Online Supplemental Materials: Specification curves and RD plots for all outcomes

1. Specification curves for all RD models linking the WFP and health outcomes

We expand our analysis by reporting specification curves for every outcome analysed in the main text. Specification curves are an analytical tool that aims to portray a set of estimates from a plausible set of valid specifications. It can be seen as a method to robustness of data to different choices available to the researcher and to attenuate researchers decisions – at the same time "defensible, arbitrary and motivated" (Simonsohn et al. 2015, p.1)– on variable selections, construction, specification forms. For every single research question – including ours – there is an abundance of such valid specifications. This analytical tool is drawn from Simonsohn et al (2015) and is based on two steps:

- Define the set of reasonable, plausible set of specifications;
- Report the estimates (and their standard errors) in a graphical form to ease comparison.

The identification of valid specifications is not an exact science, but it is rather the culmination of an open process that starts from "enumerating all of the data analytic decisions necessary to map the scientific hypothesis onto a statistical hypothesis" (Simonsohn et al. 2015, p.7) to "enumerating all the reasonable alternative ways a researcher may make those decisions" (Simonsohn et al. 2015, p.7).

It is worth noting that in our main text, we presented some robustness tests that already embrace this process of decision-making. Our tables show how our estimates vary across different samples of interest (e.g., January to March vs January to April) and models (e.g., bandwidths selected with and without regularization). Moreover, every health outcome is define in different ways when possible, i.e., using the nurse measurement as it is (its continuous form) or by constructing binary variables. However, in what follows we extend the analysis presented in the main text by adding three more specifications and an alternative ways to dichotomise two health outcomes. These approaches were discussed by the research team before and during the analysis and represent the set of non-redundant specifications culminated from our deliberation.

The legend on the bottom left of each specification curve clarifies which combination was used to obtain the point estimate plotted in each graph.

In summary, the following figures from S1 to S12 plot the average treatment effects already presented in the main text and extend this by adding three new specifications:

- RD model without covariates (i.e., when the label *Covariates* in the legend at the bottom-left is turned off). The literature on RD designs makes clear that covariates are not needed to help or hinder the causal identification of the treatment effect, and are only included to improve precision. The list of our covariates do not have an obvious causal interpretation, this is also why we do not report them. Using a *bad* set of covariates could actually decrease the precision rather than increasing it, which is why in the end there was a very short list of covariates. We discuss this in great detail in the methodology chapter. To improve robustness we estimated every single model without any covariate and present these in the following specification curves.
- Fuzzy RD model (i.e., when the label Fuzzy is turned on). Our methodology is based on the sharp eligibility criteria of the WFP. We estimated and reported *sharp* RDDs. In brief, we use the WFP indicator variable available in ELSA and instrument this using our assignment variable, constructed using quarter of births. This should attenuate measurement errors from self reports.
- RD model with bias-corrected estimates (i.e., when the label BiasCorr is turned on). In practice, estimations that utilise optimal bandwidth selection based on MSE will introduce an asymptotic bias. Calonico et al. (2014) propose a procedure to obtain a bias-corrected estimator by subtracting the estimated bias term from the estimated treatment effect.

References:

Calonico S, Cattaneo MD, Titiunik R. Robust nonparametric confidence intervals for regression-discontinuity Designs. *Econometrica* 2014;**82**(6):2295-326.

Simonsohn U, Simmons JP, Nelson LD. *Specification curve: Descriptive and inferential statistics on all reasonable specifications* 2015. Available at SSRN: https://ssrn.com/abstract=2694998 or http://dx.doi.org/10.2139/ssrn.2694998 (accessed 4 December 2017)

FIGURE S1 Specification curve for indoor air temperature

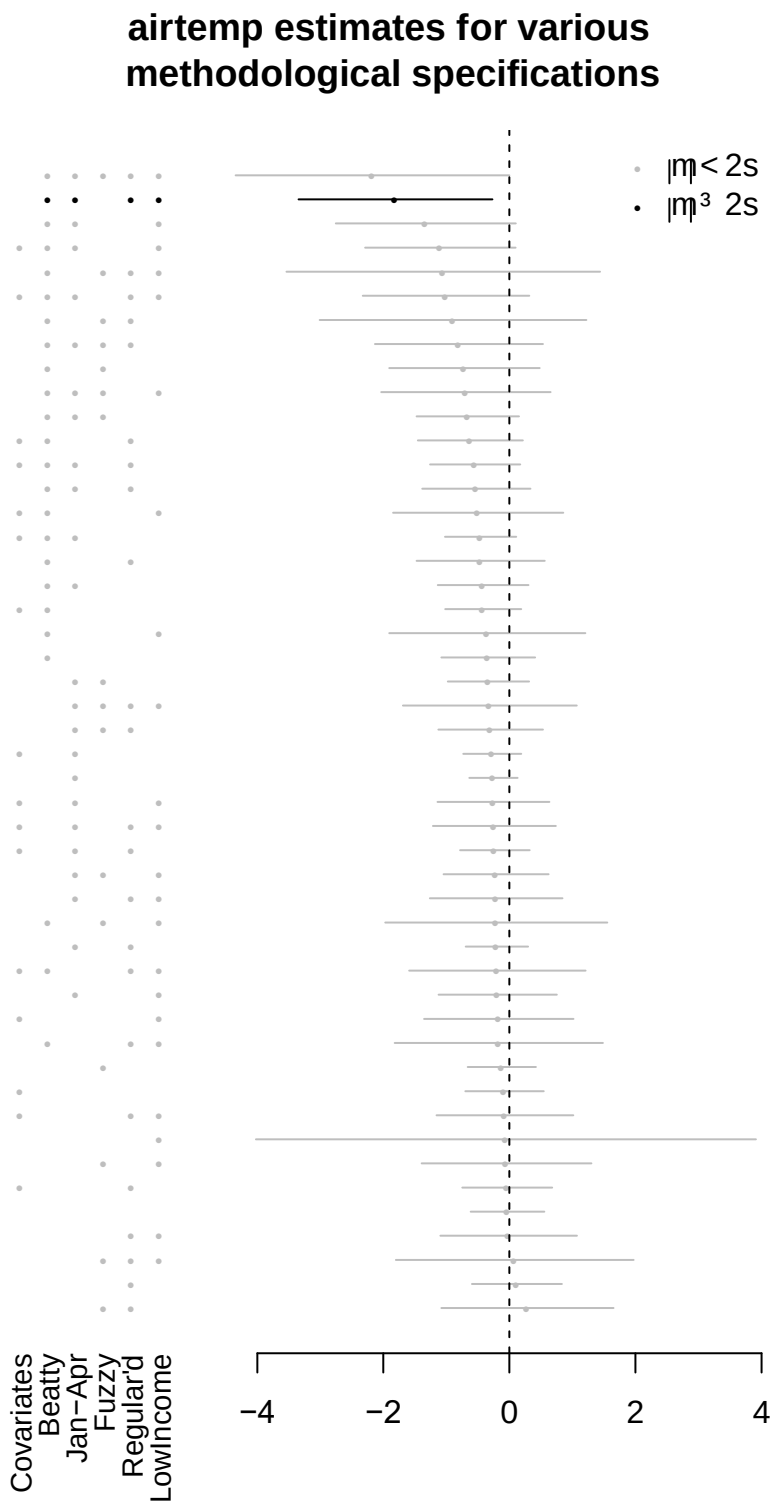


FIGURE S2 Specification curve for systolic blood pressure

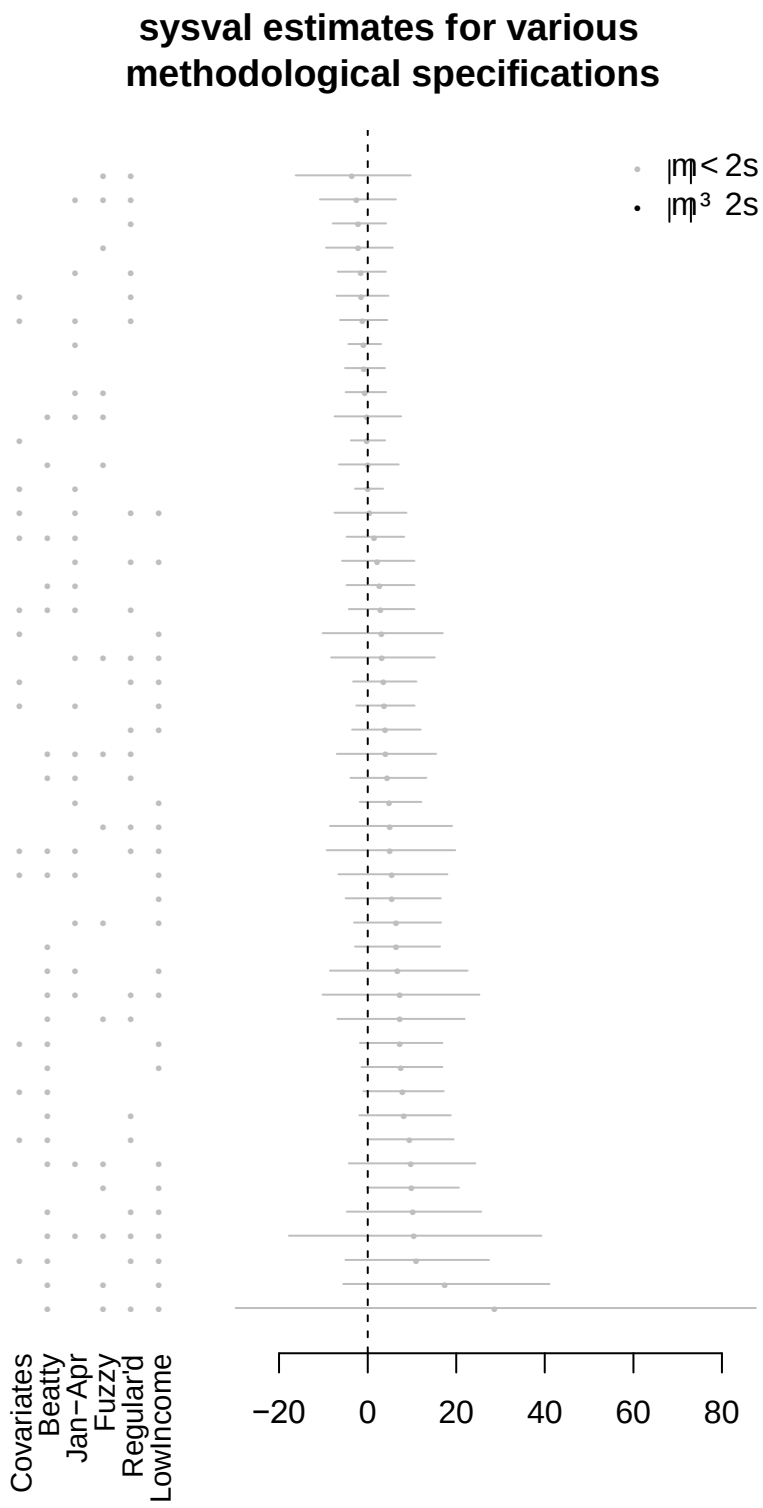


FIGURE S3 Specification curve for *binary* systolic blood pressure (dichotomised at 140)

