found strong correlations between *owner with mortgage* and *number of children*. As a result, we collapsed both categories *owner with mortgage* and *owner with no mortgage* into *owner*. *Head works* and *partner works* have correlations above 0.7 with *number of children* and therefore, if the labor participation variables are added into a regression model, it will be in a separate block to isolate their potential influence on the estimated coefficient for *number of children*. It was also decided not to add *education* and *household net earnings* in the same model due to the high correlation observed between the two variables. Note that our models do not include time to avoid perfect collinearity with cohort and age arising from the equality *time = cohort + age*. Finally, the results reported in this section are unweighted. We ran the same models using sampling weights and obtained very similar results. The regression coefficients from the weighted models are found in Tables 5 and 6 in the Appendix.

In the first set of equations we use education as regressor. The regressions are performed without interactions (model 1), interacting age and sex (model 2) and age and education (model 3). Results from the first set of models, which do not include labor participation variables or household income, are presented in Table 2. As expected, the number of children shows strong and positive within and between effects in models 1 and 2. As could be expected, the between effect is stronger than the within effect, due to household economies of scale. After adding age by sex interactions, in model 2, the within effect stays basically the same and the between effect becomes even stronger. However, when age by education interactions are included, in model 3, only the within effect remains statistically significant. In other words, after the interaction between age and education is taken into account, number of children has a positive effect only within the cohort, as a "lifecycle" predictor and no longer as a "contextual" predictor. This indicates that education and age explain most of the differences in consumption across cohorts.

Number of toddlers shows a negative and significant within effect while having a positive and much stronger between effect (model 1), suggesting that while household consumption decreases during the periods in which a household has an additional child between 0 and 4, on average, having an additional toddler increases household's consumption when compared to other cohorts. The former is probably due to time constraints associated with raising toddlers. In models 2 and 3, only the negative within effect remains significant at a 5% level. As before, the interaction between age and education explains most of the between effect.

Being married or with a partner (couple) has a consistent positive effect on household consumption with both effects significant, the between effect being stronger than the within effect. This means that the differences in consumption between couple households and one-adult households averaging across time are larger than lifecycle differences within the same household due to, for instance, marriage, divorce or widowing.

Table 2
Household Private Consumption's Within and Between Effects

	(1) Within	Between	(2) Within	Between	(3) Within	Between	(4) OLS with birth cohort dummies
Female		-0.0384*		-0.0878		-0.0625	-0.0897***
		(0.0216)		(1.211)		(0.899)	(0.0230)
Secondary		0.235***		0.224***		0.133	0.138***
		(0.0119)		(0.0114)		(1.103)	(0.0289)
University		0.456***		0.427***		0.139	0.234***
		(0.0158)		(0.0162)		(1.134)	(0.0324)
num_children	0.0868***	0.115***	0.0805***	0.146***	0.104***	-0.224	0.0492
	(0.0274)	(0.0384)	(0.0292)	(0.0448)	(0.0342)	(0.132)	(0.0302)
num_toddlers	-0.121**	0.475***	-0.116**	0.280*	-0.193***	0.0381	-0.146**
	(0.0522)	(0.119)	(0.0558)	(0.141)	(0.0658)	(0.290)	(0.0610)
couples (%)	0.119**	0.358***	0.128**	0.963***	0.112*	0.416**	0.244***
	(0.0603)	(0.0474)	(0.0619)	(0.153)	(0.0637)	(0.196)	(0.0488)
owners (%)	0.223***	-0.858***	0.226***	-1.120***	0.210***	0.0374	0.211***
	(0.0676)	(0.178)	(0.0683)	(0.188)	(0.0707)	(0.370)	(0.0651)
Constant		8.002***		7.734***		8.698***	10.67***
		(0.670)		(0.869)		(0.957)	(0.0628)
Observations	1,026	1,026	1,026	1,026	1,026	1,026	1,026
R-squared	0.555	0.977	0.558	0.985	0.570	0.996	0.897
F age	79.53***	17.10***	78.73***	17.35***	78.75***	17.85***	79.96***
F ageXsex			0.420	1.92*	0.39	1.05	0.52
F ageXedu					1.10	2.65**	7.85***
F birth cohort							52.67***

