PREDICTING WELL BEING IN OLD AGE

Carlotta Montorsi
Luxembourg Institute of Socio Economic Research
University of Luxembourg
Insubria University
21 May 2021

AIM AND SCOPE

- **Population ageing** is one of the key challenges of our times: EU population share above the age of 65 expected to reach almost 30% by 2050 (starting from 19.2% in 2016). Understanding well-being in old age is a **priority**.
- **Current life conditions** and the cumulative impacts of **past life events** influence the reported well being level [1],[2].
- **Environmental** factors relates to well-being but their effect **remains under discovered**.
- The **complexity and high dimensionality** of this mechanism challenge traditional modelling techniques, **undermining the use of well being to inform public policies**[3].
- Data-driven computational approaches open new possibilities toward the understanding of well being data generating process.

OBJECTIVES AND EXPECTED CONTRIBUTION

Focus on five well-being dimensions: **life satisfaction, happiness, depression, health frailty and perceived health**.

The short term objective and contributions are:

- Move along the trade-off between model complexity and interpretability to shed lights on well being outcomes' data generating process.
- Identify the best predictive model.
- Uncover most predictive past life events of each well being outcome.

In the long term we will:

- Integrate **Environmental Big Data** sources in the predictor set.

WELL BEING EQUATION

The well being equation we are estimating takes the following general form [4]:

$$WB_{ij} = h_j(u_j(X_i, Y_i, E_i)) + \epsilon_i$$

For individual i, WB_{ij} is self reported level in Well-Being dimension j, $u(\ldots)$ is to be thought as the person's true well being or utility (unobserved), $h(.)$ is a the function that relates the actual well being to the reported well-being. X_i, Y_i and E_i are the past life events, demographics and environmental factors. ϵ_i is an error term.

METHODOLOGY

- **Supervised learning algorithms** outperform least squares modeling technique in predictive task [5],[6].
- We optimize through 10-fold cross validation six different algorithms: least squares regression, elastic net, conditional inference trees/forest, traditional CART, Support Vector Machine and deep neural network.
- Different training and test sets in order to assess performance across different setting.
- We calculate models specific features importance for each well being outcomes and training set.

DATA

The data we used are based on the Survey of Health, Aging and Retirement in Europe (SHARE):

- **Individual-level data** on health, socio-economic status and social and family networks aged 50+ from 29 European countries and Israel.
- **Panel structure**: repeated observations every two year plus refreshment sample, from 2004 to 2018.
- **SHARELIFE questionnaire (wave 3 (2008/09) and wave 7 (2017))**: Retrospective information about individuals' life histories.

TRAINING-TEST

Training sample	Dimension (N)	Test sample	Dimension (N)
Pooled (80%)	$\sim$ 36407.8	Pooled (20%)	$\sim$ 9100
2008/09 Wave 3	$\sim$ 19068.6	2017 Wave 7	$\sim$ 24451.8
13 Countries (80%)	$\sim$ 2948.54	Country (20%)	$\sim$ 735.18

Training and test samples, average number of observations (N) among the five well being outcomes