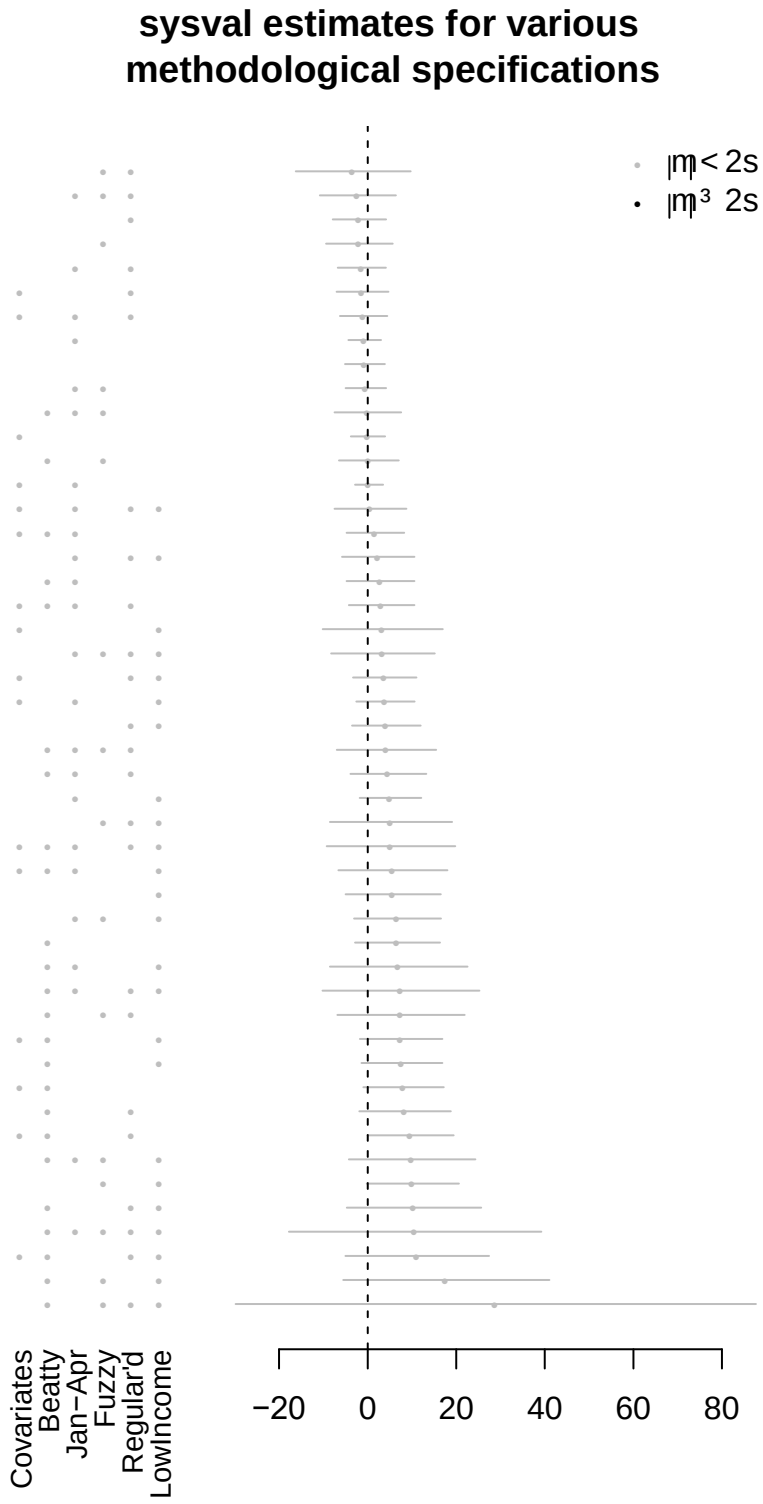


FIGURE S4 Specification curve for fibrinogen

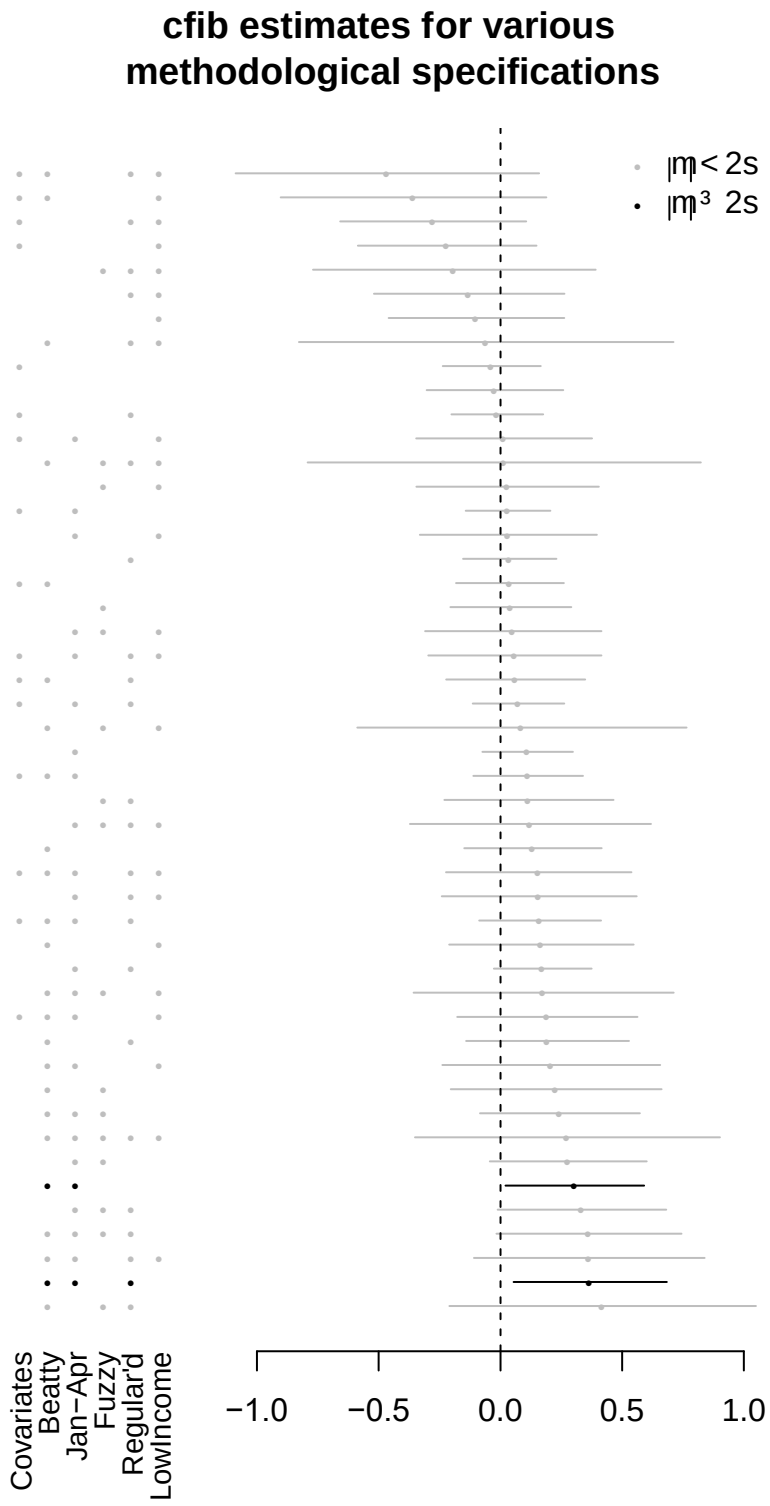


FIGURE S5 Specification curve for *binary* fibrinogen (fibrinogen dichotomised at 4 g/L)

