

# STUDYING THE RELATIONSHIP BETWEEN ANXIETY AND SCHOOL ACHIEVEMENT USING M-QUANTILE RANDOM-EFFECTS REGRESSION WITH SAMPLING WEIGHTS

Francesco Schirripa Spagnolo  
Università di Pisa

Antonella D'Agostino  
Università di Napoli "Parthenope"/ Centro Camilo Dagum

Nicola Salvati  
Università di Pisa/ Centro Camilo Dagum

## MOTIVATION OF THE STUDY

Test anxiety is referred to the specific issue of anxiety related to achieve high test scores in educational careers. Since test scores are crucial for student evaluation at school, test anxiety has become an universal issue in contemporary society. Statistical model that allows to capture the different effects of anxiety that might exist among low- and high-achieving students.

## TARGET POPULATION & DATA

Programme for International Student Assessment (PISA) 2015: Italian students enrolled in public upper secondary schools (ISCED level 3).

Two stage sampling design

- First-stage sampling units are schools having 15-years-old students
- Second-stage sampling units are students within sampled schools

## M-QUANTILE

$$MQ_y(q|x; \psi) = \mathbf{x}_i^T \boldsymbol{\beta}_{\psi q}$$

- $\psi_q$  denotes an asymmetric influence function
- $\boldsymbol{\beta}_{\psi q}$  is the vector of the regression parameters for each  $q$

Estimates of the regression parameters  $\boldsymbol{\beta}_{\psi q}$  can be obtained by minimizing:

$$\sum_{i=1}^n \rho_q(y_i - \mathbf{x}_i^T \boldsymbol{\beta}_{\psi q}).$$

The loss function  $\rho_q(u)$  is usually defined to be the Huber loss function.

## WEIGHTED M-QUANTILE RANDOM-EFFECTS MODEL

Hierarchical structure (students are nested in schools)

M-quantile random-effects model (MQRE)

$$MQ_{y_{ij}}(q|x_{ij}, \gamma_{q,j}; \psi) = \mathbf{x}_{ij}^T \boldsymbol{\beta}_{\psi q} + \mathbf{z}_{ij} \gamma_{q,j},$$

- $\gamma_{q,j}$  is the random effect for cluster  $j$  at the  $q$ th M-quantile and no distributional assumption is imposed on it

i) clusters are selected with inclusion probabilities  $\pi_j, j = 1, \dots, M$

ii) elementary units are sampled with conditional probabilities  $\pi_{ij}, i = 1, \dots, N_j$

Pseudolikelihood approach multilevel modelling

$$\log L_w = \sum_{j \in S} w_j \log \int \left[ \exp \left\{ \sum_{i \in S_j} w_{ij} \log L_{ij}(\boldsymbol{\beta}, \sigma_\gamma^2, \sigma_\varepsilon^2 | \gamma_j) \right\} \right] f(\gamma_j) d\gamma_j$$

- $s$  and  $s_j$  indicate respectively the sampled clusters and the units sampled in the cluster  $j$ .
- $\log L_{ij}(\boldsymbol{\beta}, \sigma_\gamma^2, \sigma_\varepsilon^2 | \gamma_j)$  is the log-likelihood contribution of the level 1 units conditioned on the level 2 random effects
- $w_{ij} = 1/\pi_{ij}; w_j = 1/\pi_j$

Estimating equations Weighted-MQRE

$$\sum_{j \in S} w_j \left[ \mathbf{X}_j^T \mathbf{V}_{qj}^{-1} \mathbf{U}_{qj}^{1/2} \boldsymbol{\psi}_q(\mathbf{r}_{qj}) \right] = \mathbf{0}$$

$$-\frac{1}{2} \sum_{j \in S} w_j \left[ K_{2qj} \text{tr}(\mathbf{V}_{qj}^{-1} \mathbf{Z}_j \mathbf{Z}_j^T) - \boldsymbol{\psi}_q(\mathbf{r}_{qj})^T \mathbf{U}_{qj}^{1/2} \mathbf{V}_{qj}^{-1} \mathbf{Z}_j \mathbf{Z}_j^T \mathbf{V}_{qj}^{-1} \mathbf{U}_{qj}^{1/2} \boldsymbol{\psi}_q(\mathbf{r}_{qj}) \right] = \mathbf{0}$$

$$-\frac{1}{2} \sum_{j \in S} w_j \left[ K_{2qj} \text{tr}(\mathbf{V}_{qj}^{-1} \mathbf{W}_j^{-1}) - \boldsymbol{\psi}_q(\mathbf{r}_{qj})^T \mathbf{U}_{qj}^{1/2} \mathbf{V}_{qj}^{-1} \mathbf{W}_j^{-1} \mathbf{U}_{qj}^{1/2} \mathbf{V}_{qj}^{-1} \boldsymbol{\psi}_q(\mathbf{r}_{qj}) \right] = \mathbf{0}$$

## SIMULATION STUDY

$$y_{ij} = 100 + 2x_{ij} + \gamma_j + \varepsilon_{ij}, \quad i = 1, \dots, N_j, \quad j = 1, \dots, M.$$

- $\gamma \sim 0.9N(0, 1) + 0.1N(9, 20)$
- $\varepsilon \sim 0.9N(0, 3.3) + 0.1N(10, 75)$
- $x_{ij} \sim U(0, 1)$

$$\pi_j = \begin{cases} 0.15 & \gamma < -1, \\ 0.65 & -1 \leq \gamma \leq 1, \\ 0.20 & \gamma > 1 \end{cases} \quad \pi_{ij} = \begin{cases} 0.75 & \varepsilon > 0, \\ 0.25 & \varepsilon \leq 0 \end{cases}$$

