

Some questions

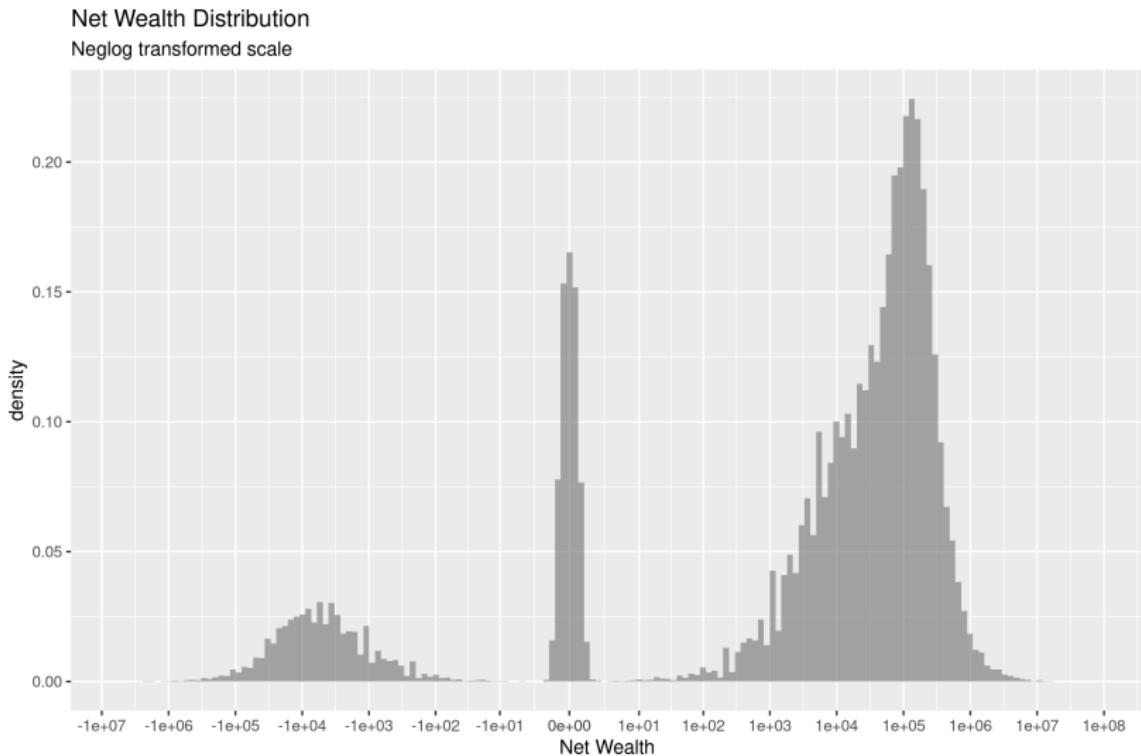
Marcelo Avila

09 August 2021

Preview on Panel Structure

- ▶ About 8.800 individuals if balancing sample for which health and wealth variables are present

