Methodol Comput Appl Probab (2015) 17:999–1014
DOI 10.1007/s11009-014-9432-5

Estimating the Model with Fixed and Random Effects by a Robust Method

Jan Ámos Víšek

Received: 29 January 2014 / Revised: 1 December 2014 /
Accepted: 4 December 2014 / Published online: 19 December 2014

Abstract Regression model with fixed and random effects estimated by modified versions of the *Ordinary Least Squares* (OLS) is a standard tool of panel data analysis. However, it is vulnerable to the bad effects of influential observations (contamination and/or atypical observations). The paper offers robustified versions of the classical methods for this framework. The robustification is carried out by the same idea which was employed when robustifying OLS, it is the idea of weighting down the large order statistics of squared residuals. In contrast to the approach based on the M-estimators this approach does not need the studentization of residuals to reach the *scale-* and *regression-equivariance* of estimator in question. Moreover, such approach is not vulnerable with respect the inliers. The numerical study reveals the reliability of the respective algorithm. The results of this study were collected in a file which is possible to find on web, address is given below. Patterns of these results were included also into the paper. The possibility to reach nearly the full efficiency of estimation - due to the iteratively tailored weight function - in the case when there are no influential points is also demonstrated.

Keywords Linear regression model · The least weighted squares · Fixed and random effects · Numerical simulations

Mathematics Subject Classification (2010) 62J02 · 62F35

1 An Introduction

Atypical observations in a data set can cause misleading conclusions of the regression analysis. That was the reason for building up the robust methods for identifying the true

The paper was written with the support of the Czech Science Foundation project P402/12/G097 DYME - Dynamic Models in Economics.

J. Ámos Víšek (✉)
Faculty of Social Sciences, Institute of Economic Studies, Charles University, Opletalova 26, Prague, The Czech Republic
e-mail: visek@fsv.cuni.cz

underlying model in the various frameworks. One such framework is the regression model with the fixed and random effects and we can meet with proposals of robust estimators of this model based on the idea of *M- estimators*, on the idea of the *Least Median of Squares* (LMS) or on the idea of the *Least Trimmed Squares* (LTS), see e. g. (Bramati and Croux 2007; Dehon et al. 2009; Kott 1989; Rocke 1991; Veradi and Croux 2009) or (Veradi and Wagner 2010).

The employment of M-estimators requires studentization of residuals - to reach the *scale-* and *regression-equivariance* of the estimators in question, see Bickel (1975). Bickel showed that - to reach the goal for the M-etimators - the studentization has to be done by a *scale-equivariant* and *regression-invariant* estimator of standard deviation of disturbances but there are only a few robust estimators of standard deviation of disturbances which are *scale-equivariant* and *regression-invariant*, see Croux and Rousseeuw (1992) and Jurečková and Sen (1984) or Víšek (2010a). Moreover, all these estimators utilize the preliminary robust *scale-* and *regression-equvivariant* estimator of regression model. So, it seems more reasonable to use the estimator which is directly *scale-* and *regression-equvivariant*.

The latter proposals based on LMS and LTS can be (and typicaly are) very sensitive to "inliers", see Víšek (1996b) or Víšek (2011a). The discussion on the sensitivity of highly robust estimators to inliers has been started by Hettmansperger and Sheather (1992). Although their shocking results appeared later to be due to the bad algorithm they used (see again Víšek (1996b) and also Boček and Lachout (1993)), they gave an inspiration for studies of sensitivity of robust methods to the changes of data inside the main cloud of them, i. e. sensitivity to the inliers[1] (Víšek 1996b; 2002; 2006b). The *Least Weighted Squares* (LWS) which employs the idea of smooth decrease of the influence of atypical observations by means of prescribing the weights to the order statistics of the squared residuals rather than to the squared residuals directly, rid of both these problems, i. e. the problem of residual studentization as well as with inliers. Moreover, the advantage of LWS is that it can be adjusted to the level and/or the character of contamination by an adaptive selection of the weights (which is possible to perform iteratively in a reasonable time due to the fast algorithm we have for computing the estimator and the speed of the computational technique).

2 The Model and Estimators

Sometimes we model panel date containing several groups of observed objects (patients in hospital, industries in the national economy, countries in EU) and we believe that the slope parameters are (or are likely to be) the same for all of them but the objects have different "intercepts", i. e. their regression planes are "parallel" but shifted "up or down". An approach via dummy variables would be cumbersome. In fact such a model has to contain the same number of "intercepts" as is the number of objects and hence the design matrix can be pretty "wide" and what is really important these intercepts would be (usually) estimated on the base of small number of observations (i. e. in the text below, if T is small and only n goes to infinity).

These reasons led (approximately) in sixties of the past century to the idea of considering a model in which the different intercepts for different groups of observations are assumed to be realization of one latent variable, see e. g. Hausman (1978) and Hausman and Taylor

[1]The *inliers* are similarly as the *outliers* or the *leverage points* the observations which have the large influence on the estimate in question in the sense that their small shift can cause a large change of values of estimator .

(1981) or Maddala (1971). The ideas which already hinted that such model can be a plausible solution of the problem however appeared even earlier, see e.g. Balestra and Nerlove (1966) or Wallace and Hussain (1969). The situation (we have mentioned a few lines above) was discussed also in a bit different framework e. g. in Telser (1964) and Zellner (1962) or (Zellner 1965). So, let us give the framework we will consider in the form of formalized model.