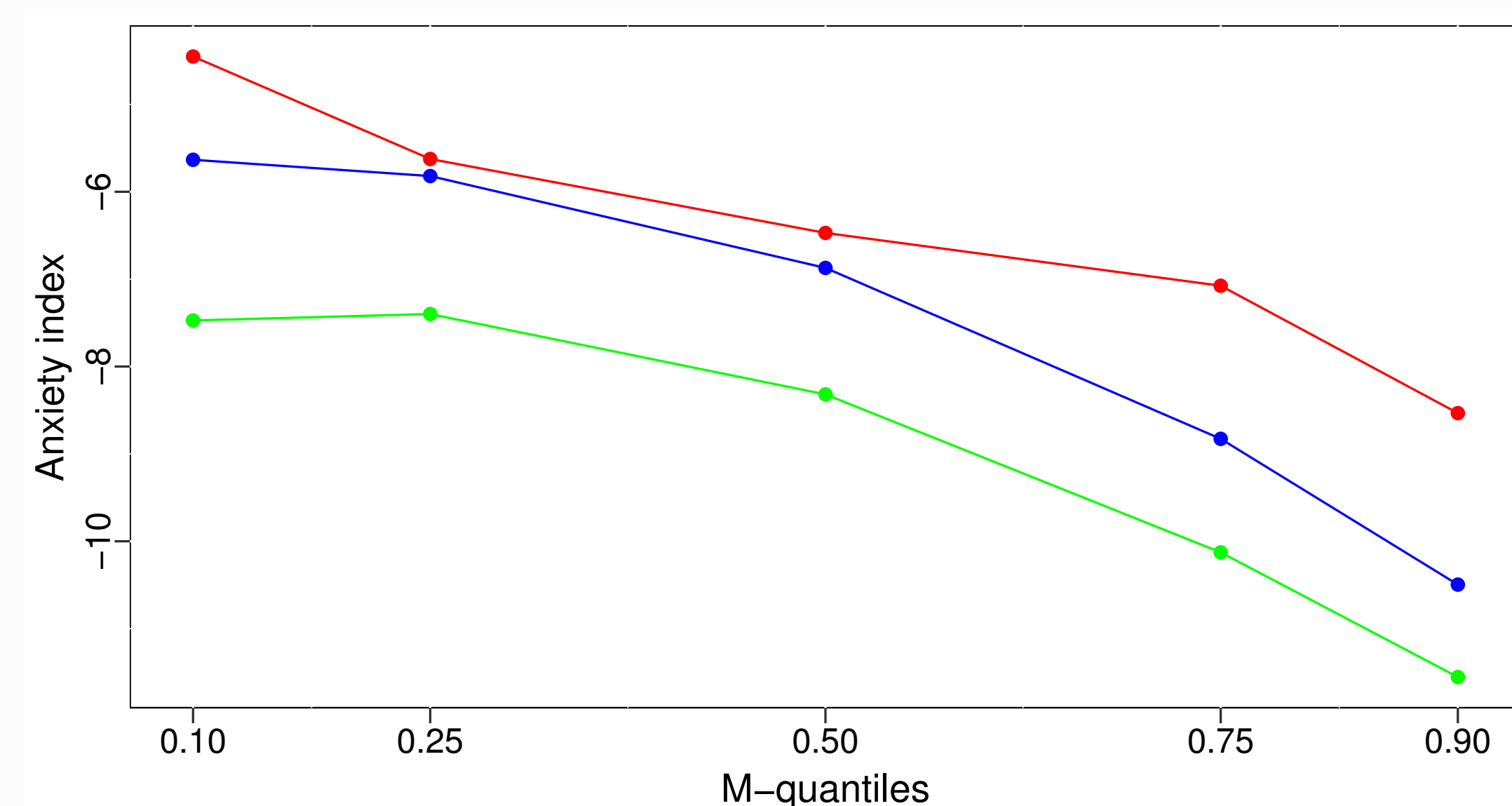
$$n = 1500$$

Method	ARB					
	$q = 0.10$		$q = 0.25$		$q = 0.50$	
	$\hat{\beta}_0$	$\hat{\beta}_1$	$\hat{\beta}_0$	$\hat{\beta}_1$	$\hat{\beta}_0$	$\hat{\beta}_1$
W-MQRE	0.797	-0.025	0.675	-0.013	0.626	-0.011
MQRE	1.406	-0.025	1.258	-0.016	1.118	-0.013

## VARIABLES

- Outcome variables: mathematics, science and reading score
- Anxiety index
- Gender
- Immigration Background
- Grade repetition
- Lack of punctuality
- Ambition of the student
- Index of socio-economic status (ESCS)
- Macro-region (South Italy vs North Italy)
- Type of school (Lyceums vs Other schools)
- School-mean of individual variables

## ANXIETY INDEX



Effect of index of anxiety on mathematics (blu line), science (green line) and reading (red line).

## ESSENTIAL BIBLIOGRAPHY

- OECD (2016). *PISA 2015 Results (Volume III): Students' Well-Being*, PISA. OECD Publishing, Paris.
- Rabe-Hesketh S. and Skrondal A. (2006). Multilevel modelling of complex survey data. *J Roy Stat Soc A Sta*, 169(4): 805–827
- Tzavidis N., Salvati N., Schmid T., Flouri E. and Midouhas E. (2016). Longitudinal analysis of the strengths and difficulties questionnaire scores of the Millennium Cohort Study children in England using M-quantile random-effects regression. *J Roy Stat Soc A Sta*, 179(2): 427–452.