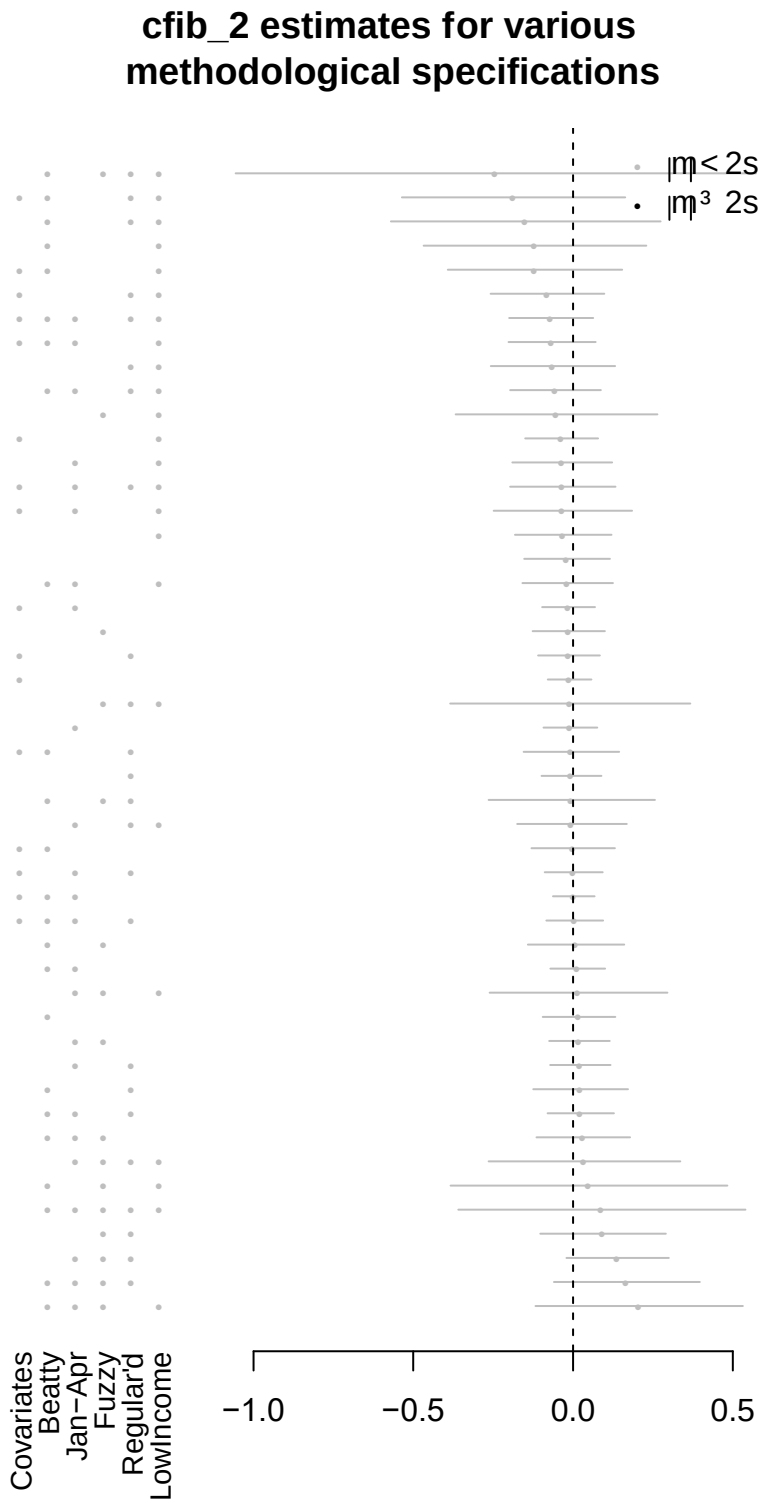


FIGURE S6 Specification curve for (log of) high sensitive c-reactive protein

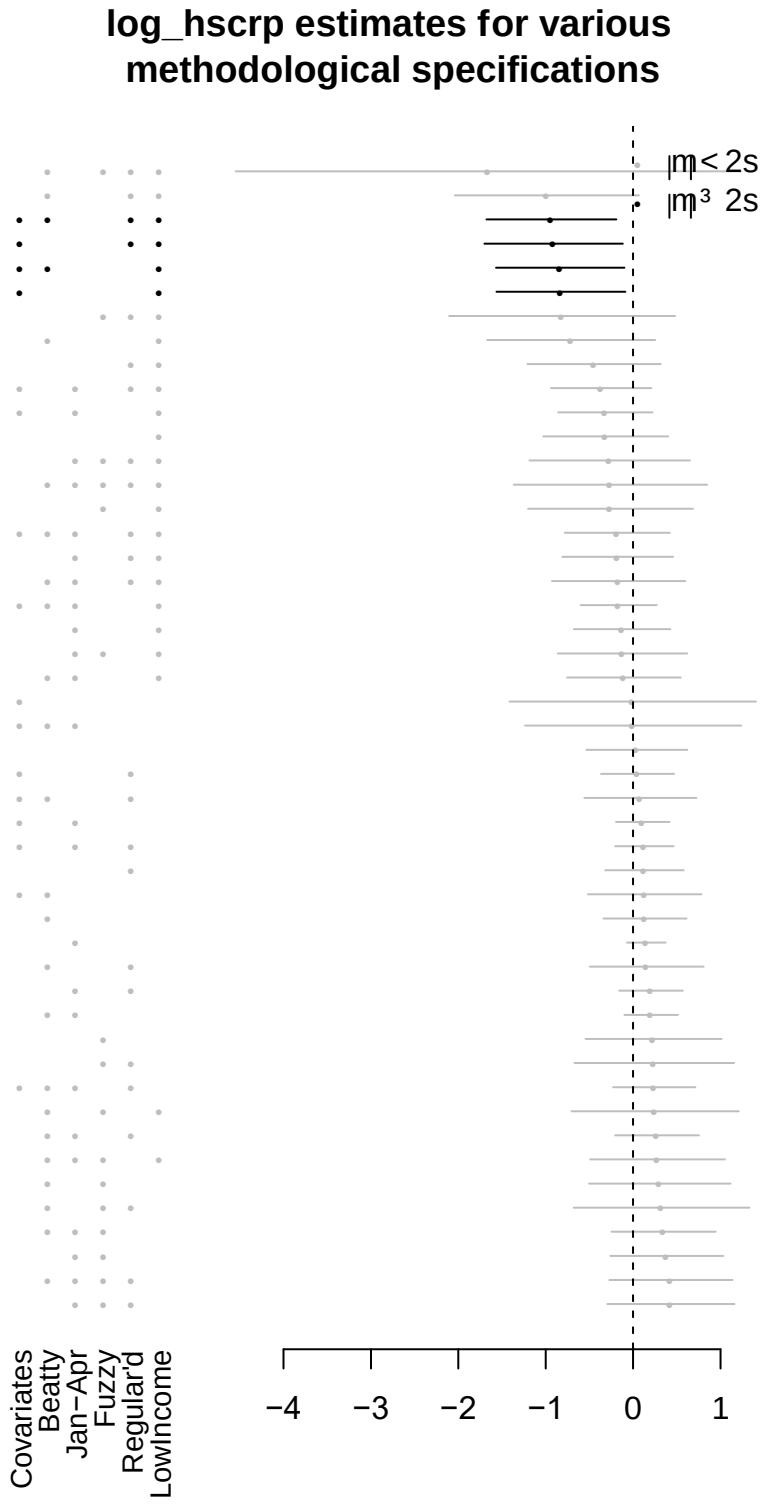


FIGURE S7 Specification curve for *binary* c-reactive protein (dichotomised at 3 g/L)

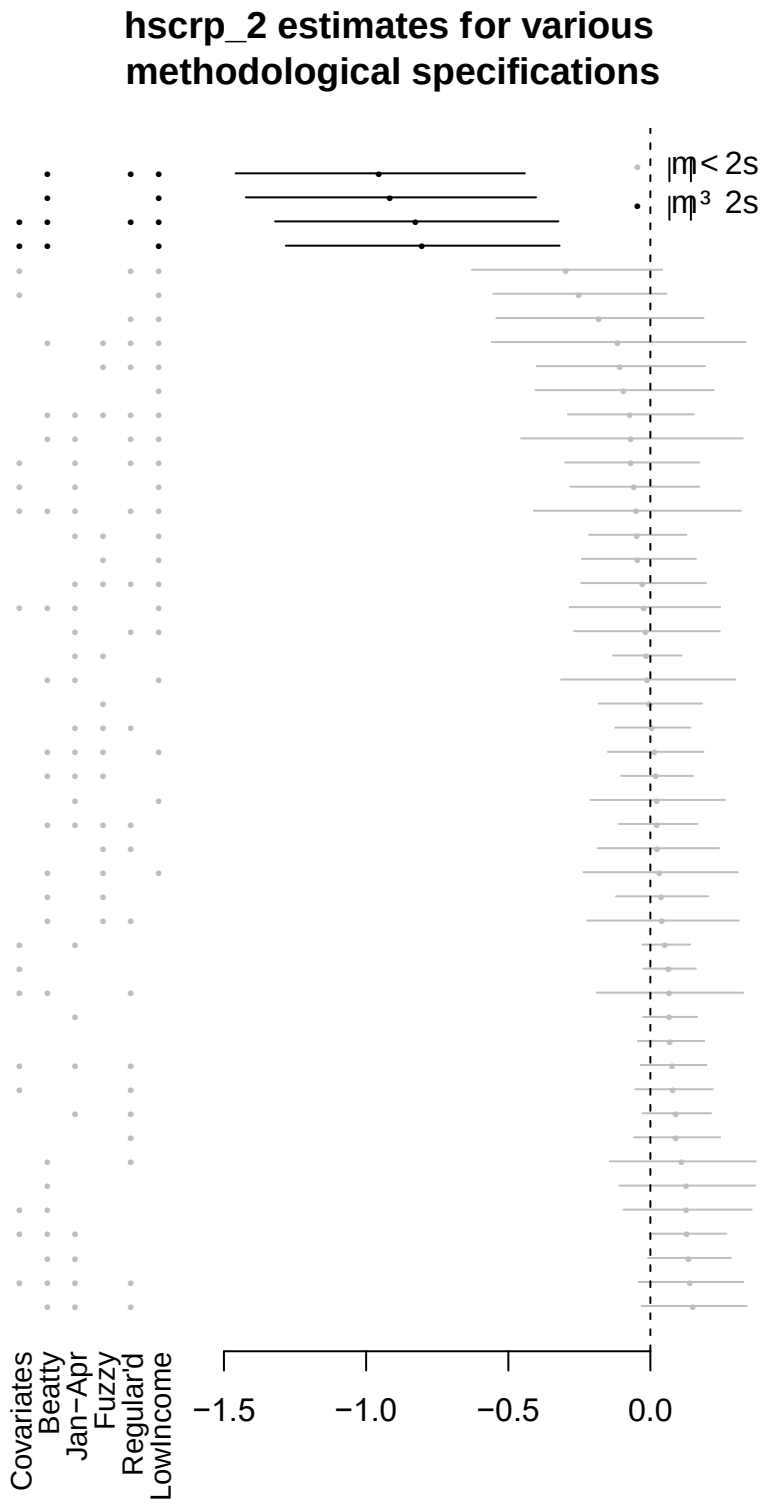


FIGURE S8 Specification curve for self-reported chest infection

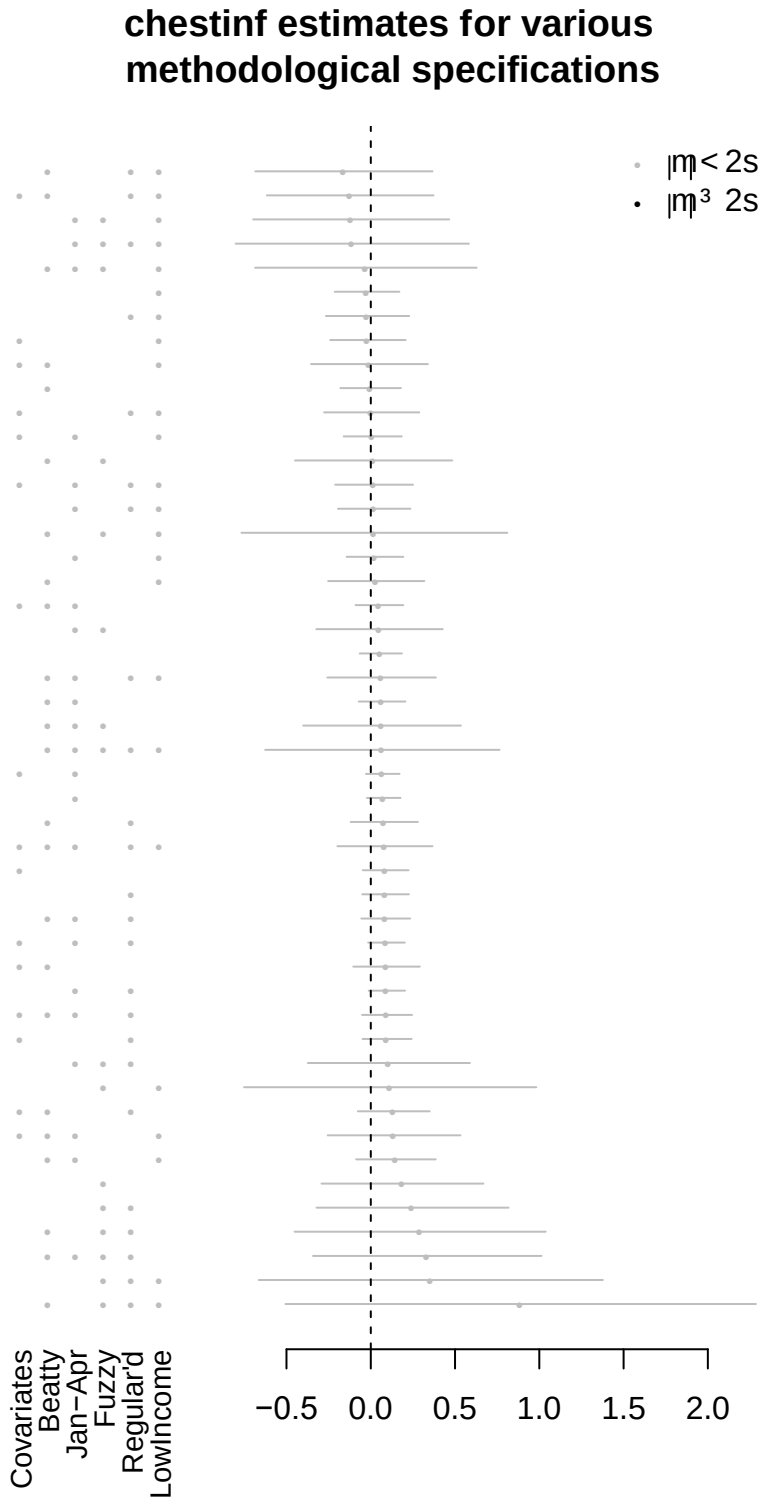


Figure S9 Specification curve for Forced Expiratory Volume (z-score)

