# Modelling Household Consumption: a long-term forecasting approach

Rossella Bardazzi
University of Florence

# A Multi-Sectoral Approach to model Household Consumption

- **Cross-section Analysis (Income and Demographic Characteristics)**
- **Cohort Analysis (Age and Cohort Effects)**
- **Time Series Analysis (Price Effects)**

# Cross-section Analysis

## Key Equation

Household consumption of product $i$ =

( $f_i$ (household's per capita income) + $g_i$ (demographic factors))

* (family size for product i)

**or**

$$c_{ih} = \left( a_i + \sum_{k=1}^{K} b_{ik} Y_k + \sum_{l=1}^{L} d_{il} D_l \right) \left( \sum_{g=1}^{G} w_{ig} n_g \right)$$

# PLEC (Piecewise-Linear Engel Curve)

The relationship between consumption and income is described by a linear spline, a piecewise-linear Engel curve.
Per capita household income is divided into several income brackets; within each bracket, it is assumed that consumption responds linearly to income.

Households may change their consumption propensity over different brackets: the slopes of these Engel curves represent the 'specific' propensity to consume both with reference to specific goods and to each income level.

# Demographic Characteristics

- Region of Residence
- Family Size
- Age of Household Head
- Education of Household Head
- Number of Income Earners

# Age Structure of the Household

*group* 1: *age* 0 ~ 5

*group* 2: *age* 6 ~ 15

*group* 3: *age* 16 ~ 20

*group* 4: *age* 21 ~ 30

*group* 5: *age* 31 ~ 40

*group* 6: *age* 41 ~ 50

*group* 7: *age* 51 ~ 65

*group* 8: *age* 66 ~ 99

$n_g$ is the number of household members in age category $g$

# To summarize

$$c_{ih} = \left( a_i + \sum_{k=1}^{K} b_{ik} Y_k + \sum_{l=1}^{L} d_{il} D_l \right) \left( \sum_{g=1}^{G} w_{ig} n_g \right)$$

- The estimated ***b*** parameters represent the specific propensities to consume
- The ***d*** parameters measure the effects of demographic variables
- The ***w*** parameters are the adult equivalency weights for measuring the product-specific weighted family size

# Data for Cross-section Analysis

Survey of Italian Household Budgets Survey (IHBS)

- Sample size – 22,000 households approximately;
- Annual total consumption (proxy of income), demographic information, and monthly out-of-pocket expenditures on goods and services, aggregated into 56 categories (to match the *INFORUM* INTIMO model in time-series analysis)

# Two relevant issues

- zero expenditures
- cohort effects

# Handling the Problem of "Zero Expenditures"

- Most expenditure categories have zero observations for more than 50% of the households
- Four types of reasons for zero expenditure (Bardazzi and Barnabani, 1998):
  - Infrequency of purchase
  - Economic decision
  - Conscientious abstention
  - Misreporting

# Consumption vs. Expenditure

- **Consumption** refers to how much of a product, in monetary terms, is consumed;
- **Expenditure** means how much the family spends on certain product

An example: a person **consumes** his car everyday, however, he only **spends** on a new vehicle several times during his lifetime.

# Non only durables…
# The case of transport fuel expenditure

| | N | | No.zeros | | Zeros (%) | |
|---|---|---|---|---|---|---|
| | 1997 | 2013 | 1997 | 2013 | 1997 | 2013 |
| **Whole sample** | 20.604 | 18.672 | 4.879 | 5.870 | 24% | 31% |
| **Given household has a vehicle** | 17.083 | 14.949 | 1.607 | 2.480 | 9% | 17% |
| **Given household has no vehicle** | 3.521 | 3.723 | 3.272 | 3.390 | 93% | 91% |

# Handling the Problem of "Zero Expenditures"

Nonlinear Probability Model

Step one, use a *probit* model to estimate the probability of purchase;

Step two, apply nonlinear least-squares on the non-zero observation to predict the expenditure;

Step three, calculate the expected consumption by multiplying the probability and expenditure, then apply nonlinear least-squares with the expected consumption as the dependent variable.

# When I was your age…

Ageing has effects on many issues, including consumption choices.

There are additional effects related to the passage of time: similarities in experiences and social influences across a particular generation affect its members' choices (**cohort effects**).