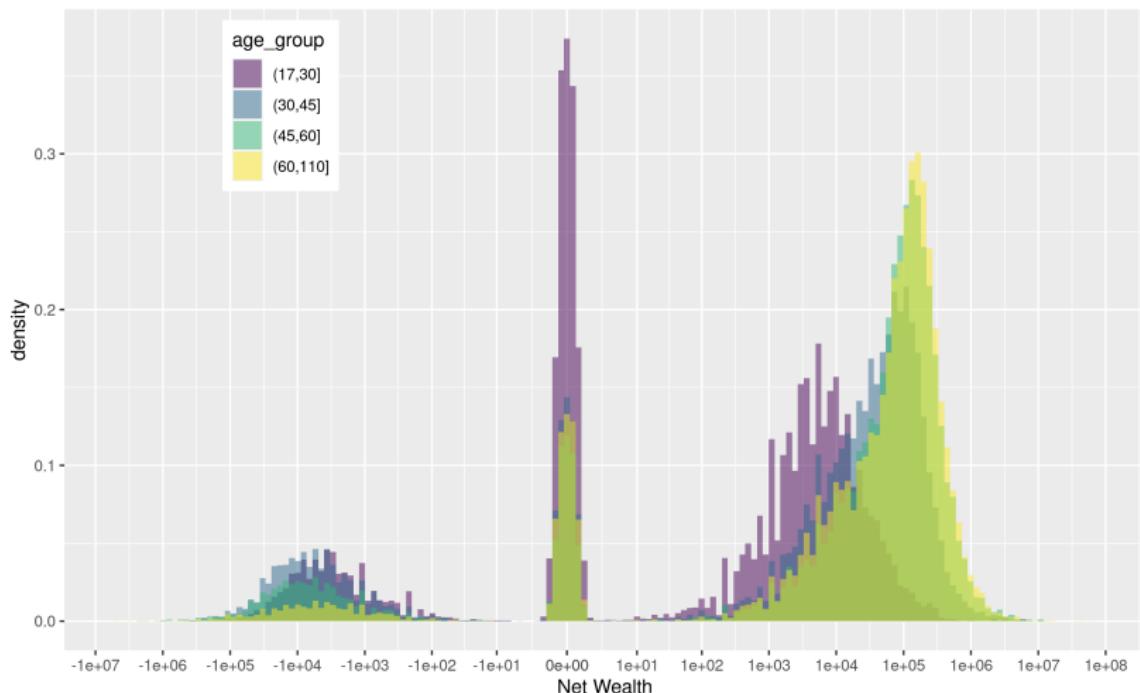
Wealth I



Wealth II

Net Wealth Distribution

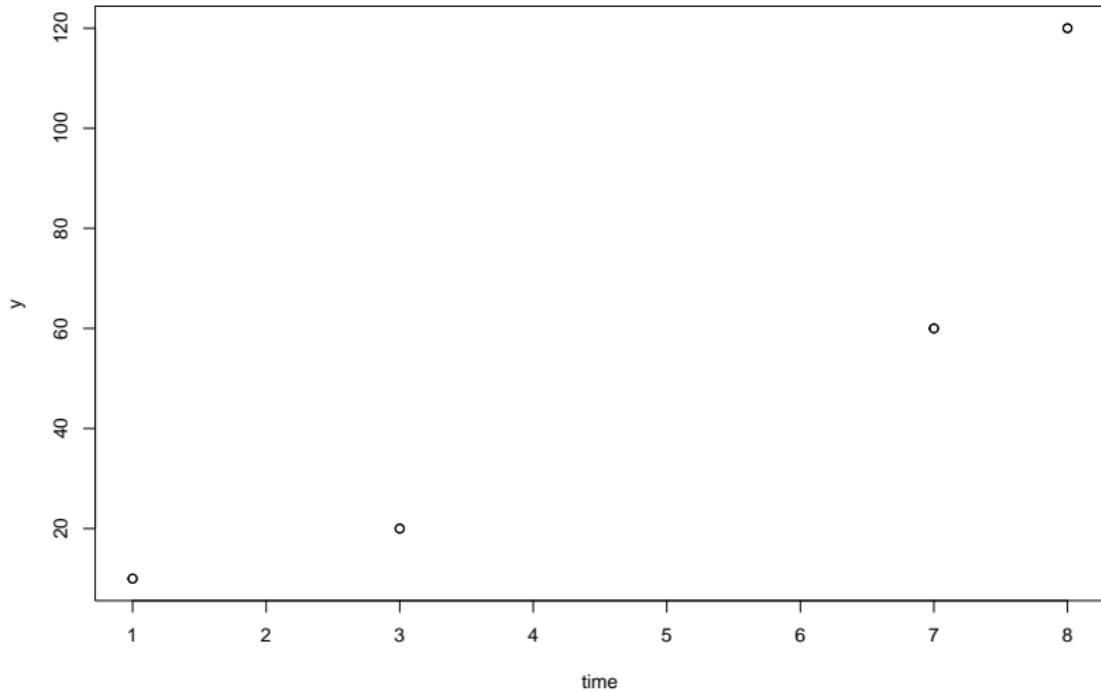
Neglog transformed scale



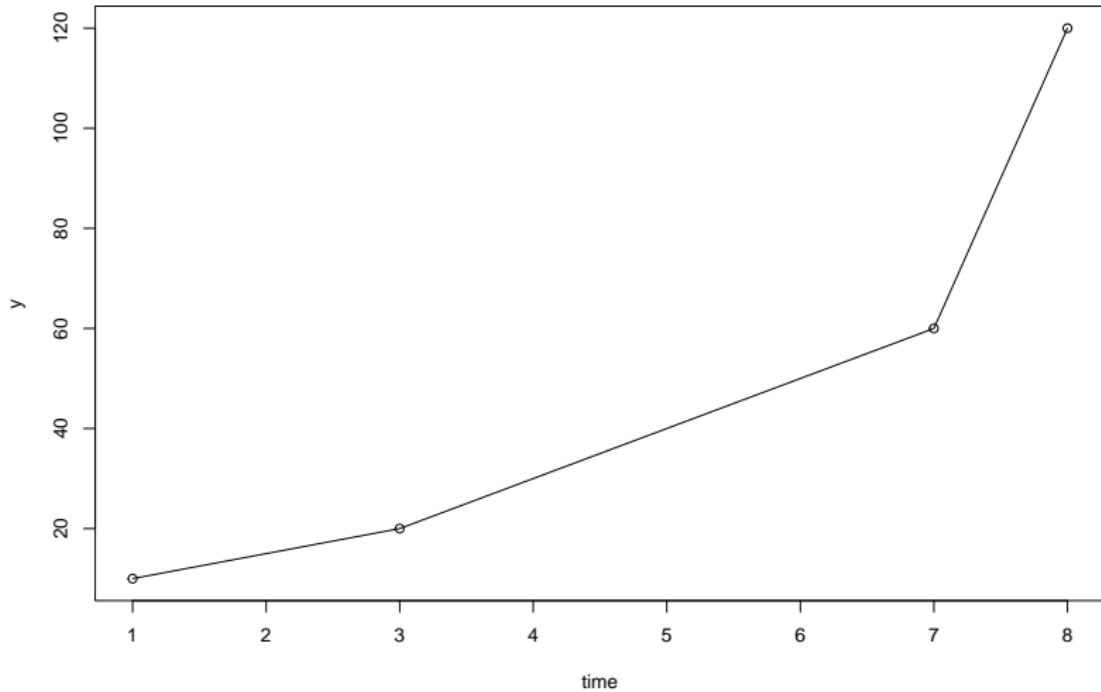
Interpolation of Wealth variables

- ▶ Wealth variables not available for all (odd) years, so I imputed by linear interpolation for the missing years.
- ▶ It is somewhat based on strong assumptions (linear development of wealth) but I believe better than other simple alternatives, such as last-observation-carried-forward (LOCF)