Standard errors in parentheses. num_children: number of dependent children; num_toddlers: number of children 0 to 4 years of age. l_earnings: log of household net earnings. Couples and owners are measured as a share of households in the cohort. Calculations are unweighted.
*** p<0.01, ** p<0.05, * p<0.1

For owner, the estimated within effects are positive and significant while the between effects are negative and significant and of a greater magnitude in models 1 and 2. The discrepancy in signs can be interpreted as follows: an increase in consumption when a family changes status from renting to owning makes sense as a family may decide to own when their household income has increased enough.[15] On the other hand, the negative between effect for owner means that on average families that own their dwelling consume less than those that rent. This is consistent with the fact that there is a positive association between older ages and owning (about 90% of households whose head is older than 70 own) and consumption is lower among the older cohorts. After adjusting for the interaction between age and education (model 3), the between effects disappear and only the within effects are significant.

The F associated with the age by education interaction is statistically significant only for the between effects formulation, which means that while there is no significant interaction within the household there is a significant one when comparing different households which differ by the education and age of the head. No interaction is found between age and sex either within or between.

Regarding the effect of time-invariant characteristics sex and education, obviously only the between effects are estimated. Female household head is only marginally significant in model 1 and is not significant in models 2 and 3, showing that education explains most of the differences. The education dummies secondary and university have significant and positive coefficients in models 1 and 2, with university having the largest effect, as expected. The coefficients are no longer significant after adding the interaction age by education, as the interaction absorbs this effect. However, it is important to note that, after adding the interaction term, the interpretation of the education dummies is no longer as a main effect but as the effect of education for ages 22 to 25 (reference category). Finally, the OLS model estimates shown as a benchmark give a negative effect for female and positive and strong effects for secondary and university.

Table 3 presents the results from the second set of models. In this case the models differ by including labor participation variables and/or household net earnings as additional predictors of household consumption, instead of education. Model 2 omits earnings and hence loses explanatory power, while

[15] This is confirmed in Table 3 where earnings are an explanatory variable.

offering some interesting insights. The results on the number of children and toddlers are similar to those in Table 2, except for the fact that the negative and smaller within effect for toddlers is not significant any more.

Table 3
Household Private Consumption's Within-Between Effects with Earnings and Labor Participation

	(1)		(2)		(3)		(4)
	Within	Between	Within	Between	Within	Between	OLS with birth cohort dummies
Female		0.0520***		0.0616		0.0453***	0.0223**
		(0.0117)		(0.0626)		(0.0152)	(0.0101)
num_children	0.0674***	0.0976***	0.0110	0.0217	0.0403**	0.0957***	0.0869***
	(0.0201)	(0.0209)	(0.0241)	(0.0858)	(0.0197)	(0.0209)	(0.0124)
num_toddlers	-0.0386	0.191***	-0.0406	0.974***	-0.0125	0.192**	0.0125
	(0.0385)	(0.0622)	(0.0455)	(0.300)	(0.0372)	(0.0770)	(0.0308)
couples (%)	0.0364	0.136***	0.0788	0.577***	0.0233	0.136***	0.0951***
	(0.0444)	(0.0263)	(0.0525)	(0.121)	(0.0430)	(0.0326)	(0.0210)
owner (%)	-0.147***	-0.931***	0.0857	-1.751***	-0.141***	-0.874***	-0.434***
	(0.0514)	(0.0924)	(0.0605)	(0.490)	(0.0505)	(0.122)	(0.0440)
l_earnings	0.651***	0.818***			0.558***	0.805***	0.718***
	(0.0230)	(0.0149)			(0.0260)	(0.0264)	(0.0133)
Head works (%)			0.743***	2.719***	0.306***	0.00798	0.121***
			(0.0428)	(0.289)	(0.0405)	(0.113)	(0.0367)
Partner works (%)			-0.0482	0.832***	0.0228	0.0737	0.0829***
			(0.0393)	(0.267)	(0.0323)	(0.0692)	(0.0290)
Constant		1.229***		4.308***		1.447***	3.845***
		(0.397)		(1.404)		(0.353)	(0.111)
Observations	1,026	1,026	1,017	1,017	1,017	1,017	1,017
R-squared	0.761	0.993	0.669	0.884	0.780	0.993	0.939
F age	56.98***	12.58***	37.18***	8.020***	34.48***	12.62***	26.83***
F birth cohort							64.42***