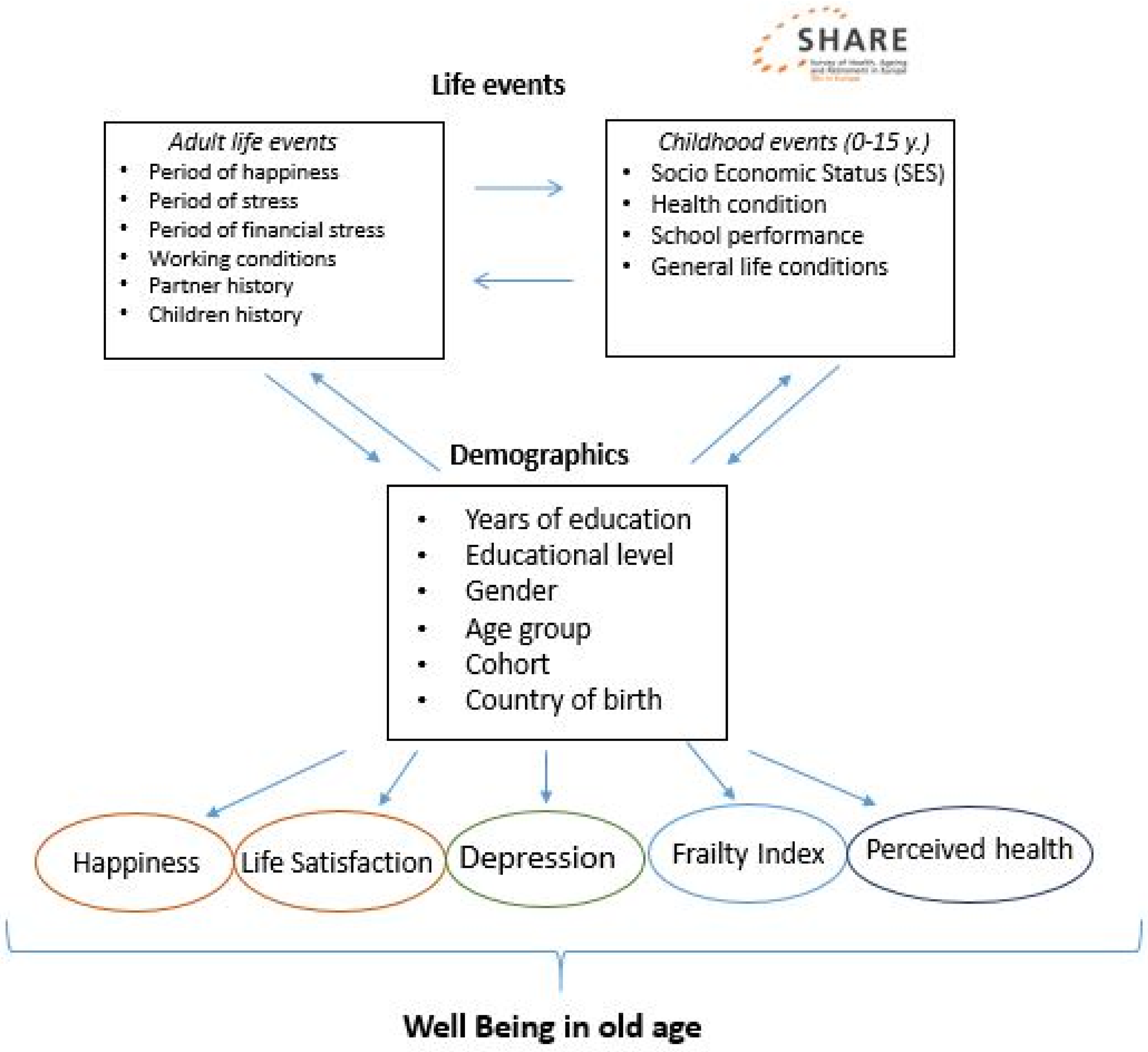

Figure: Illustration of the framework

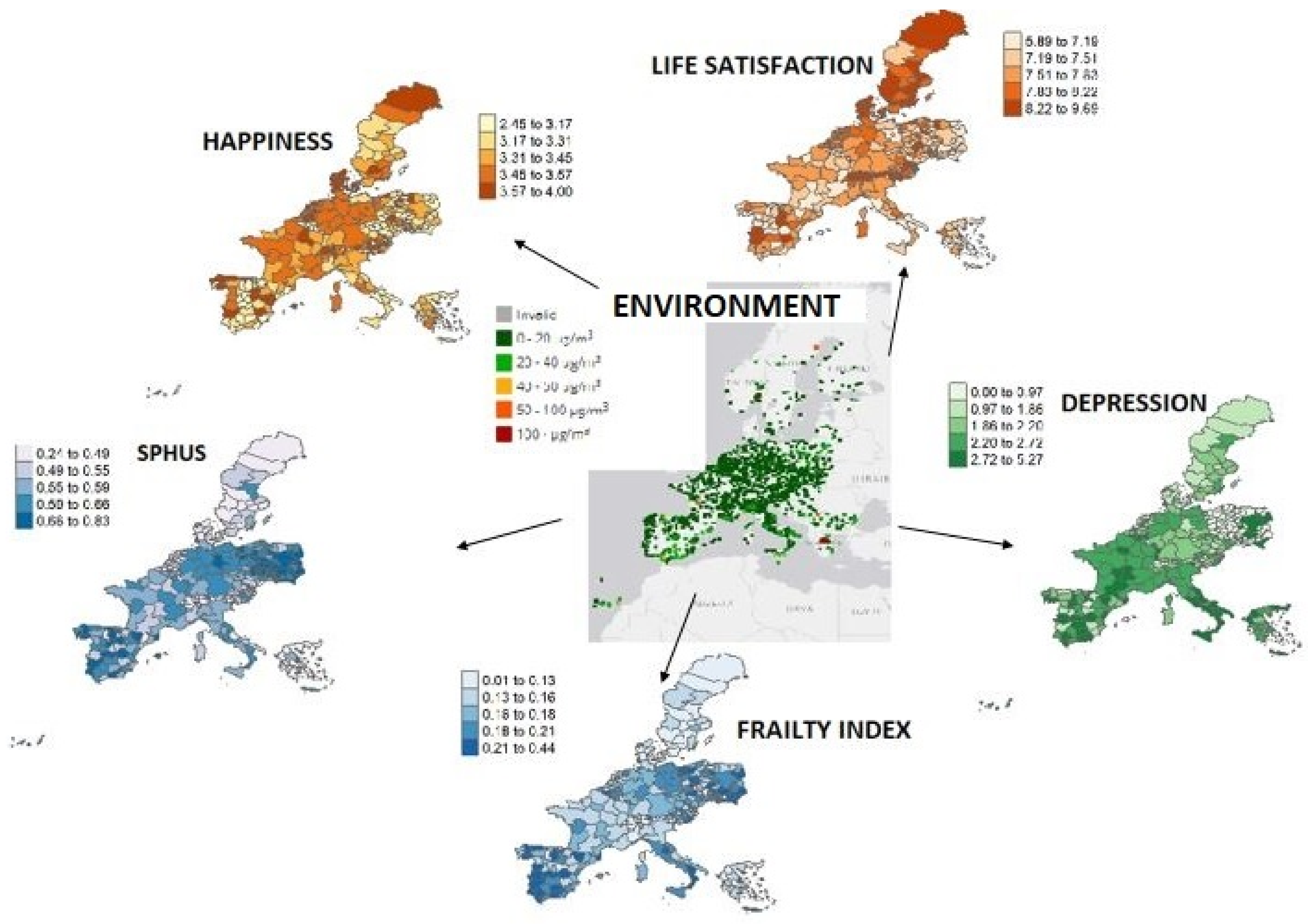

Figure: What is the role of Environment?

ACKNOWLEDGEMENT

The Doctoral Training Unit **Data-driven computational modelling and applications** (DRIVEN) is funded by the Luxembourg National Research Fund under the PRIDE programme (PRIDE17/12252781). `https://driven.uni.lu`

REFERENCES

[1] Andrew E Clark and Tom Lee. "Early-life correlates of later-life well-being: Evidence from the Wisconsin Longitudinal Study". In: *Journal of Economic Behavior & Organization* (2017).

[2] Peter Martin. "Individual and social resources predicting well-being and functioning in the later years: Conceptual models, research and practice". In: *Ageing International* 27.2 (2002), pp. 3–29.

[3] Ed Diener et al. *Well-being for public policy*. Series in Positive Psychology, 2009.

[4] David G Blanchflower and Andrew J Oswald. "Well-being over time in Britain and the USA". In: *Journal of public economics* 88.7-8 (2004), pp. 1359–1386.

[5] Trevor Hastie, Robert Tibshirani, and Jerome Friedman. *The elements of statistical learning: data mining, inference, and prediction*. Springer Science & Business Media, 2009.

[6] Brett Lantz. *Machine learning with R*. Packt publishing ltd, 2013.