- ▶ some robustness checks would be necessary here, specially on dealing with missing data at the one or both “ends”.

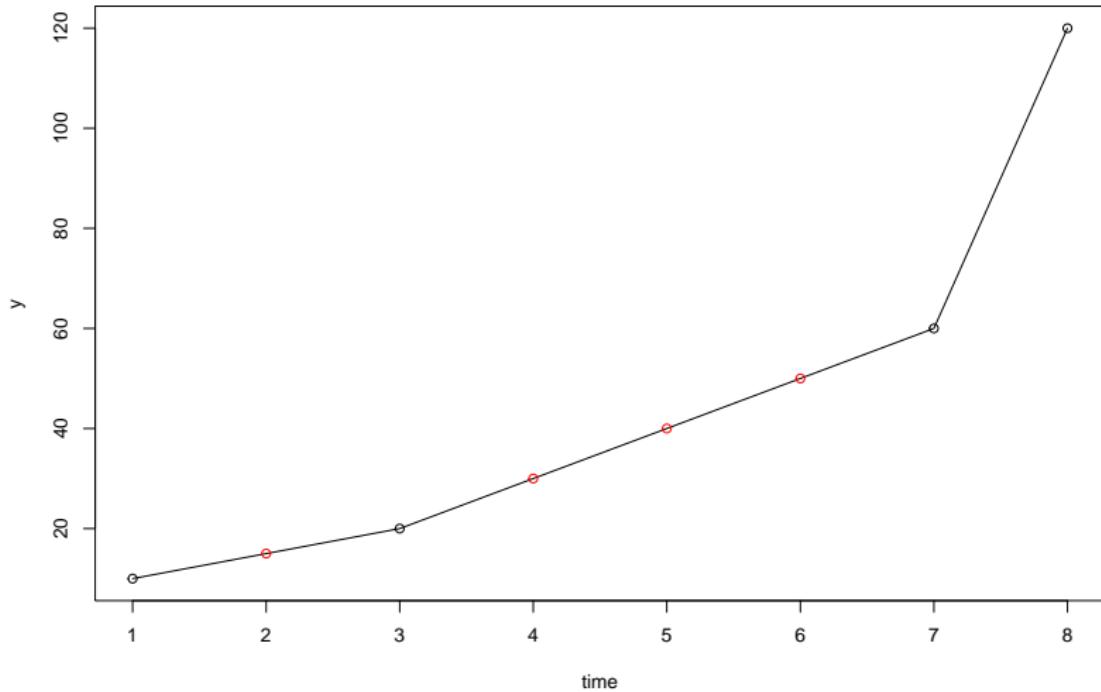
Linear interpolation example



Linear interpolation example



Linear interpolation example



Questions

- ▶ keep other samples as controls?
 - ▶ including sample that starts later? -> including more observations slightly less noisy (due to sample size) -> could be better for inference (controls more similar to)