Standard errors in parentheses. num_children: number of dependent children; num_toddlers: number of children 0 to 4 years of age. Head/partner works indicates head/partner participates in the labor market; l_earnings: log of household net earnings. Calculations are unweighted. *** p<0.01, ** p<0.05, * p<0.1

Among some of the interesting aspects to highlight, we find that female being the household head has a positive and significant sign where earnings are present. This is true for models 1 and 3, but not for model 2. This result indicates that for households with similar income and family composition, those where the head is a woman tend to consume more. This contrasts with the models in Table 2 where female had negative or no effect, and is probably because consumption tracks income closer than education level, used as regressor in Table 2. Hence, we identify relevant effects of women being household head, contrary to previous papers focusing on only male-headed households.

Another result worth mentioning is that the positive within effects for owner reported in Table 2 become negative in the presence of earnings, which is interpreted as meaning that, keeping earnings the same, changing status from renting to owning has an associated decrease in consumption. On the other hand, the negative between effect seen in the models in Table 2 is maintained, confirming our interpretation of Table 2 results. In order to better understand the negative within coefficients for owner we rerun the same models excluding housing expenditure from household consumption. Results, reported in Table 7 in the Appendix, show that once rent and imputed rents are removed, the within effect for those that own stops being significant when we control for earnings (models 1 & 3), while it becomes positive and significant when earnings are not included (model 2).

If we look at couple, the effect is positive as in Table 2, but only the between effect is significant in all three models, with a much larger effect for model 2 which only contains labor participation variables.

Head works has both positive and significant within and between effects when household earnings are not present and only significant within effects when earnings are included in the model. In other words, head works as a lifecycle predictor has a positive effect on consumption, while, when comparing households with similar income, their consumption levels are not different whether or not the head is active. Partner works has only positive and significant between effects when earnings are not included in the model (model 2).

5. Conclusions

In this paper we employ pseudo-panel techniques to exploit a series of cross-sections of the Spanish Household Budget Survey from 2006 to 2020. Our aim is to obtain a reduced form estimation of the lifecycle consumption profile, capturing the impact of the main socioeconomic characteristics –mainly education and family structure– which could be used to project future consumption in micro-based simulation models investigating the impact of ageing.

We start by replicating some previous attempts to adjust cohort averages to time-invariant and time-variant characteristics using OLS estimations. We go a step further in the estimation of the effects of those characteristics by employing the WBRE method, which allows us to identify the impact of time-invariant household characteristics and to separate the within from the between effects for time-variant characteristics.

As could be expected, education and income cannot be used as regressors at the same time due to the strong correlation between them. In the first set of regressions (Table 2) education is included, and in successive models the interaction between age and gender and age and education are added, the latter being more significant. In the second set of regressions (Table 3) we perform three specifications: adding income, labor participation or both. It is observed that when education is replaced with income the explanatory power of the within model substantially increases.

Women being a household head has a negative or no effect in the first set of regressions, when education is used as regressor. Interestingly, when income is used as regressor in the second set of regressions, the effect of women being a household head becomes positive showing that, in households with similar income and family composition, those where women are the head consume more. This is probably because women working externalize home production to a greater extent. Education shows the expected positive effect when included in the first set of regressions, the effect being absorbed when age and education interactions are added.

Regarding time-variant characteristics, when education is included in the model, we identify a positive effect of number of children on household consumption which turns out to be bigger between cohorts than as a lifecycle predictor. Interestingly, the number of toddlers has positive between effects while lifecycle effects are negative, probably reflecting time constraints affecting consumption of raising small kids. Both for children in general and for toddlers, the between effect is no longer significant once we take into account the interaction between age and education of the household head. We also identify a positive impact of living as a couple on household consumption, which is always stronger between cohorts than over the lifecycle. The effects of number of children, toddlers and cohabitation remain similar when income is an explanatory variable instead of education, although there are some differences. The within effects of toddlers and partnership status are no longer significant when we control for income and/or labor participation.

When estimating the impact of being an owner versus a renter, one would expect a positive impact on consumption, probably reflecting an increase in income. This is well captured by the positive within effect in models not controlled by income. On the other hand, we obtain consistent negative between effects for ownership, possibly reflecting the large concentration of owners among the older cohorts, these cohorts having a low consumption profile.