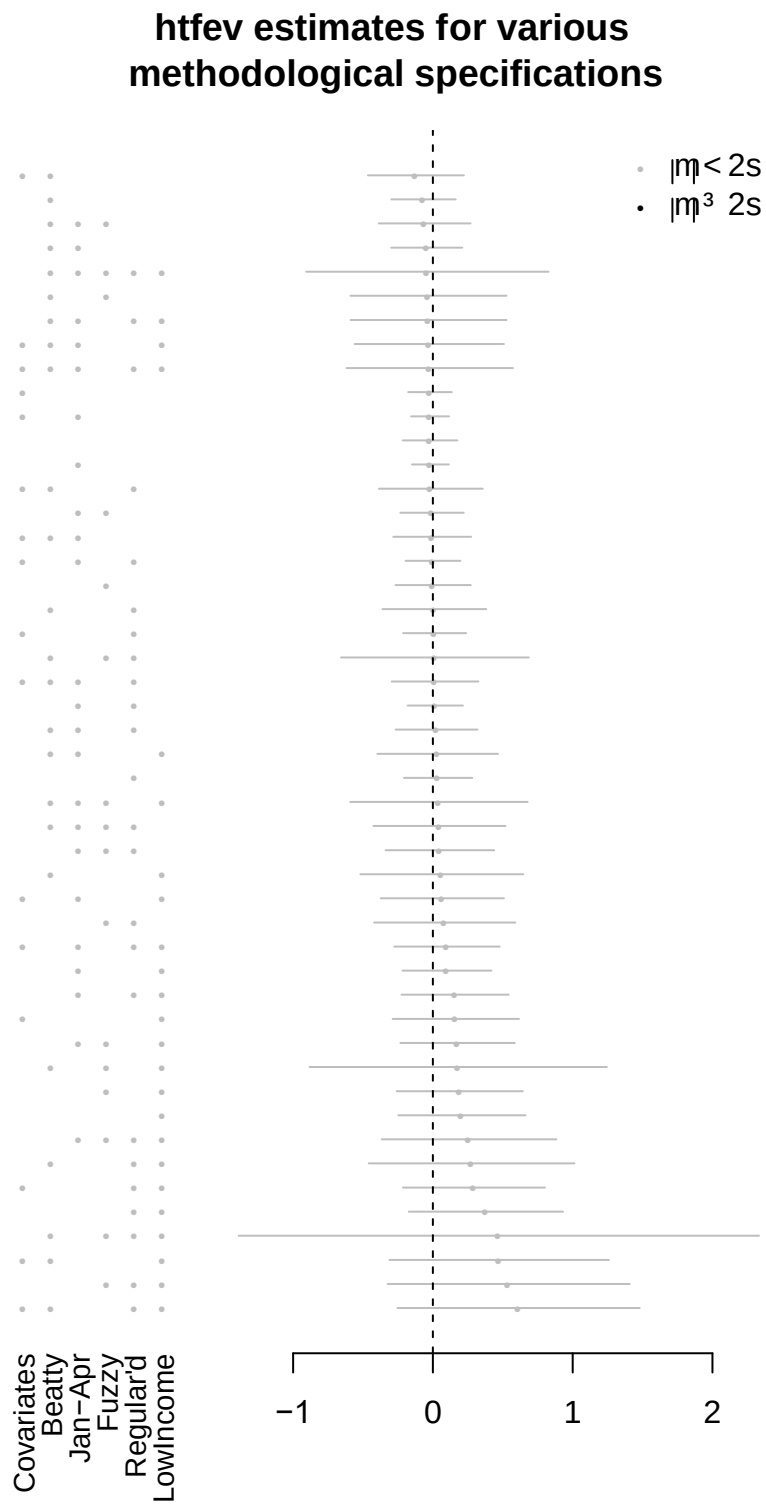


FIGURE S10 Specification curve for *binary* Forced Expiratory Volume

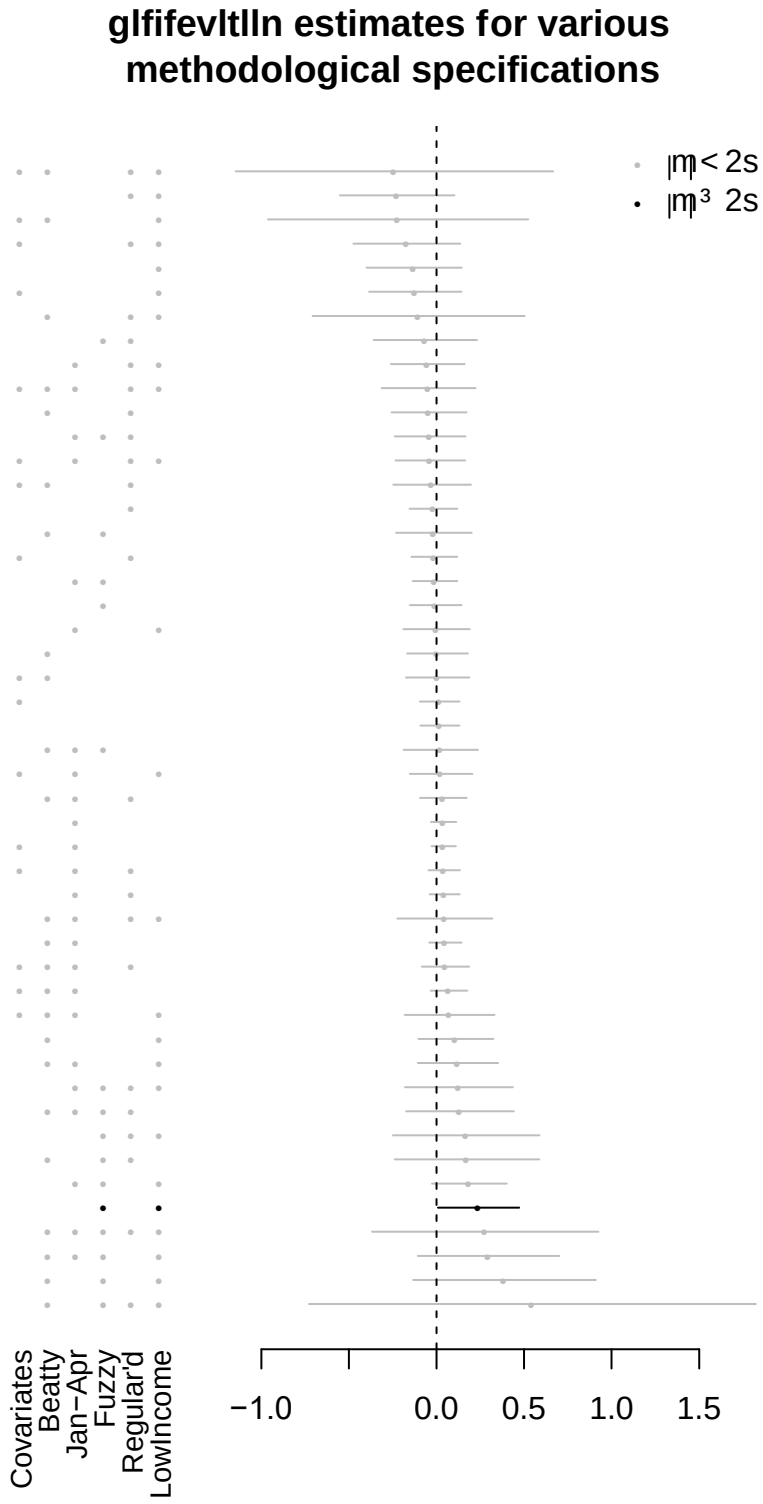


FIGURE S11 Specification curve for self-rated health

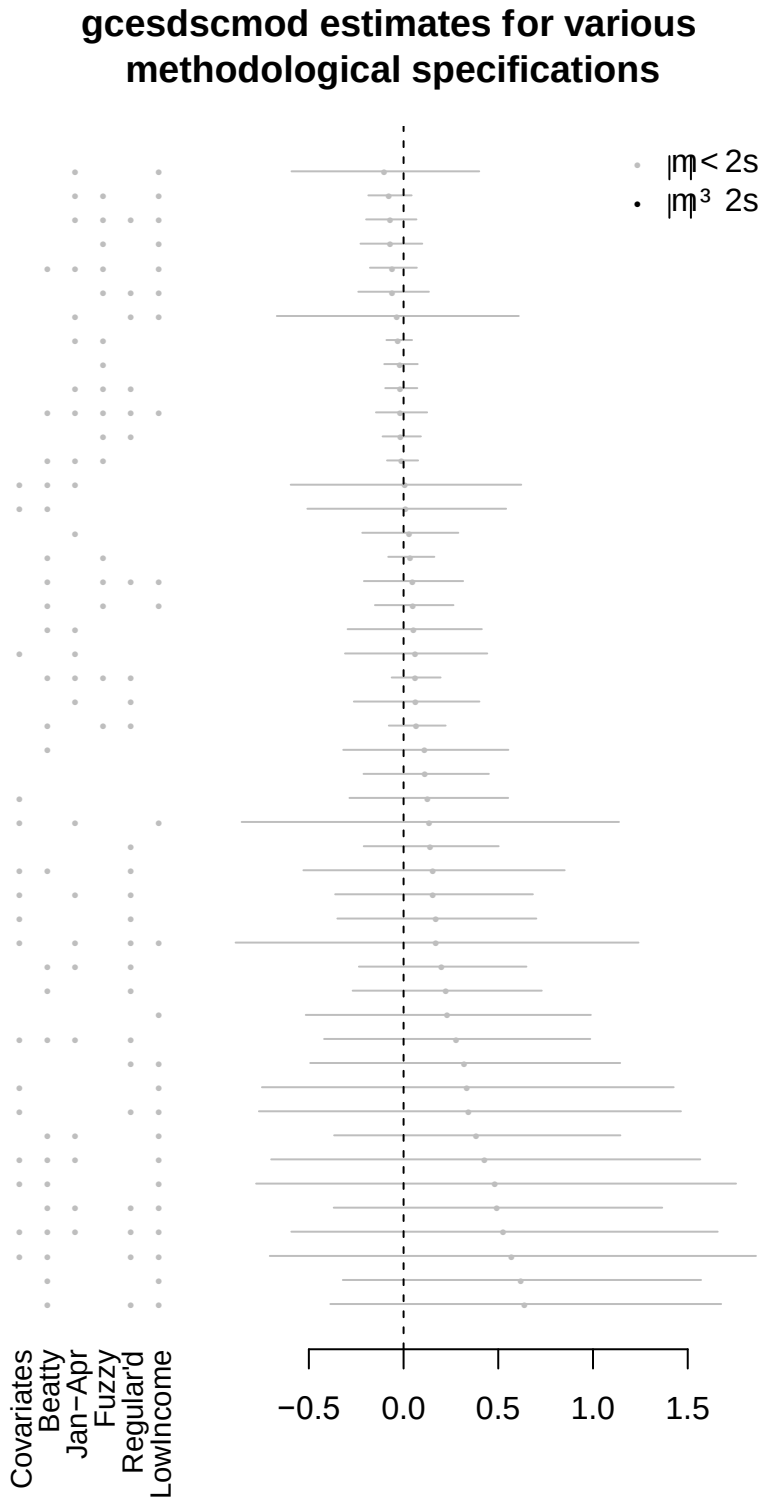


FIGURE S12 Specification curve for depressive symptoms (8-item index)