Survey	to_keep		
Year	0	1	Total
2011	12,586	8,768	21,354
2013	10,597	8,768	19,365
2015	8,681	8,768	17,449
2017	18,024	8,803	26,827
Total	66,798	35,107	101,905

- ▶ Wealth module for 2019?
- ▶ Sample at risk for each disease?
- ▶ On the box-plots

On the box-plots | Original

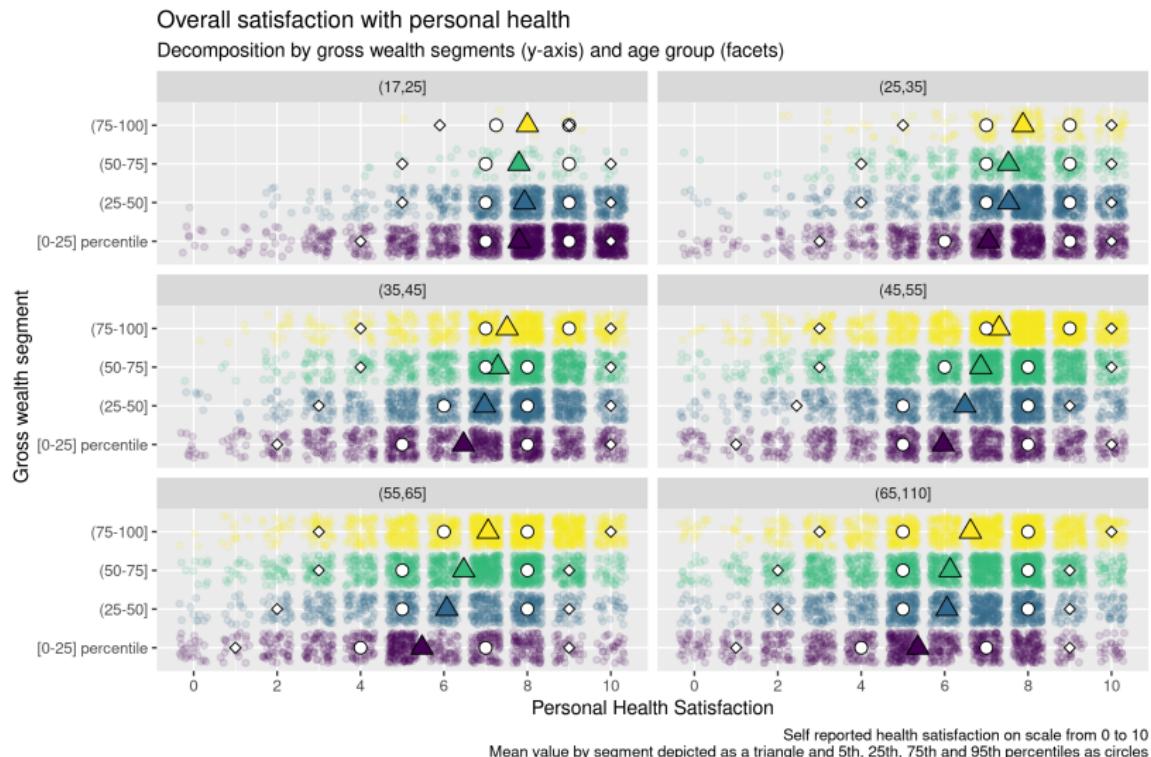


Figure 1: fig_boxplot_plus_mean.png

On the box plots | Box Plot (incl. average)