Example: transport fuel consumption

When only considering age effects on consumption, we assume that people of different generations may demand less fuel as they age.

However, if a cohort effect exists, then, for instance, the members of the millennial generation may start from a lower base level of consumption and continue to demand relatively less fuel at every stage in their lives

# When I was your age…

Several works in the literature are focussing on this type of effects:

- In studying energy consumption:

*Different* ***social norms****, including individual* ***expectations*** *and* ***aspirations****, interplay with* ***material culture*** *and* ***energy practices*** *in shaping individual behaviour, subject to the external influences that form the context where energy cultures develop.* (Stephenson et al., 2014)

# For example: a new transport culture?

| **Baby boom generation** | |
|---|---|
| **Material culture/ Public Policies** | Automobile-dominated infrastructure |
| **Norms** | Car as a status symbol |
| **Practices** | Big cars, Home purchasing choices and commuting practices |
| **Millennials** | |
| **Material culture/Public Policies** | Public transport infrastructure;<br>Limited Traffic Zones;<br>Emission/Consumption limits |
| **Norms** | New source of prestige; Environmental concern |
| **Practices** | IT innovation widely used to improve transport efficiency and share transport costs;<br>IT technology limits learning/work commuting |

Bardazzi – Pazienza (2018)

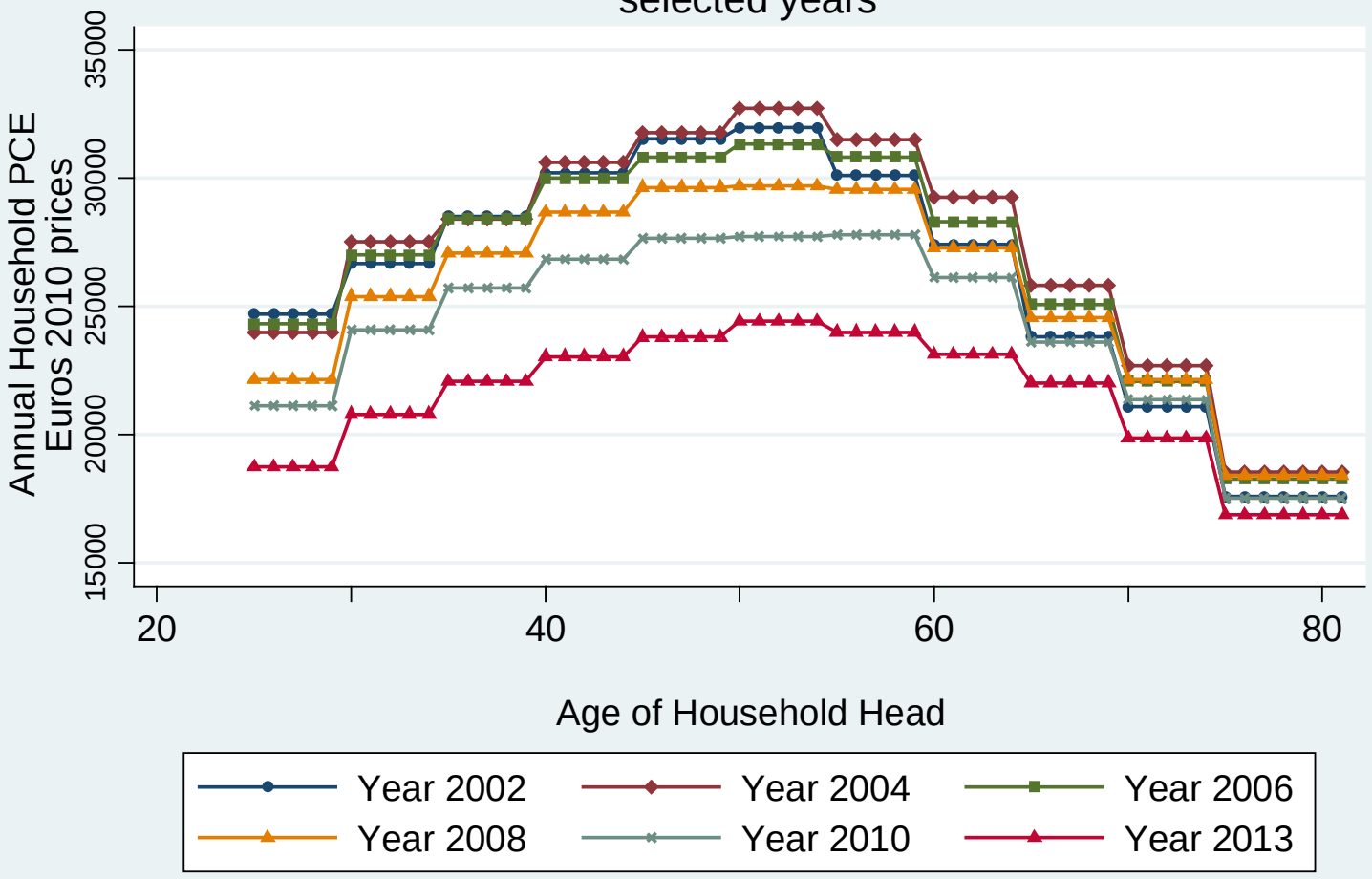
# When I was your age…

- In studying the impact of the economic crisis on household consumption:

*Personal experiences of macroeconomic shocks can have a lasting influence on individual consumer. Individuals who have lived through difficult economic times spend less and reduce the quality of their consumption. The cross-sectional differences vary over time (across cohorts) as households accumulate different experiences.* (Malmendier and Shen (2018), Scarred consumption, NBER wp24696)

# Cohort Analysis

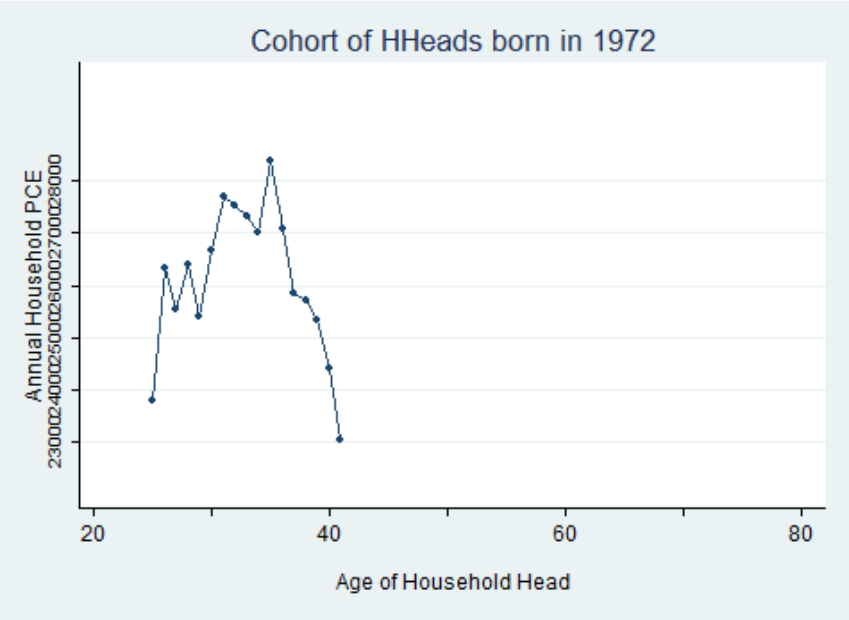
- *Cohort* means a group of people born in the same year.
- *Age*, *Year*, and *Cohort* effect
  - *Age* effect means the behavioral pattern related with the progression of one's life cycle;
  - *Cohort* effect describes the inter-cohort or intergenerational difference;
  - *Year* effect refers to the changes brought by idiosyncratic events that affect all cohorts in a particular year, however it does not represent a time trend.


Cross-sectional PCE profiles
selected years
Annual Household PCE
Euros 2010 prices
15000
20000
25000
30000
35000
20
40
60
80
Age of Household Head
Year 2002
Year 2004
Year 2006
Year 2008
Year 2010
Year 2013