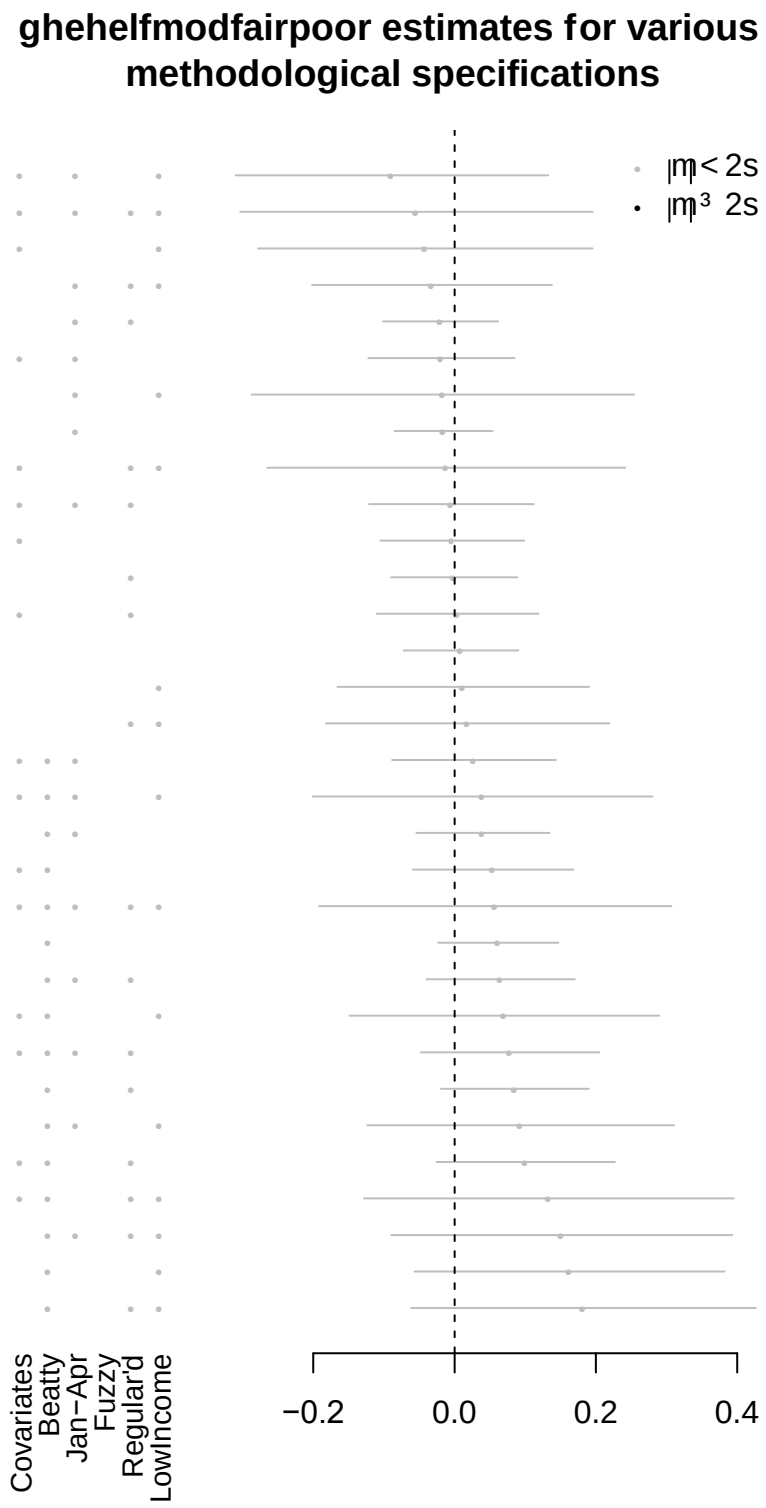
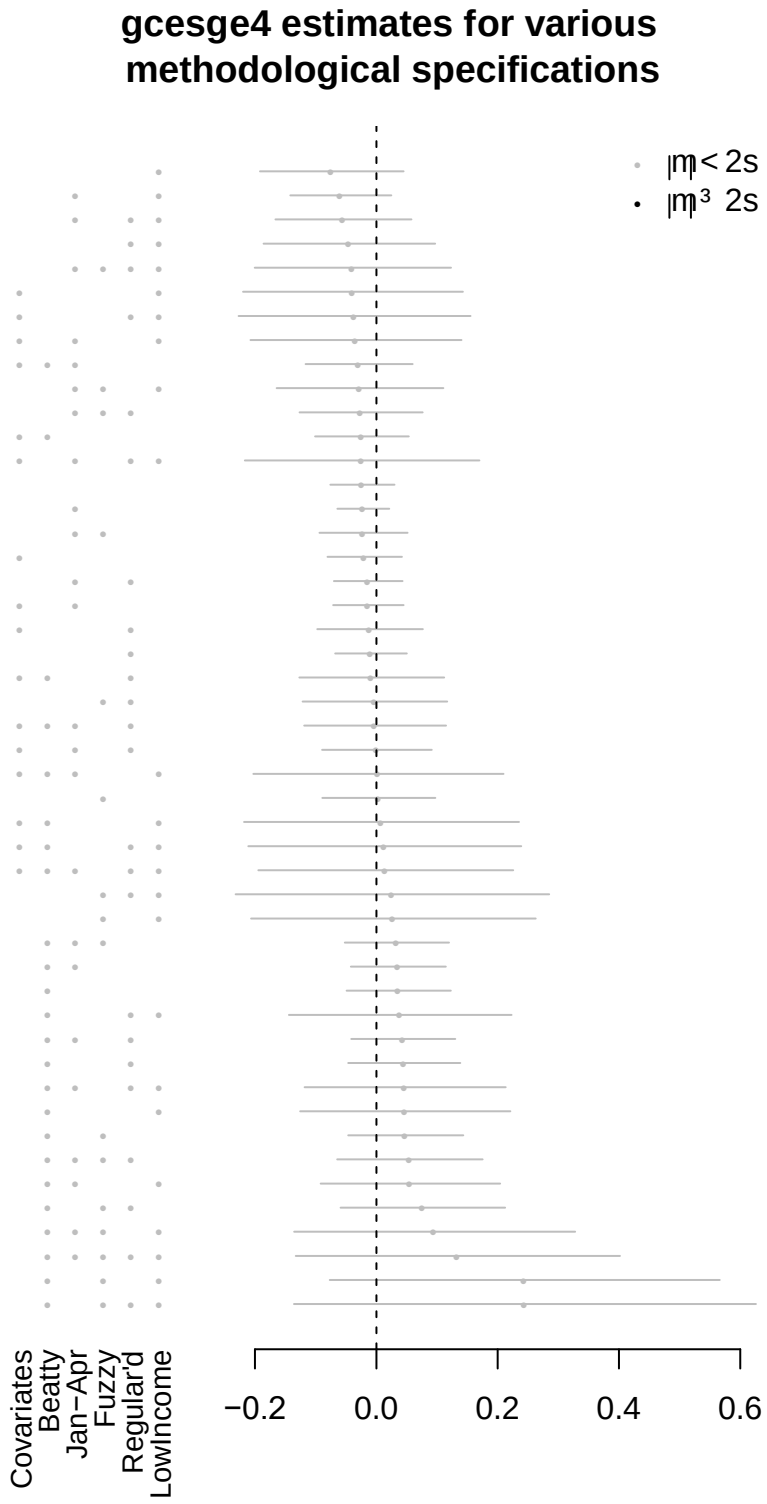


FIGURE S13 Specification curve for binary depressive symptoms (index dichotomised at 4)



2. RD Plots

This appendix complements the analysis provided in the main text by graphically representing the discontinuities using diagrams that plots the relationship between each health outcome and each covariate and age in quarters of birth measured relative to the WFP eligibility age (which is set to 0). These RD plots provide both easy presentation and transparent validation of the design. As reported in the main text, in the absence of the WFP, one should expect these relationships to be continuous around the cut-off age. These plots are constructed by using all the information available in the usable sample, i.e., every observation appearing in the dataset from say people age 50 to people age 80+. Each dot is the sample mean of the cloud of points within evenly spaced bins, which number is computed optimally by a data-driven algorithm aimed at minimising the MSE. [92] The two smooth global polynomial lines are of order four and summarise the relationship on either side of the WFP eligibility date and will capture the presence of a discontinuity.

This graphical analysis is consistent with the more formal analysis presented in the main text. None of the health outcomes show evidence of strong discontinuity at the cut-off. We construct RD plots for each covariates and show that none of the covariates is discontinuous at the WFP receipt age, which validate the design.