Lastly, the impact of household head labor participation tends to be positive, while the labor status of the partner has only significant positive between effects when only labor status is included (income absorbs this effect).

Our results illustrate the importance of using the WBRE estimator, particularly when the within and between effects go in opposite directions, as is the case for number of toddlers and ownership status. The use of the traditional random effects estimator would not allow us to distinguish the lifecycle from the contextual effect, which makes these results much richer.

Finally, our results are limited by the nature of the data. Spain has not collected panel data on household consumption since 2005 when the rotatory quarterly panel was transformed into a yearly cross-sectional survey. Even before 2005, individuals were only followed for 8 quarters, rendering data inadequate for a lifecycle analysis. A pseudo-panel was created in order to empirically analyze household consumption over the lifecycle. This synthetic panel involves a loss of information as the micro data are aggregated at the cohort level. However, it provides the advantage of applying panel data techniques to estimate lifecycle profiles, which would not be possible working with the original cross-sections.

References

Alegre Martín, J. & Pou Garcias, Ll. (2008). El consumo y la tasa de ahorro privados de los hogares españoles: una descomposición de los efectos edad y cohorte. *Investigaciones Económicas*, *32*(1), 87-121.

Alessie, R. & de Ree, J. Explaining The *Hump* In Lifecycle Consumption profiles. *De Economist* **157**, 107–120 (2009). https://doi.org/10.1007/s10645-009-9119-4

Allison, P. D. (2009). *Fixed effects regression models*. SAGE publications.

Attanasio, O. P. (1999). Consumption. *Handbook of macroeconomics*, *1*, 741-812.

Attanasio, O.P. & Browning, M. (1994). Testing the Lifecycle Model of Consumption: What can we Learn from Micro and Macro Data?. *Investigaciones Económicas*, 18 (3), 433-463.

Attanasio, O. P., & Browning, M. (1995). Consumption over the Life Cycle and over the Business Cycle. *The American Economic Review*, *85*(5), 1118–1137. http://www.jstor.org/stable/2950978

Attanasio, O. P., Banks, J., Meghir, C., & Weber, G. (1999). Humps and bumps in lifetime consumption. *Journal of Business & Economic Statistics*, *17*(1), 22-35.

Bafumi, J. & Gelman, A. (2007). Fitting Multilevel Models When Predictors and Group Effects Correlate. https://ssrn.com/abstract=1010095 or http://dx.doi.org/10.2139/ssrn.1010095

Beck, N. & Katz, J.N. (2007). Random Coefficient Models for Time-Series-Cross-Section Data: Monte Carlo Experiments, *Political Analysis*, 15(2), 182–95.

Bell, A. & Jones, K. (2015). Explaining Fixed Effects: Random Effects Modeling of Time-Series Cross-Sectional and Panel Data, *Political Science Research and Methods*, 3 (1), 133-153.

Blundell, R., Browning, M., & Meghir, C. (1994). Consumer Demand and the Lifecycle Allocation of Household Expenditures, *Review of Economics Studies*, 61, 57-80.

Blundell, R. & Windmeijer, F. (1997). Cluster effects and simultaneity in multilevel models. *Health Econ*, 6, 439-443.

Blundell, R. Pistaferri, L., & Saporta-Eksten, I. (2016). Consumption Inequality and Family Labor Supply, *The American Economic Review*, 106(2), 387-435.

Browning, M., Deaton, A., & Irish, M. (1985). A profitable Approach to Labor Supply and Commodity Demands over the Lifecycle, *Econometrica* 53, 921-951.

Browning, M. & Meguir, C. (1991). The effects of male and female labor supply on commodity demands, Econometrica, 59 (4), 925-951.

Bullard, J., & Feigenbaum, J. (2007). A leisurely reading of the lifecycle consumption data. *Journal of Monetary Economics*, *54*(8), 2305-2320.

Carrasco, R., Labeaga, J. M., & David López-Salido, J. (2005). Consumption and habits: evidence from panel data. *The Economic Journal*, *115*(500), 144-165.

Cutanda, A., & LABEAGA, J. M. (2001). Simulación de perfiles de consumo a partir de un pseudo-panel de la ECPF. *Revista de Economía Aplicada*, *9*(26), 95-123.