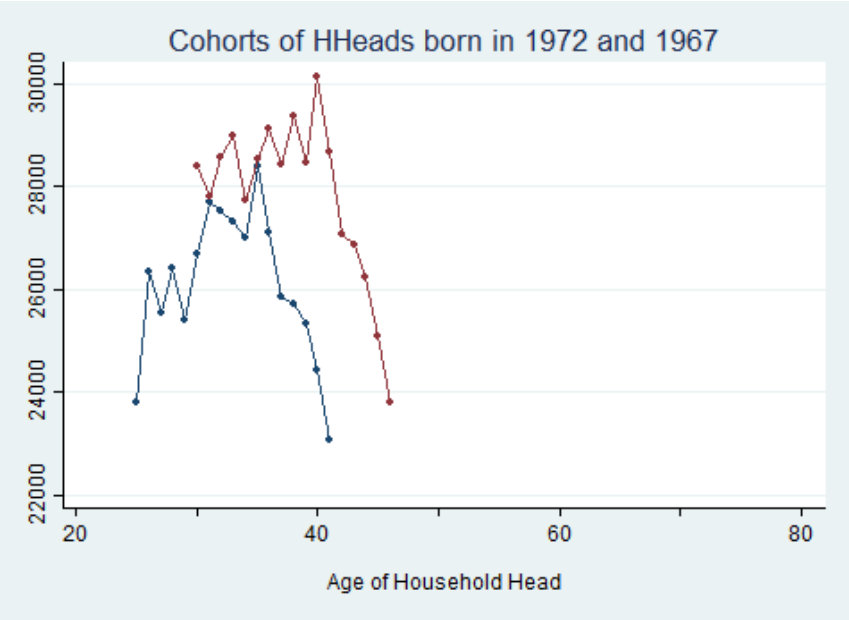
# Constructing Cohort Data

- Use IHBS 1997 – 2013 to build a pseudopanel (repeated cross-sections)
- Cohorts are defined by date of birth of the household head, or more conveniently, age in 1997
- Average the expenditures by age of household head and then track the sample from the same cohort one year older in the next year. For example, we can look at the average consumption of 25-year-olds in the 1997 survey, of 26-year-olds in the 1998 survey and so on

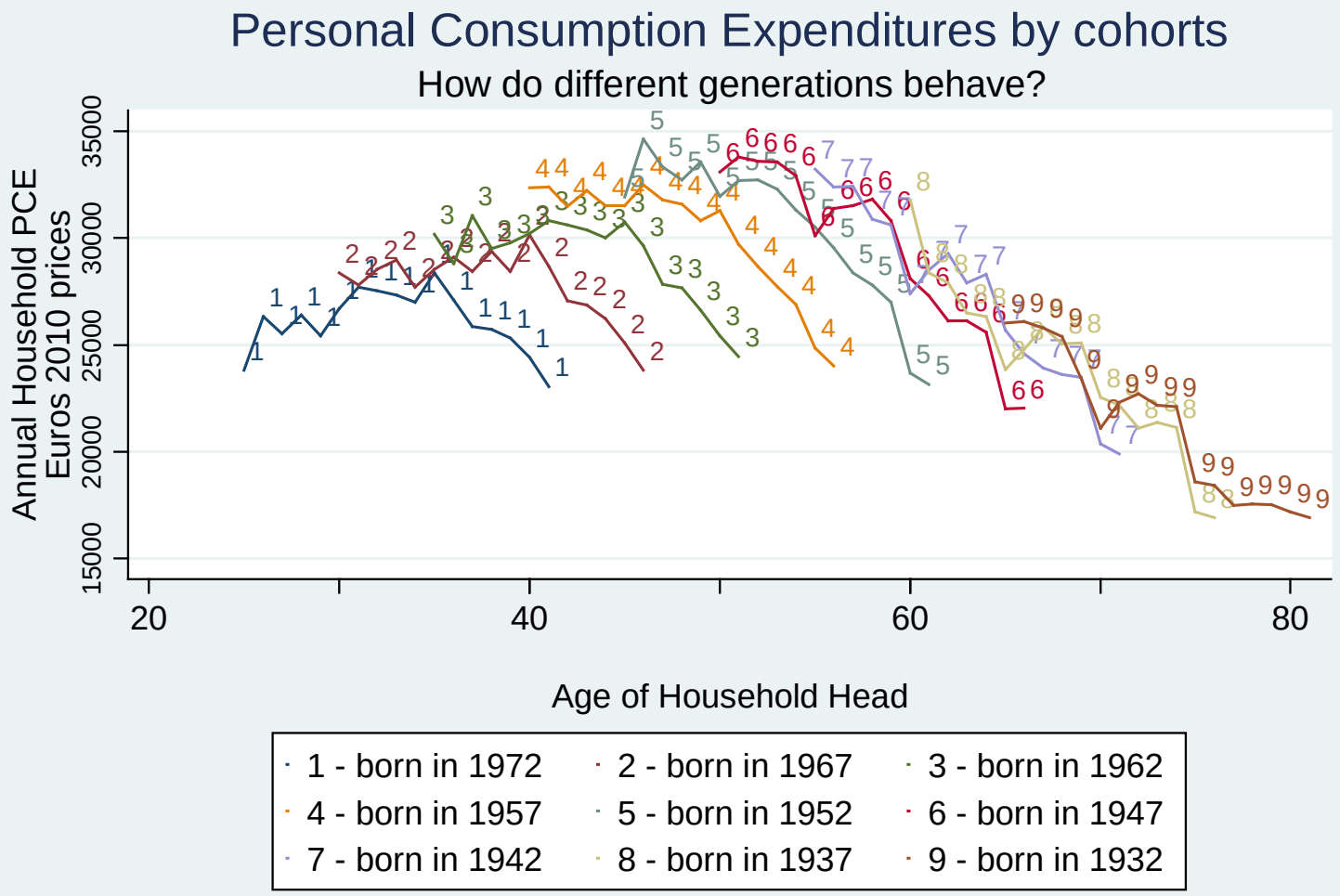
## 25 years old in 1997 to 41 years old in 2013