FIGURE S14

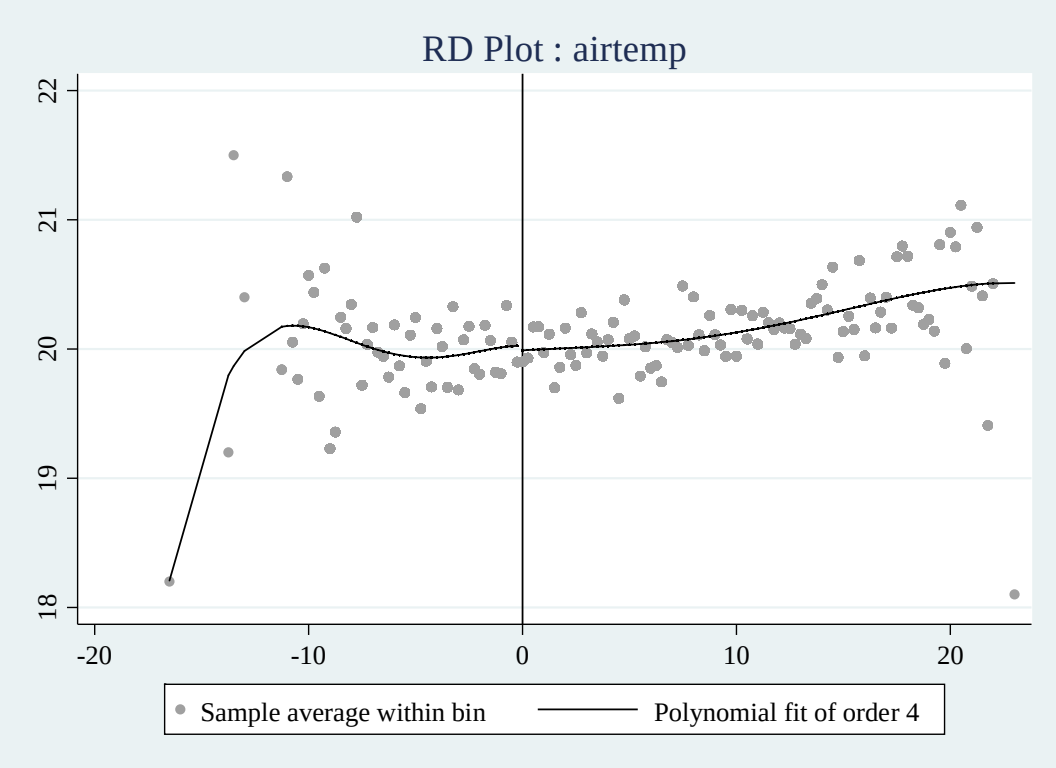


FIGURE S15

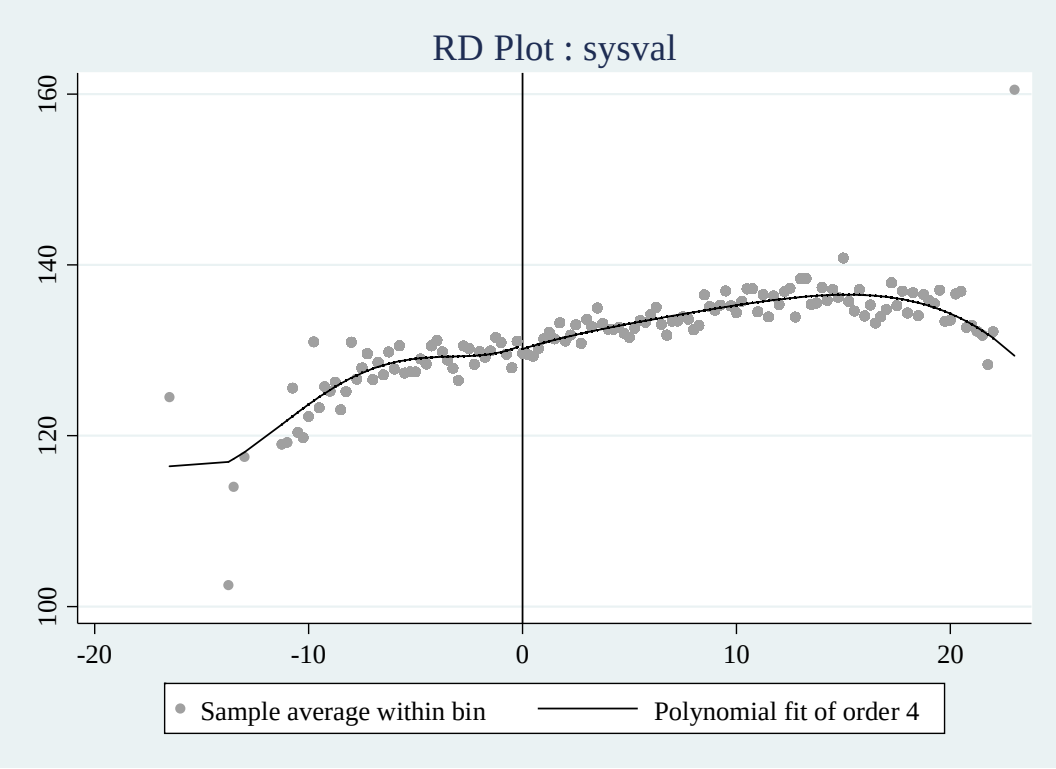


FIGURE S16

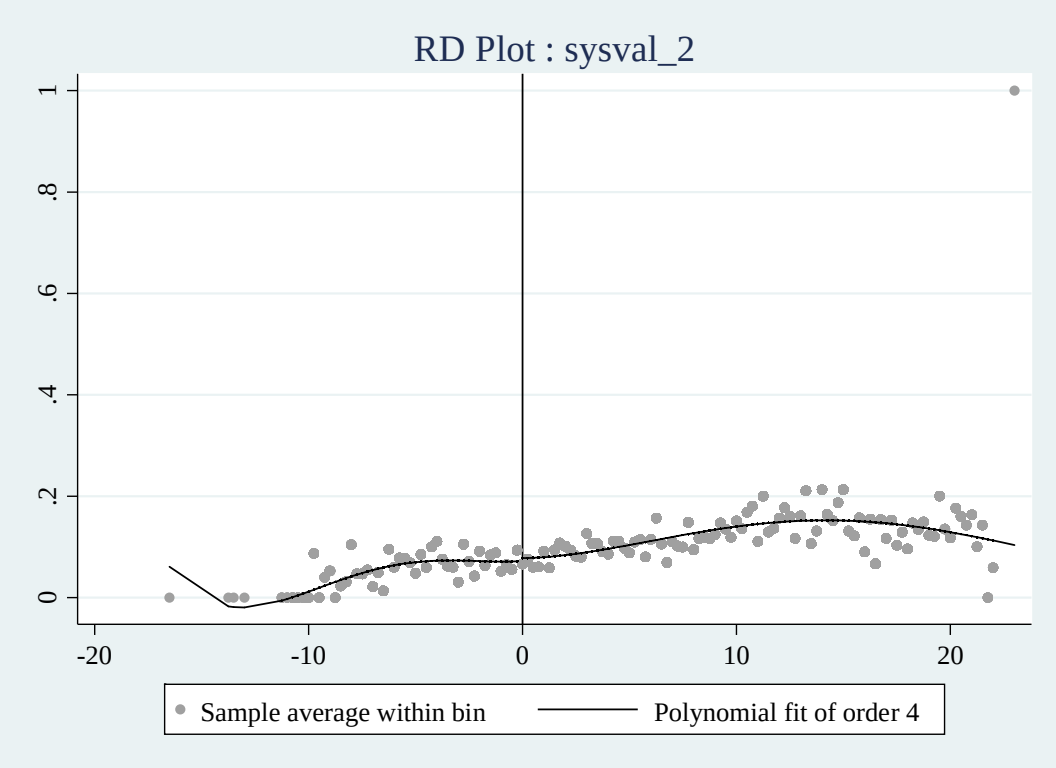


FIGURE S17

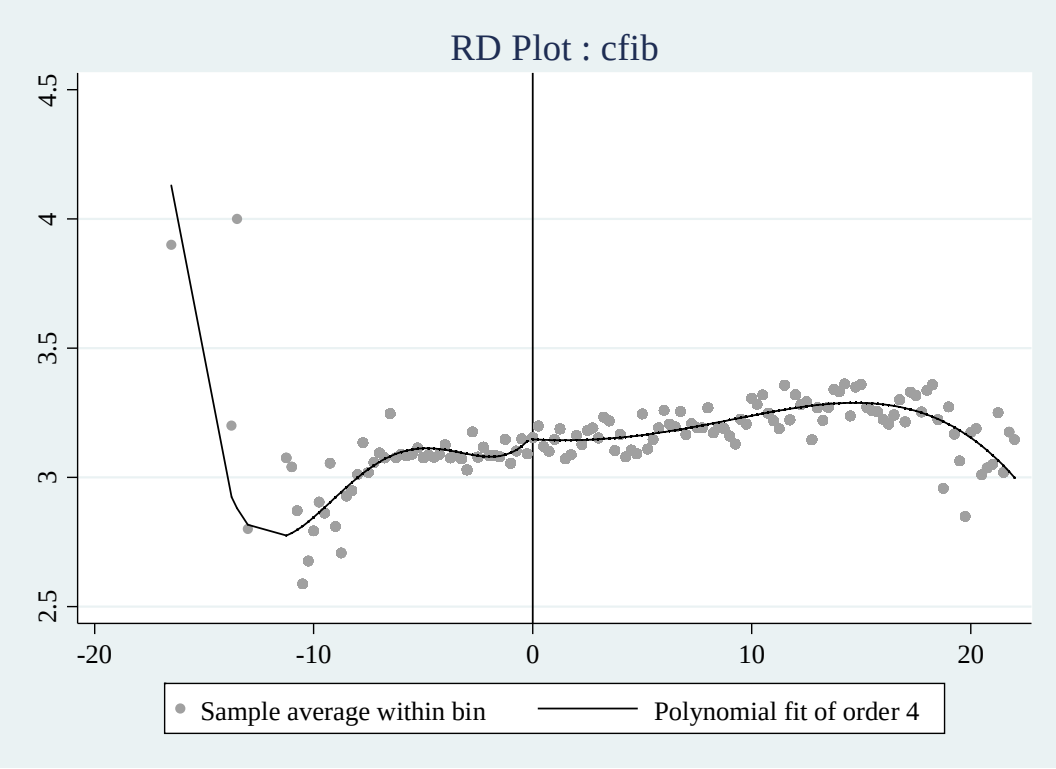


FIGURE S18

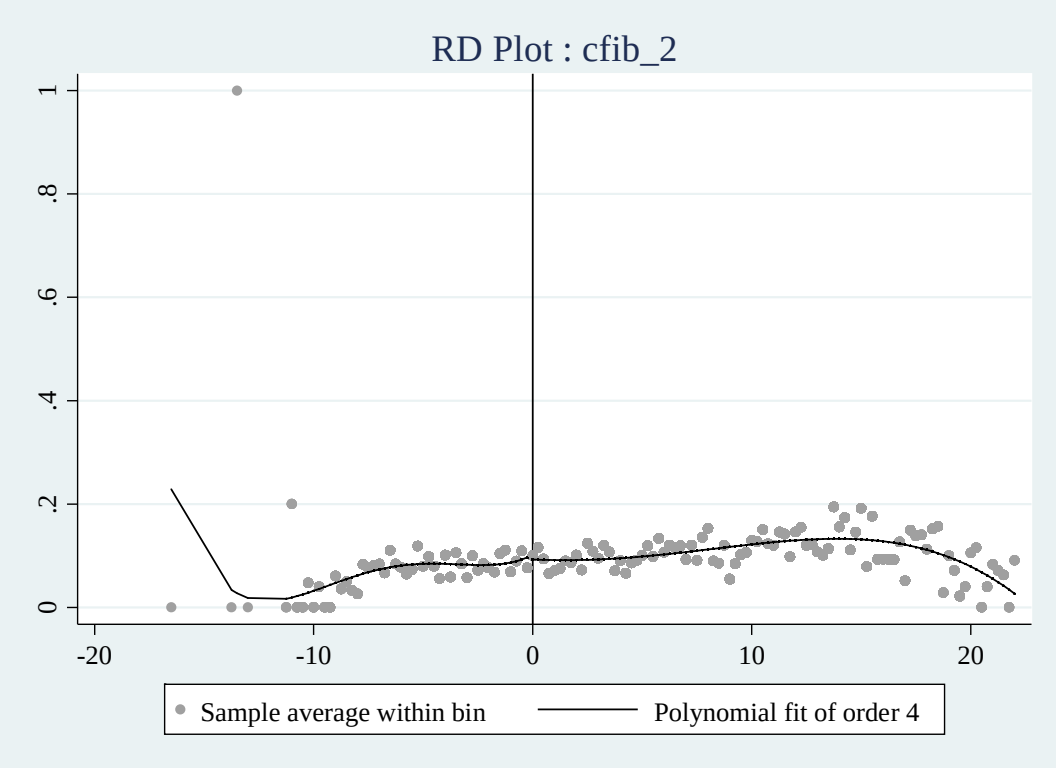


FIGURE S19

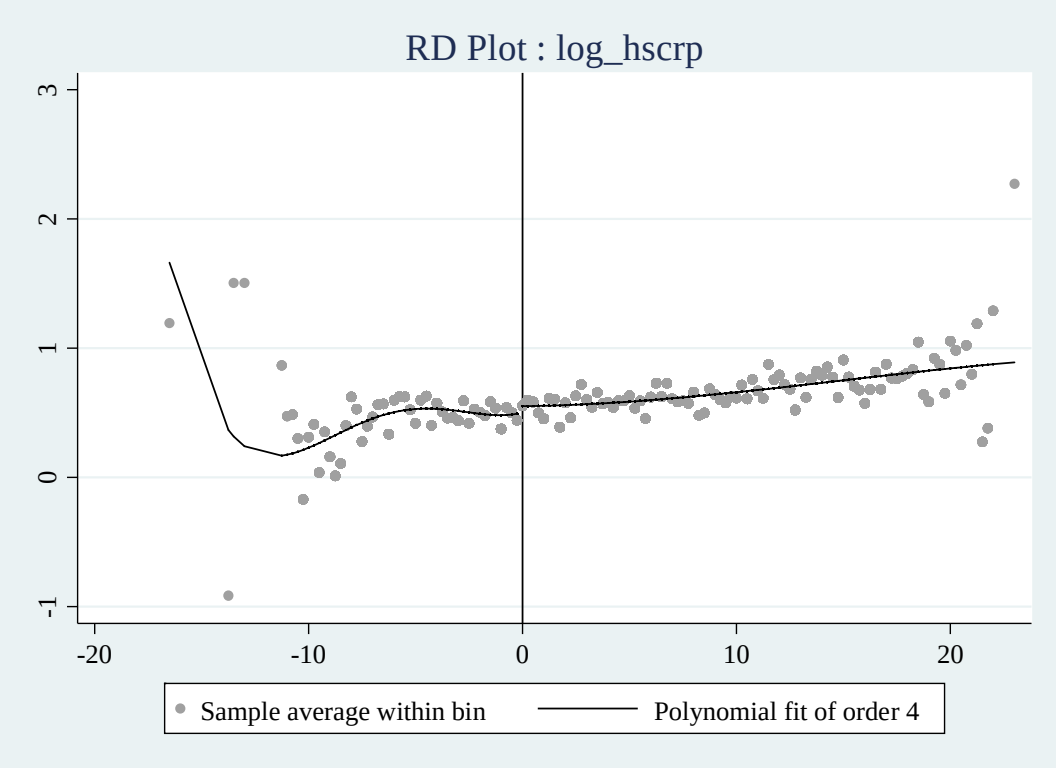