Deaton, AS. (1985). Panel Data from Time Series of Cross Sections, *Journal of Econometrics*, 30, 109-126.

Eurostat (2022), *Population Statistics at regional level*. (online): European Comission. Retrieved August 08, 2022 from https://ec.europa.eu/eurostat/statistics-explained/index.php?title=Population_statistics_at_regional_level#Fertility

Fernández-Villaverde, J., & Krueger, D. (2007). Consumption over the lifecycle: Facts from consumer expenditure survey data. *The Review of Economics and Statistics*, *89*(3), 552-565.

Gourinchas, P. O., & Parker, J. A. (2002). Consumption over the lifecycle. *Econometrica*, *70*(1), 47-89.

Hausman, J.A. (1978). Specification Tests in Econometrics, *Econometrica*, 46(6),1251-71.

INE (2006-2020). Encuesta de Presupuestos Familiares (EPF). Madrid: Instituto Nacional de Estadística. Recovered from https://www.ine.es/dyngs/INEbase/es/operacion.htm?c=Estadistica_C&cid=1254736176806&menu=resultados&secc=1254736195147&idp=1254735976608#!tabs-1254736195147

Lee, R. D., & Mason, A. (Eds.). (2011). *Population aging and the generational economy: A global perspective*. Edward Elgar Publishing.

Li, X. (2011). *Approaches to modelling heterogeneity in longitudinal studies* (Doctoral dissertation, Open Access Te Herenga Waka-Victoria University of Wellington). http://researcharchive.vuw.ac.nz/bitstream/handle/10063/1695/thesis.pdf?sequence=1

Liu, L., Wolfe, R.A., & Kalbfleisch, J.D. (2007). A shared random effects model for censored medical cost and mortality, *Statistics in Medicine*, 26 (1), 139-155.

Moffit, R. (1993), "Identification and Estimation of Dynamic Models with Time Series of
Repeated Cross Sections", *Journal of Econometrics*, 59, 99-123.

Mundlak, Y. (1978). "On the pooling of time series and cross section data", *Econometrica* 46, 69-85.

Rice, N. & Jones, A. (1997). Multilevel models and health economics, *Health economics* 6, 561-575.

Robinson, W.S. (1950). Ecological correlations and the behaviour of individuals, *American Sociological Review*, 15, 351–57.

Spielauer, M., Horvath, T., Fink, M., Abio, G., Souto, G., Patxot, C., & Istenič, T. (2022). Measuring the lifecycle impact of welfare state policies in the face of ageing. *Economic Analysis and Policy*, *75*, 1-25.

Stram, D.O. & Lee, J.W. (1994). Variance Components Testing in the Longitudinal Mixed Effects Model, *Biometrics*, 50 (4), 1171-1177.

Twisk W.R. & Vente W. (2019). Hybrid models were found to be very elegant to disentangle longitudinal within- and between-subject relationships, *Journal of Clinical Epidemiology*, 107, 66-70.

United Nations (2013). *National transfer accounts manual: Measuring and analysing the generational economy*. UN.

van de Ven, J. (2017). SIDD: An adaptable framework for analysing the distributional implications of policy alternatives where savings and employment decisions matter, *Economic Modelling*, 63, 161-174.

Yu, H., Jiang, S. & Land, K.C. (2015). Multicollinearity in hierarchical linear models, *Social Science Research*, 53, 118-136.

Appendix A. Correlations between study variables.