## 30 years old in 1997 to 46 years old in 2013



# Italian Household Expenditures by Cohorts



# Cohort Analysis

Key Equation

$$x = \beta + A\alpha + C\gamma + Y\psi + u$$

$x$ – log of cohort consumption of each of 56 categories;

$A$ – matrix of age dummies;

$C$ – matrix of cohort dummies;

$Y$ – matrix of year dummies.

# Cohort Analysis

Identification Problem – linear dependency of age, cohort, and year.

If we define cohorts *c* as the age of the household head in year t = 0 and *t* refers to the date, we can infer the cohort's age a as

$$a = c + t$$

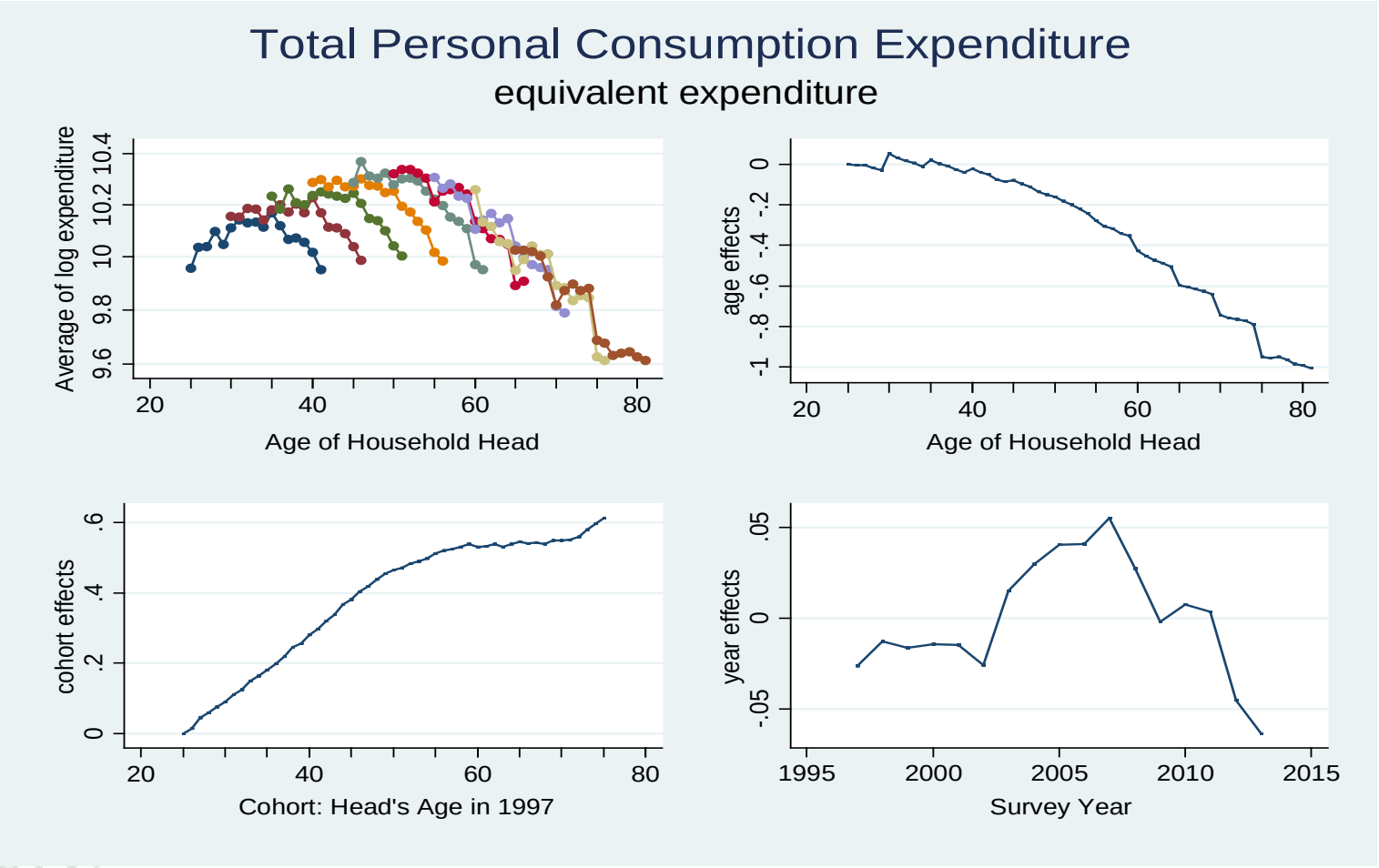
# Solution to the Identification Problem