FIGURE S20

Figure S21

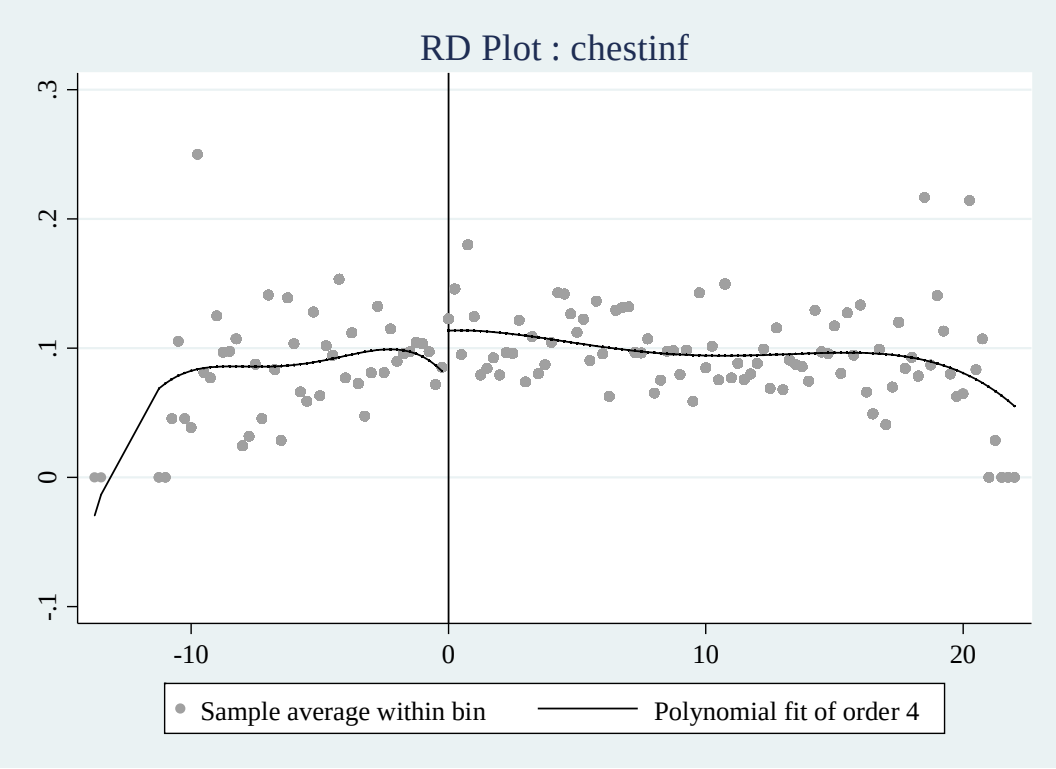


FIGURE S22

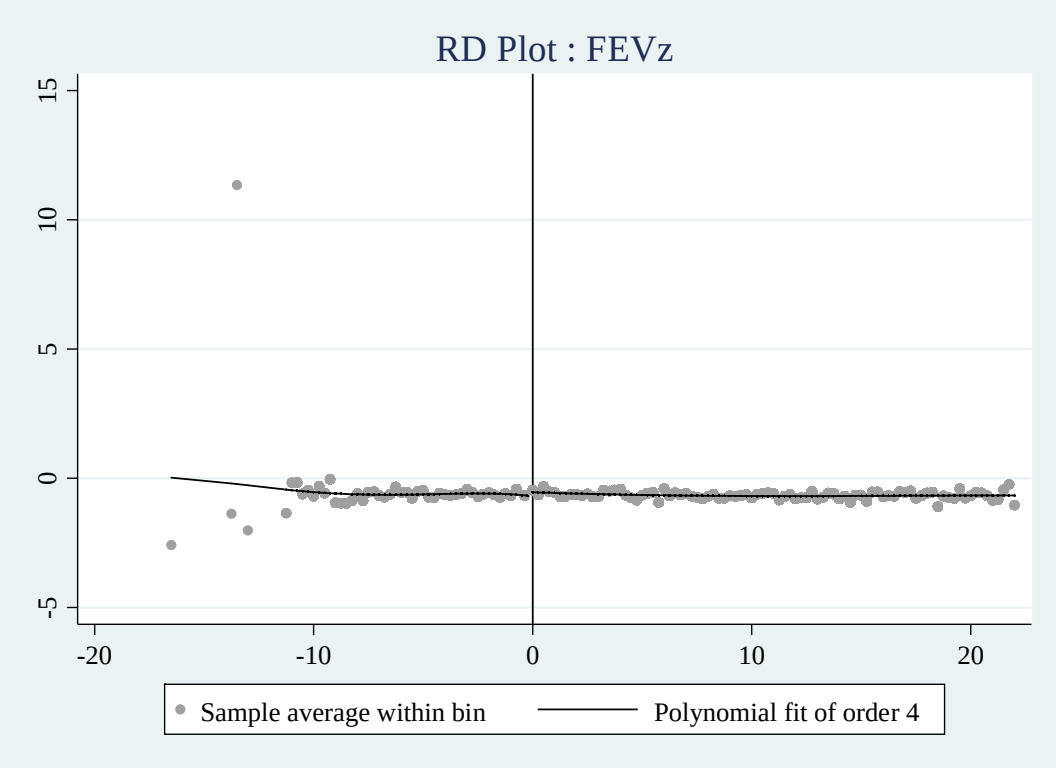


FIGURE S23

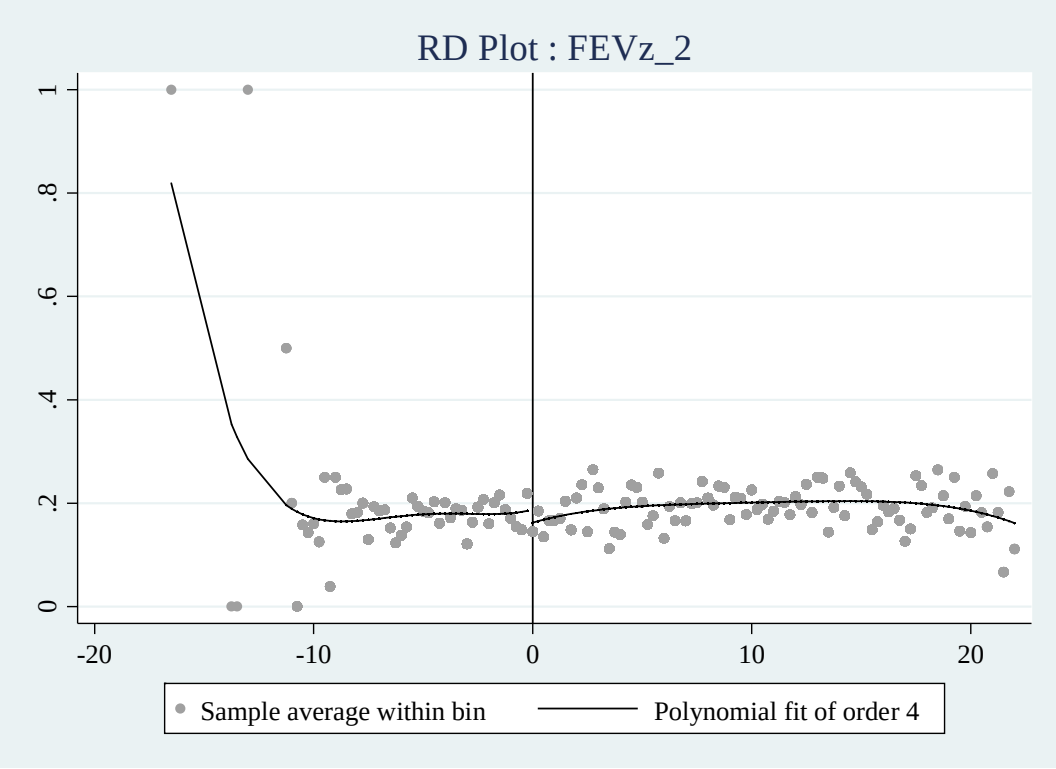


FIGURE S24

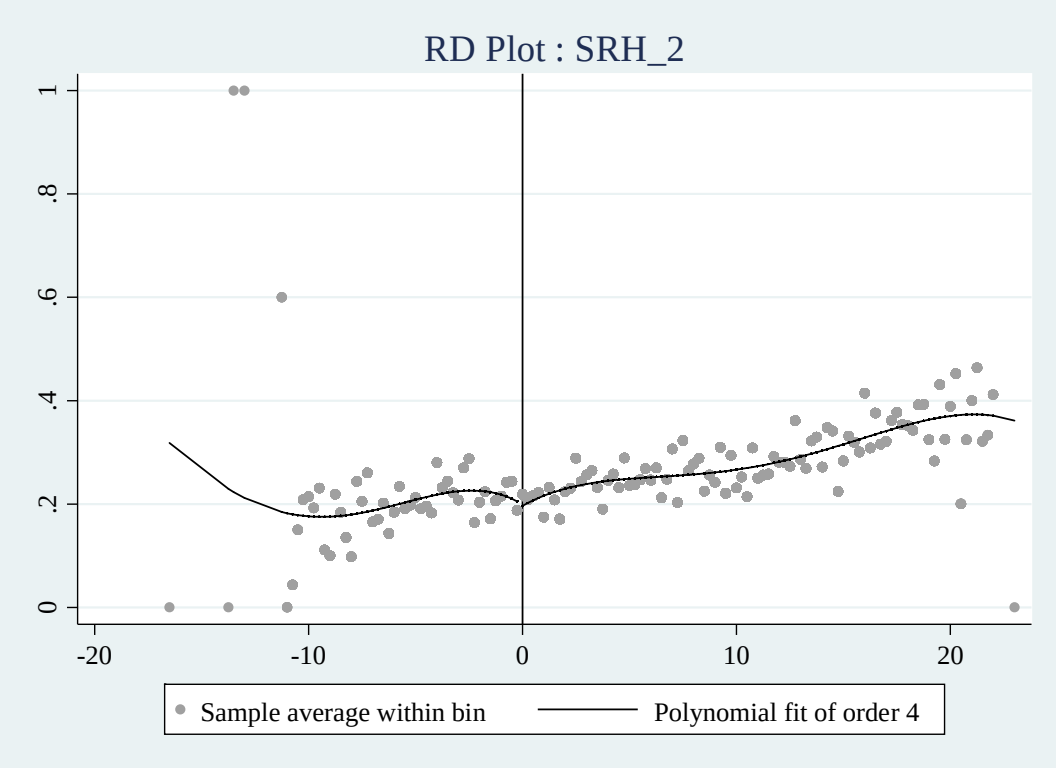


FIGURE S25

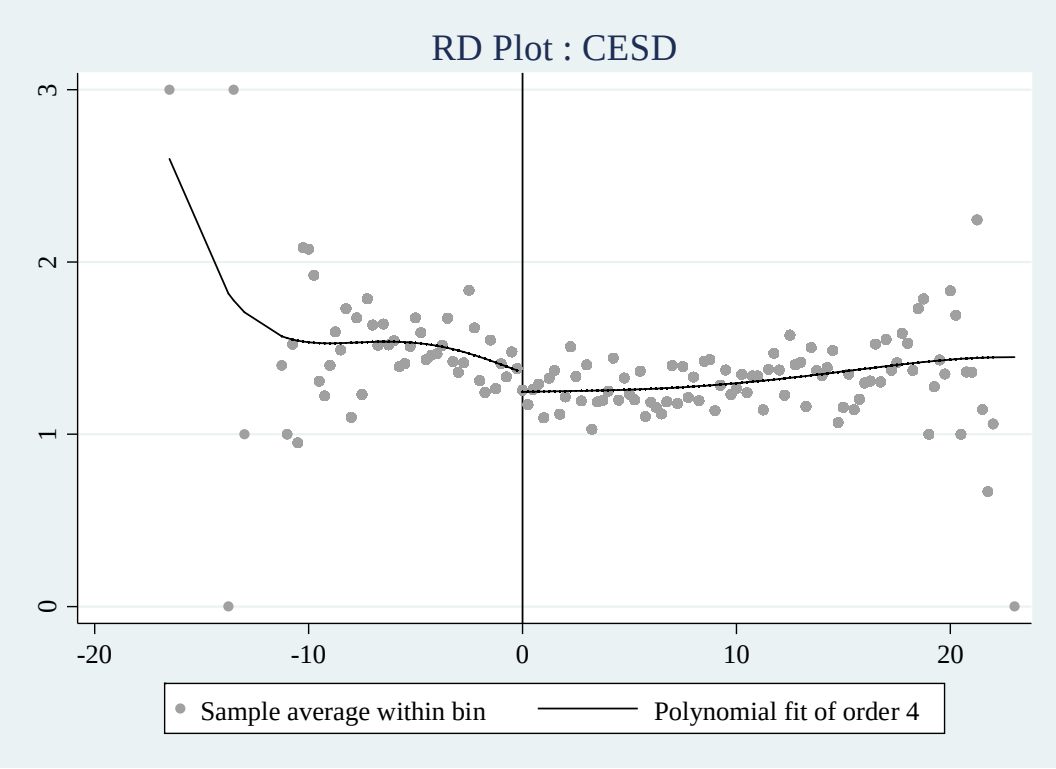


FIGURE S26

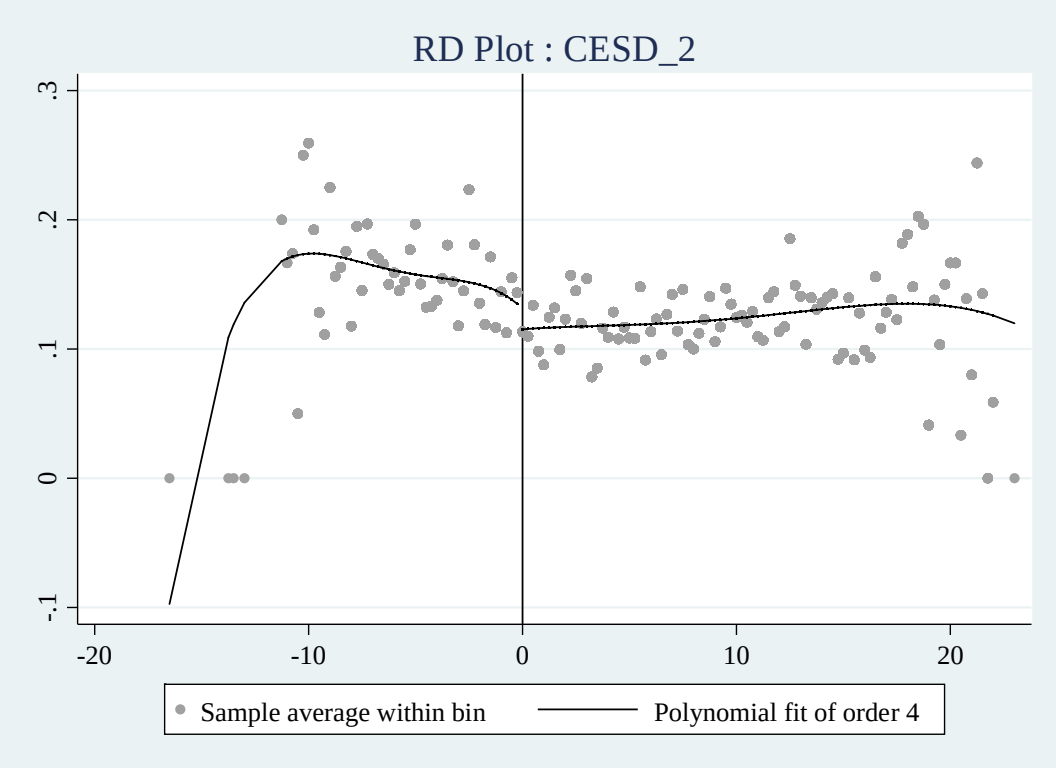