Table 4
Level -1 and cross-level pairwise correlations. 2006-2020 pseudo-panel

Variable	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17
1. log of consumption																	
2. education	.65***																
3. log of earnings	.93***	.76***															
4. number of children	.15***	-.08**	.08**														
5. num. children 0-4	-.08**	-.05*	-.08**	.58***													
6. single person	-.36***	.09***	-.31***	-.54***	-.32***												
7. couple	.45***	.11***	.48***	.37***	.34***	-.83***											
8. couple no children	.11***	.16***	.21***	-.52***	-.14***	-.28***	.48***										
9. couple with children	.41***	-.00	.37***	.82***	.48***	-.70***	.73***	-.25***									
10. one adult with children	-.30***	-.21***	-.41***	.08**	-.19***	.25***	-.71***	-.54***	-.37***								
11. other family	-.29***	-.36***	-.40***	-.22***	-.16***	.16***	-.46***	-.16***	-.38***	.34***							
12. owner	.40***	.12***	.44***	-.22***	-.51***	.07**	.05	.05	.02	-.10***	-.26***						
13. owner no mortgage	.08**	-.02	.13***	-.67***	-.78***	.35***	-.23***	.26***	-.45***	-.04	.01	.78***					
14. owner with mortgage	.28***	.15***	.26***	.81***	.70***	-.49***	.41***	-.35***	.73***	-.04	-.29***	-.19***	-.76***				
15. rent	-.35***	-.09***	-.40***	.24***	.52***	-.09***	-.03	-.03	-.01	.10***	.26***	-.98***	-.79***	.22***			
16. reduced or free rent	-.44***	-.20***	-.47***	.13***	.35***	.03	-.10***	-.08***	-.05*	.09***	.22***	-.81***	-.57***	.05***	.68***		
17. head works	.28***	.16***	.22***	.73***	.55***	-.48***	.32***	-.39***	.65***	.06*	-.07**	-.48***	.83***	.81***	.49***	.30***	
18. partner works	.06*	.08**	-.00	.70***	.57***	-.41***	.18***	-.39***	.50***	.21***	.01	-.62***	-.89***	.76***	.63***	.43***	.93***

*** p<0.01, ** p<0.05, * p<0.1

Appendix B. Regression coefficients using sampling weights

In this section, we present the same regression models reported in Section 4 corrected using sampling weights.

Table 5
Household Private Consumption's Within and Between Effects

	(1)		(2)		(3)		(4)
	Within	Between	Within	Between	Within	Between	OLS with birth cohort dummies
female		-0.0351		-0.0693		-0.377	-0.108**
		(0.0228)		(0.644)		(0.730)	(0.0515)
secondary		0.234***		0.223***		0.113	0.182***
		(0.00904)		(0.00903)		(1.020)	(0.0645)
university		0.455***		0.425***		0.678	0.303***
		(0.0152)		(0.0141)		(1.001)	(0.0637)
num_children	0.120***	0.116***	0.117***	0.144***	0.136***	0.00368	0.0841**
	(0.0271)	(0.0271)	(0.0285)	(0.0233)	(0.0354)	(0.148)	(0.0328)
num_toddlers	-0.0440	0.475***	-0.0363	0.273**	-0.0545	-0.0191	-0.0369
	(0.0544)	(0.110)	(0.0573)	(0.122)	(0.0661)	(0.289)	(0.0631)
couple	0.146*	0.365***	0.167*	1.018***	0.150*	0.705***	0.382***
	(0.0844)	(0.0428)	(0.0865)	(0.150)	(0.0814)	(0.190)	(0.0570)
owner	0.167*	-0.861***	0.170*	-1.162***	0.160*	-0.135	0.154**
	(0.0910)	(0.130)	(0.0943)	(0.175)	(0.0889)	(0.317)	(0.0754)
Constant		8.005***		7.697***		8.434***	10.47***
		(0.517)		(0.541)		(0.873)	(0.0859)
Observations	1,026	1,026	1,026	1,026	1,026	1,026	1,026
R-squared							0.929
F age	145.7***	30.36***	449.4***	48.85***	178.2***	228.6***	60.91***
F ageXsex			0.68	3.22***	0.79	5.64***	1.530
F ageXedu					3.91***	15.57***	8.590***
F birth cohort							68.90***

Standard errors in parentheses. num_children: number of dependent children; num_toddlers: number of children 0 to 4 years of age. *** p<0.01, ** p<0.05, * p<0.1

Table 6
Household Private Consumption's Within-Between Effects with Earnings and Labor Participation