- Attribute growth to age and cohort effects not year, and to use year effects to capture cyclical fluctuations that average to zero over the long run (Deaton and Paxson, 1994)
- This means we impose the constraint that year dummy coefficients are orthogonal to a time-trend and sum to zero
- A set of T – 2 year dummies is defined as following, from $t = 3, \ldots, T$

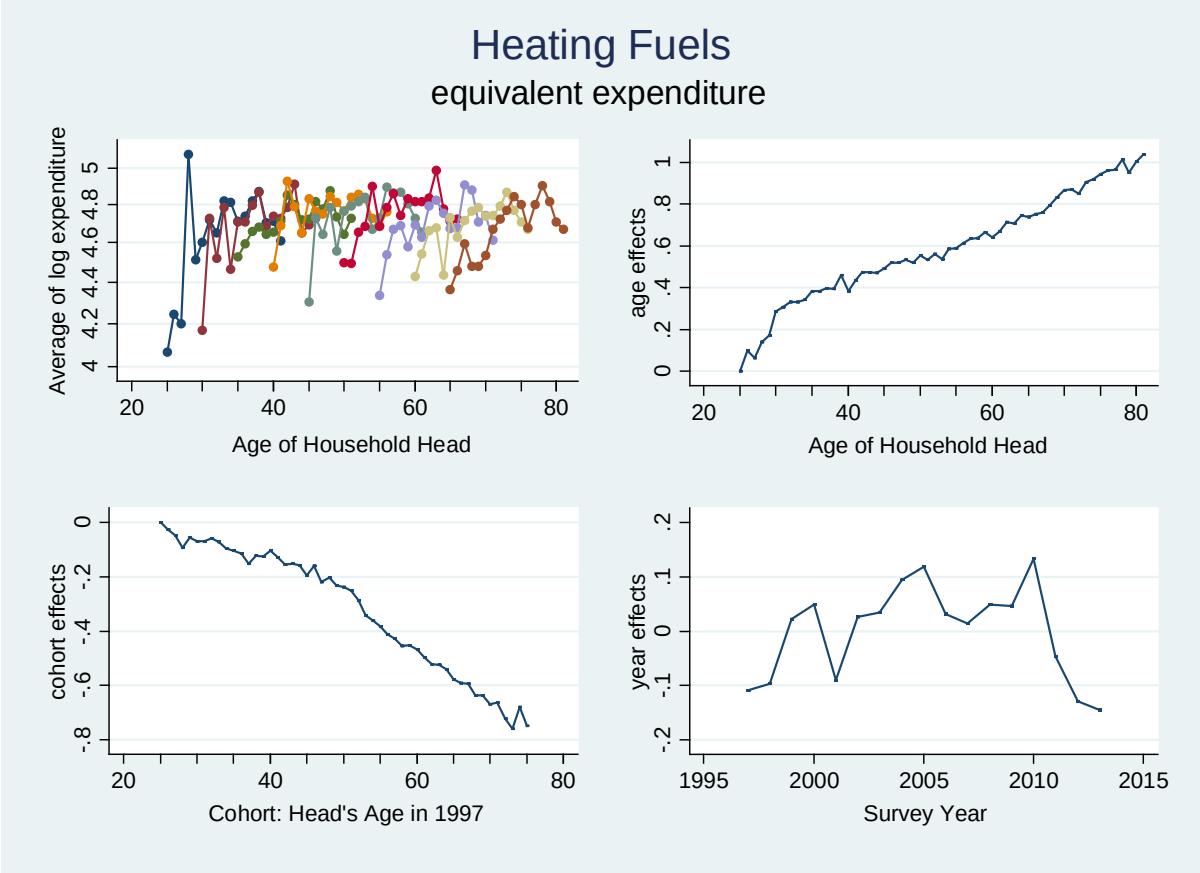
$$d_t^* = d_t - [(t-1)\, d_2 - (t-2)\, d_1]$$