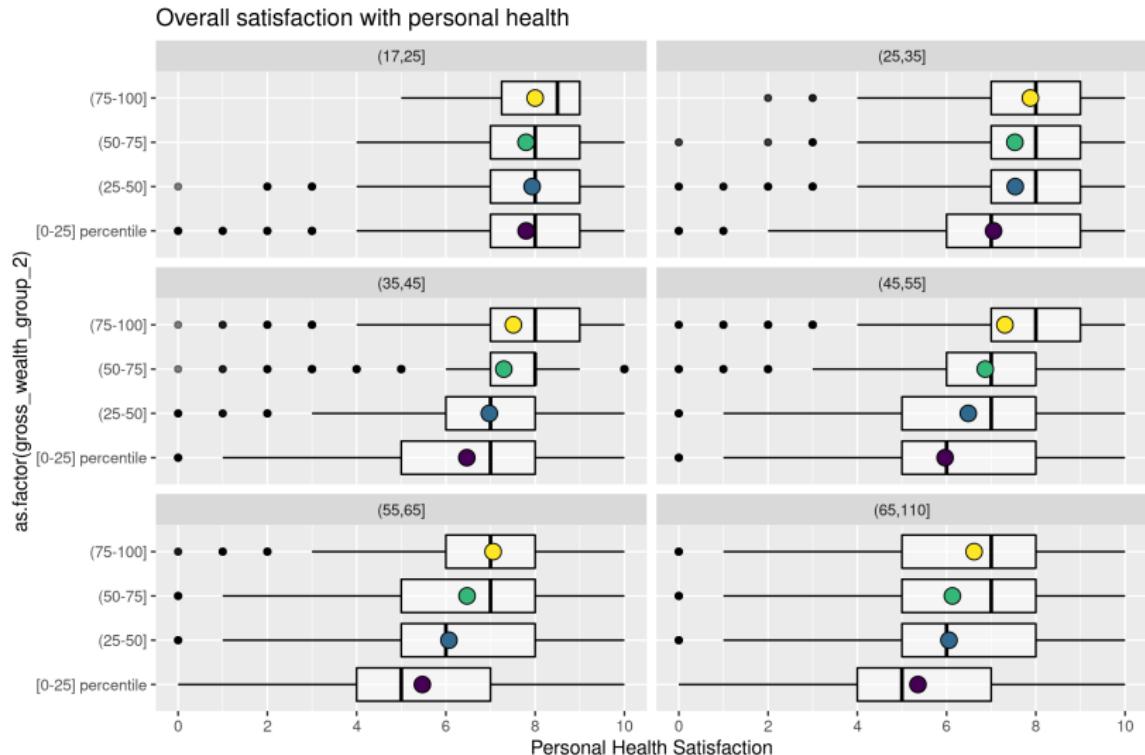


Figure 2: fig_boxplot_plus_mean.png

On the box plots | Both together

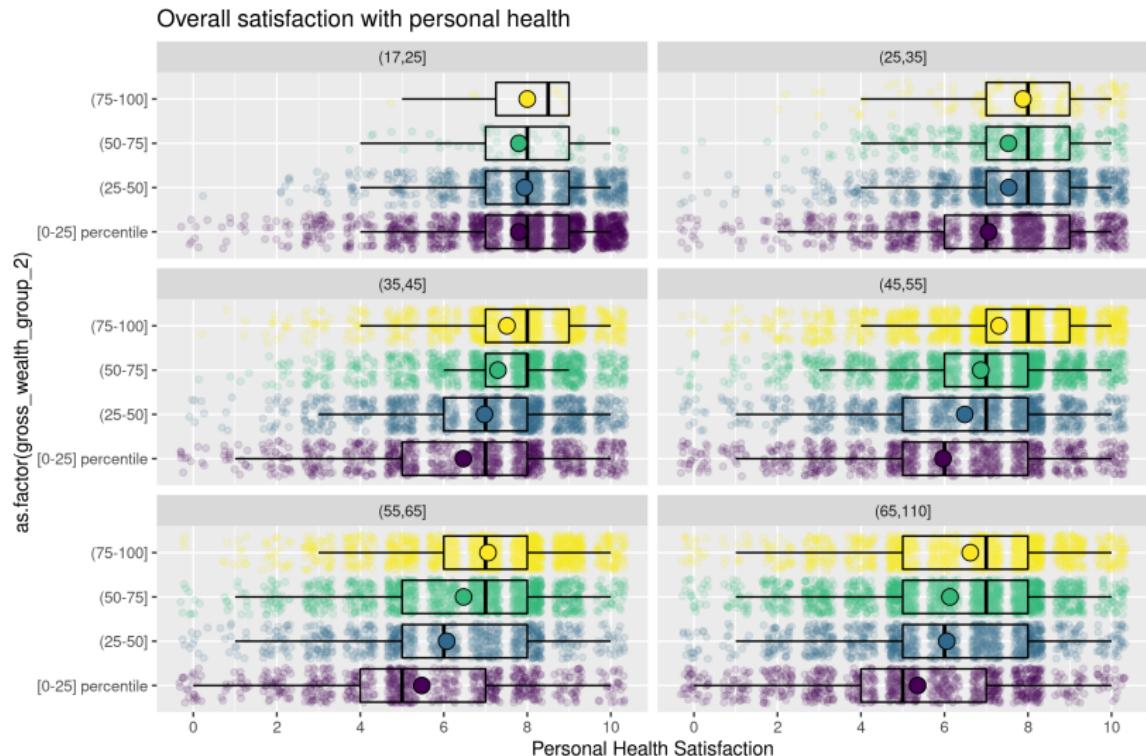


Figure 3: fig_boxplot_plus_mean.png

Outlook

- ▶ gather control variables
 - ▶ education, gender, labour-related, etc)
- ▶ run the regressions