	(1)		(2)		(3)		(4)
	Within	Between	Within	Between	Within	Between	OLS with birth cohort dummies
female		0.0505***		0.102**		0.0452***	0.0308***
		(0.0127)		(0.0444)		(0.0161)	(0.0104)
num_children	0.0792***	0.0973***	0.0193	-0.0123	0.0482***	0.0967***	0.0861***
	(0.0206)	(0.0141)	(0.0198)	(0.0445)	(0.0186)	(0.0146)	(0.00954)
num_toddlers	0.0198	0.188***	0.0220	1.012***	0.0377	0.245***	0.0605**
	(0.0345)	(0.0564)	(0.0418)	(0.218)	(0.0311)	(0.0724)	(0.0256)
couple	0.0916	0.135***	0.166**	0.679***	0.104*	0.139***	0.133***
	(0.0592)	(0.0257)	(0.0715)	(0.106)	(0.0584)	(0.0314)	(0.0206)
owner	-0.358***	-0.933***	0.0215	-1.948***	-0.318***	-1.018***	-0.691***
	(0.0742)	(0.0703)	(0.0790)	(0.436)	(0.0700)	(0.0952)	(0.0434)
l_earnings	0.697***	0.819***			0.572***	0.798***	0.732***
	(0.0235)	(0.0114)			(0.0312)	(0.0249)	(0.0116)
head works			0.797***	2.815***	0.314***	0.105	0.143***
			(0.0578)	(0.263)	(0.0424)	(0.0988)	(0.0267)
partner works			-0.0471	0.641***	0.0332	0.0616	0.0377
			(0.0438)	(0.158)	(0.0382)	(0.0776)	(0.0241)
Constant		1.224***		4.751***		1.184***	3.840***
		(0.328)		(1.205)		(0.336)	(0.0980)
Observations	1,026	1,026	1,017	1,017	1,017	1,017	1,017
R-squared							0.969
F age	119.4***	30.68***	55.74***	16.09***	96.05***	24.04***	32.41***
F birth cohort							71.48***

Standard errors in parentheses. num_children: number of dependent children; num_toddlers: number of children 0 to 4 years of age; head/partner works indicates head/partner participates in the labor market; l_earnings: log of household net earnings. *** p<0.01, ** p<0.05, * p<0.1

Appendix C.

In this section, we present the same regression models reported in Section 4 excluding housing expenditure from household consumption.

Table 7
Household Private Consumption's Within-Between Effects, excluding housing from consumption

	(1)		(2)		(3)		(4)
	Within	Between	Within	Between	Within	Between	OLS with birth cohort dummies
female		0.0348**		0.0362		0.0188	0.0104
		(0.0166)		(0.0679)		(0.0210)	(0.0126)
num_children	0.0921***	0.0785**	0.0279	-0.00109	0.0622**	0.0777***	0.0782***
	(0.0258)	(0.0297)	(0.0300)	(0.0931)	(0.0253)	(0.0290)	(0.0156)
num_toddlers	-0.0429	0.130	-0.0441	0.882***	-0.0112	0.0504	0.0162
	(0.0494)	(0.0886)	(0.0566)	(0.325)	(0.0476)	(0.107)	(0.0386)
couple	0.0593	0.185***	0.0957	0.616***	0.0307	0.146***	0.140***
	(0.0570)	(0.0374)	(0.0654)	(0.131)	(0.0551)	(0.0453)	(0.0264)
owner	0.0153	-0.715***	0.300***	-1.452***	0.0344	-0.519***	-0.278***
	(0.0659)	(0.132)	(0.0753)	(0.531)	(0.0647)	(0.169)	(0.0553)
l_earnings	0.765***	0.824***			0.653***	0.856***	0.731***
	(0.0296)	(0.0212)			(0.0333)	(0.0366)	(0.0167)
head works			0.881***	2.651***	0.368***	-0.232	0.178***
			(0.0534)	(0.313)	(0.0519)	(0.157)	(0.0461)
partner works			-0.0396	0.867***	0.0437	0.0613	0.0880**
			(0.0490)	(0.289)	(0.0414)	(0.0960)	(0.0365)
Constant		0.799		3.822**		0.778	3.627***
		(0.565)		(1.523)		(0.490)	(0.140)
Observations	1,026	1,026	1,017	1,017	1,017	1,017	1,017
R-squared	0.764	0.989	0.691	0.893	0.782	0.990	0.930
F age	76.88***	6.36***	43.76***	6.39***	45.29***	8.82***	41.06***
F birth cohort							63.00***

Standard errors in parentheses. num_children: number of dependent children; num_toddlers: number of children 0 to 4 years of age; head/partner works indicates head/partner participates in the labor market; l_earnings: log of household net earnings. *** p<0.01, ** p<0.05, *p<0.1