- Where $d_t$ is the usual year dummy, equal to 1 if the year is $t$ and 0 otherwise.

# Estimation Results for Household's Total PCE



# Age and cohort effects differ by products



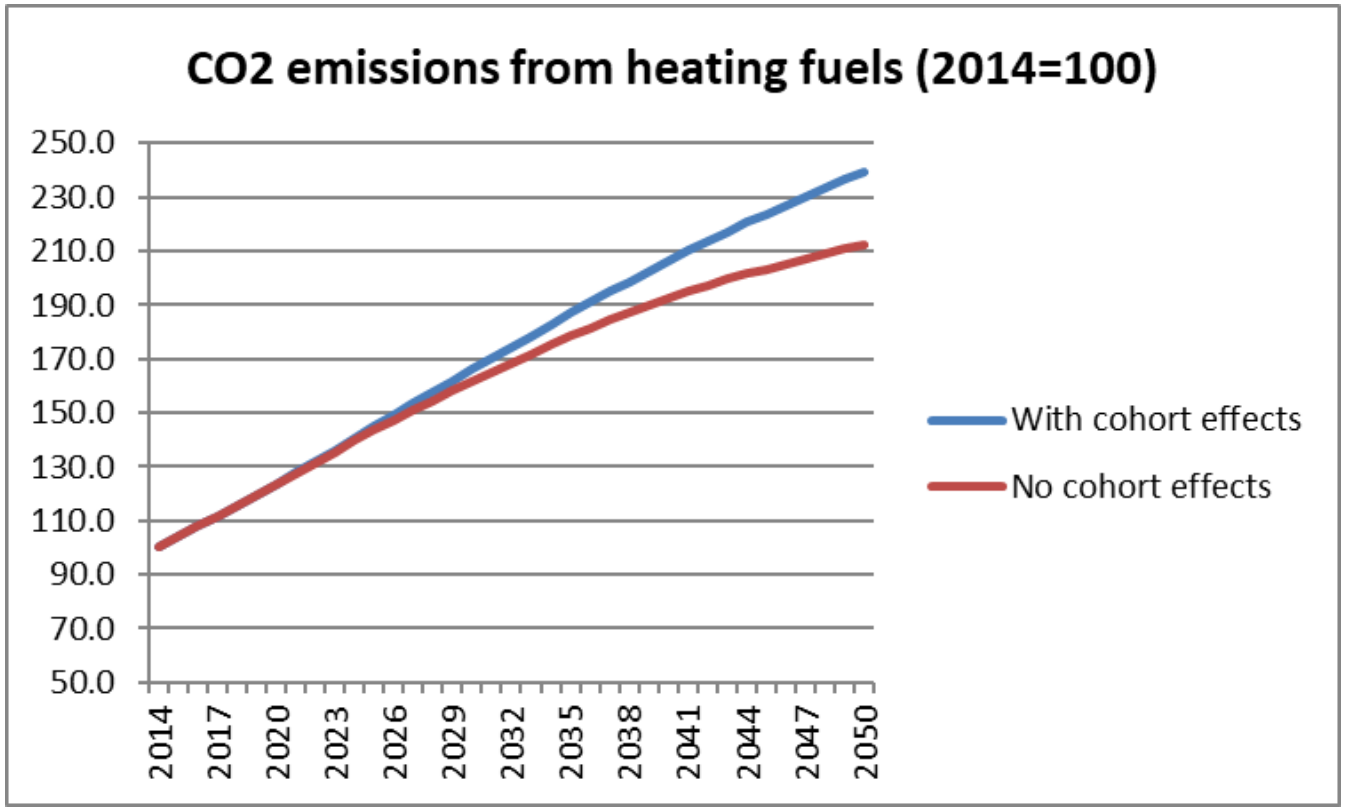

the expenditure on heating fuels at age 50 of someone born in 1963 is on average more than 20 per cent higher than the expenditure at the same age of someone born in 1947

- Population aging does not imply that individuals at the same (old) age but in different generations should behave similarly

- Therefore, **it could be misleading** to attribute the elderly current level of energy consumption to future generations: when the Millennial cohort, which reached adulthood after 2000, will get 60 years old, their consumption choices will be determined by **a set of experiences, norms and technologies largely different** from their parents and grandparents.

- Therefore in modelling PCE in the long run we should try to take into account of these cohort effects besides the pure age effects

# A projection to 2050: heating fuels



# Time Series Analysis

- Study the price effect (own price effects, substitution and complementarity effects between different products)
- Incorporate the information obtained from cross-section study;
- Estimate 56-component Italian Personal Consumption Expenditure in the *INFORUM* INTIMO Model

# A Perhaps Adequate Demand System (PADS, Almon 1996)

- Allows complementarity and substitution effects among different goods.
- Homogeneous of degree zero in price and income.
- Adds up, i.e. the sum of the expenditures on all products should be equal to the total expenditure.
- As income rises, the marginal propensities to consumption should be different for different goods and should depend upon relative prices.