FIGURE S27

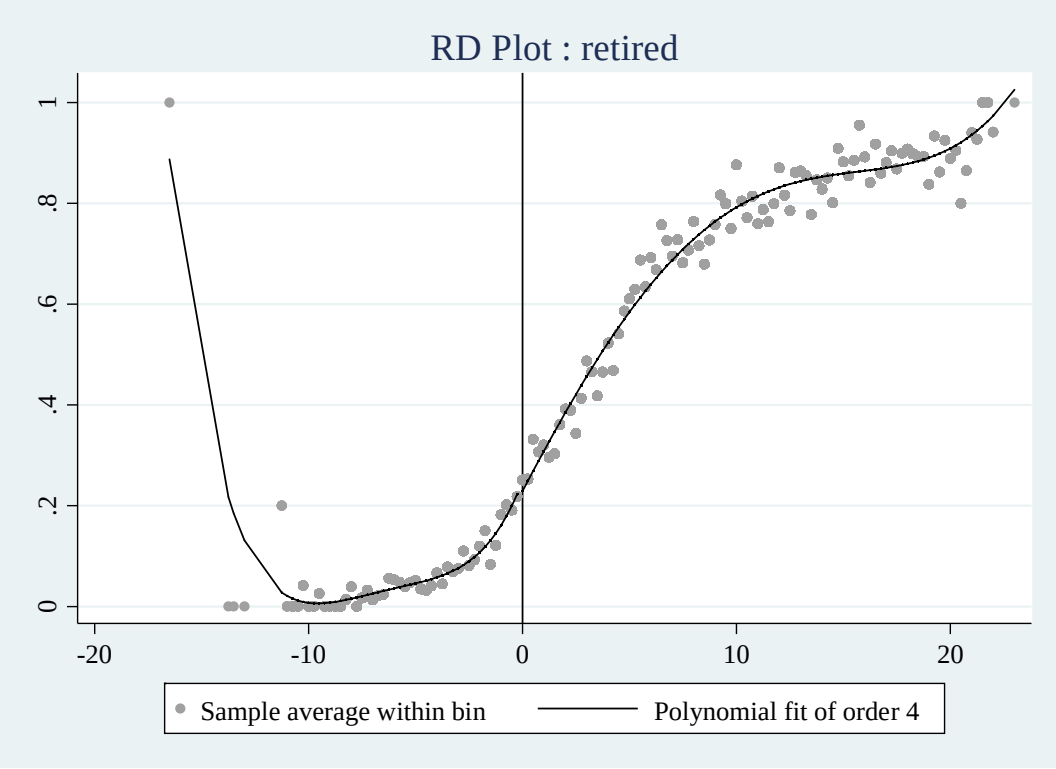


FIGURE S28

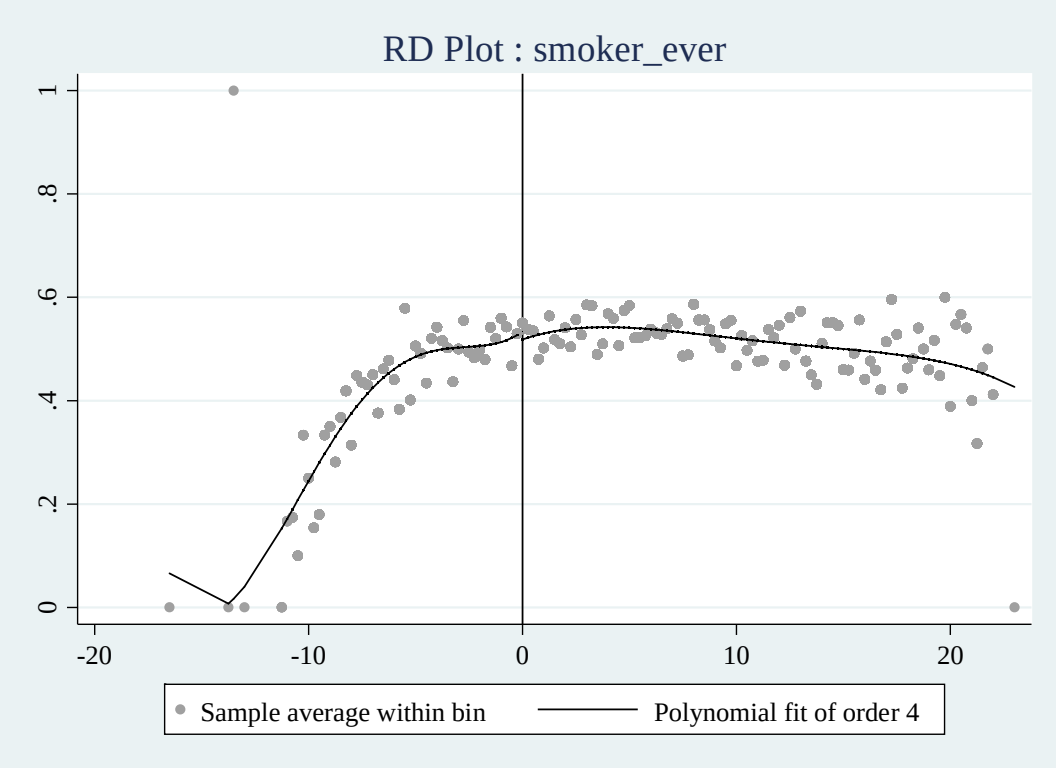


FIGURE S29

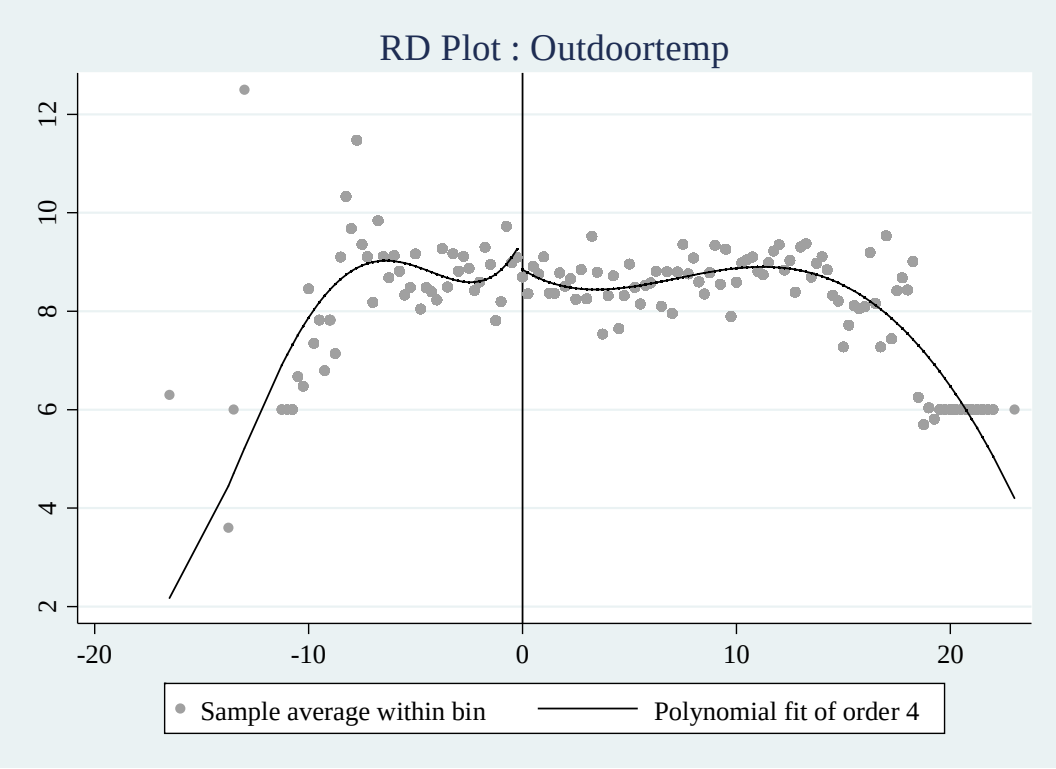
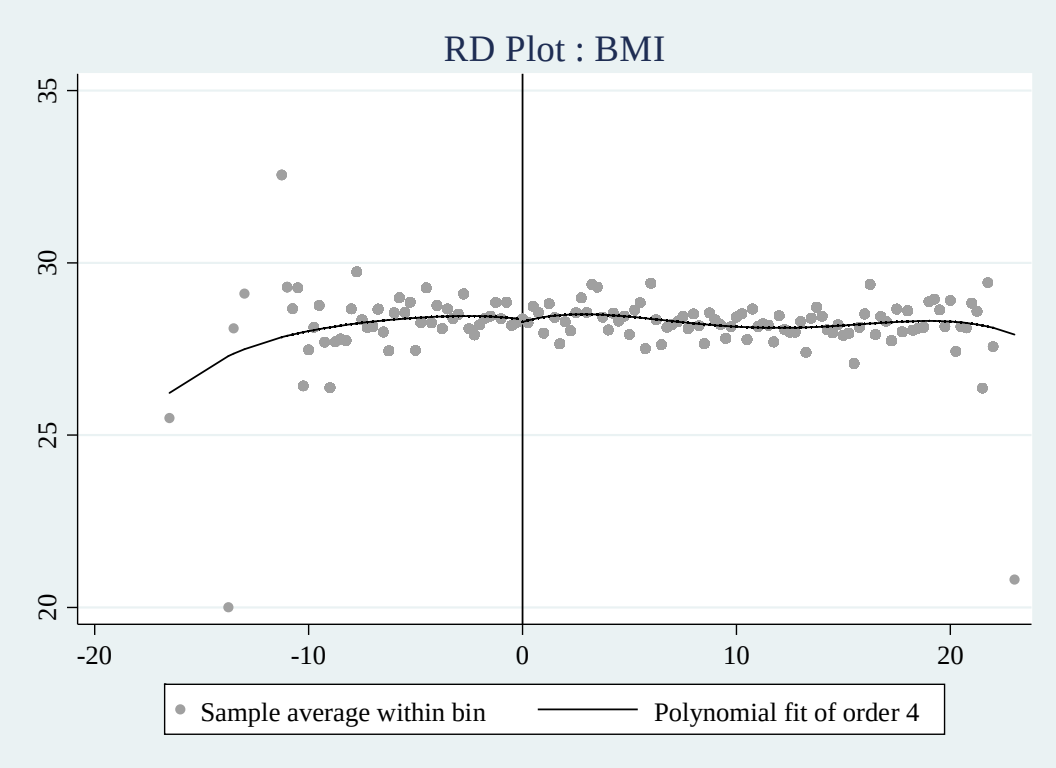


FIGURE S30


RD Plot : BMI
35
30
25
20
-20
-10
0
10
20
Sample average within bin
Polynomial fit of order 4