- Allows effects of variables other than prices and income, such as time, and demographic factors.
- Not too complicated to estimate.

# A Perhaps Adequate Demand System (PADS)

$$x_i(t) = \left( \alpha_i + \beta_i \frac{y(t)}{P(t)} + \sum_k^{K_i} \theta_{ik} T_{ik} \right) \prod_{n=1}^{N} p_n^{\delta_{in}}$$

$x_i(t)$ is the consumption *per capita* of product in period t ;

$y(t)$ is a measure of nominal per capita income ;

$P$ is the overall consumer price index;

$T$ are additional variables important to the product *i*;

$p_n$ is the price of product *n* and is equal to one in an arbitrary base year

# A Two-Step Scheme - Step One

- create a time-series variable C* which summarizes all the income, demographic, and cohort information for each product;

$$C_{it}^{*} = b_{i0} + \sum_{k=1}^{K} \frac{b_{ik} y_{kt}}{P_t} + \sum_{l=1}^{L} d_{il} D_{lt} + \sum_{m=1}^{M} c_{im} C_{mt}$$

- Create a time-series of weighted population to consider the age structure

$$WP_{it} = \sum_{g=1}^{G} w_{ig} N_{gt}$$

In red the estimated cross-section and cohort parameters.

# A Two-Step Scheme - Step Two

Estimate PADS using C* to replace the income variable and WP to replace the population variable.

$$\frac{q_{it}}{WP_{it}} = \left( \alpha_i + \beta_i C^*_{it} + \sum_k^{K_i} \theta_{ik} T_{ik} \right) \prod_{n=1}^{N} p_n^{\delta_{in}}$$

# Conclusions

The PCE modelling in our INFORUM model is at present based on information at the cross-sectional, (pseudo)-panel and time-series level

- Is this indirect linkage between micro-based evidence and the time-series analysis the most effective?
- Are there alternative modelling schemes to include all long-run determinants of PCE into multisectoral INFORUM models?

# Thank you for your attention!