For any $n,\ T\ \in\ N$ (the set of all positive integers) and $\beta^0\ \in\ R^p$ (p-dimensional Euclidean space) we will study the linear regression model (we employ formalism as in Wooldridge (2006))

$$Y_{it} = X_{it}'\beta^0 + u_i + e_{it}, \quad i = 1, 2, ..., n, \ \ t = 1, 2, ..., T \tag{1}$$

with Y_{it}'s being the response variables, X_{it}'s staying for p-dimensional random explanatory variables, u_i's for the effects and e_{it}'s for the disturbances.

For the case when $cov(X_{itj}, u_i) = 0$ for all $i = 1, 2, ..., n$, $t = 1, 2, ..., T$ and $j = 1, 2, ..., p$, the model (1) will be called the *random effects model*, otherwise we speak about the *fixed effects model*, see e. g. (Judge et al. 1985).

We will consider the residuals in a bit more general form than it is frequently done. So, for any $\beta \in R^p$, $i \in N$ and $t \in \{1, 2, ..., T\}$ we will consider the residual of the (i, t)-th observation given as

$$r_{it}(\beta) = Y_{it} - X_{it}' \cdot \beta \tag{2}$$

and we denote by $r^2_{(\ell)}(\beta)$ the ℓ-th order statistic among the squared residuals $r^2_{it}(\beta)$, $i = 1, 2, ..., n$, $t = 1, 2, ..., T$, i. e. we have

$$r^2_{(1)}(\beta) \le r^2_{(2)}(\beta) \le ... \le r^2_{(n\cdot T)}(\beta).$$

Finally, let's recall the classical *Ordinary Least Squares* and their robust version, the *Least Weighted Squares*.

Definition 1 Let $w_\ell \in [0, 1]$, $\ell = 1, 2, ..., n \cdot T$ be weights. The estimators

$$\hat\beta^{(OLS,n,T)} = \underset{\beta \in R^p}{arg\,min} \sum_{i=1}^{n}\sum_{t=1}^{T} r^2_{it}(\beta) \ \text{ and } \ \hat\beta^{(LWS,n,T,w)} = \underset{\beta \in R^p}{arg\,min} \sum_{\ell=1}^{n\cdot T} w_\ell\, r^2_{(\ell)}(\beta)$$

are called the *Ordinary Least Squares* (*OLS*) and the *Least Weighted Squares* (*LWS*) estimator, respectively (Víšek 2000).

Denoting the empirical distribution function of the absolute values of residuals $r_{it}(\beta)$'s by $F^{(n)}_\beta(r)$, a straightforward derivation shows that $\hat\beta^{(LWS,n,T,w)}$ is one of solutions of the *normal equations*

$$\sum_{i=1}^{n}\sum_{t=1}^{T} w\left(F^{(n\cdot T)}_\beta(|r_{it}(\beta)|)\right) X_{it}(Y_{it} - X_{it}' \cdot \beta) = 0. \tag{3}$$

Then the generalization of the classical Kolmogorov-Smirnov result on uniform convergence of empirical d. f. for the *regression model framework* (see Víšek (2011b)[2])

$$\sup_{\beta\in R^p} \sup_{-\infty<r<\infty} \sqrt{n} \left| F^{(n\cdot T)}_\beta(r) - (n\cdot T)^{-1} \sum_{i=1}^{n}\sum_{t=1}^{T} F^{(it)}_\beta(r) \right| = O_p(1)$$

[2]The main technical tool for the proof of this generalization was the Skorohod embedding into Wiener process, see Portnoy (1983).

allows to prove the consistency of $\hat{\beta}^{(LWS,n,T,w)}$ under the following conditions:

Condition *C* 1. *The sequence* $\left\{\left\{\left(X_{it}', e_{it}\right)'\right\}_{t=1}^{T}\right\}_{i=1}^{\infty}$ *is a sequence independent and identically distributed* $(p+1)$*-dimensional random variables (r.v.'s) distributed according to distribution functions (d.f.)* $F_{X,e}(x,r) = F_X(x) \cdot F_e(r)$ *with* $\mathbb{E}e = 0$ *and* $var(e) = \sigma^2$. *Moreover,* $F_e(r)$ *is absolutely continuous with bounded density* $f_e(r)$. *Further, there is* $q > 1$ *so that* $\mathbb{E}\,\|X_1\|^{2q} < \infty$. *Finally,* $\{u_i\}_{i=1}^{\infty}$ *is a sequence of independent and identically distributed r. v.'s, independent from the sequence* $\left\{\left\{\left(X_{it}', e_{it}\right)'\right\}_{t=1}^{T}\right\}_{i=1}^{\infty}$, *with d.f.* $F_u(s)$ *with finite variance* σ_u^2.

The weights are usually generated as $w_\ell = w\left(\frac{\ell-1}{n\cdot T}\right)$ *and we typically assume:*

Condition *C* 2. *The weight function* $w(u)$ *is continuous, nonincreasing and* $w : [0,1] \to [0,1]$ *with* $w(0) = 1$.

Condition *C* 3. *Put* $F_\beta^{(it)}(r) = P\left(|r_{it}(\beta)| \le r\right)$ *(remember (2)). For any* $n \in N$ *there is the only solution of*

$$\mathbb{E}\left\{\sum_{i=1}^{n}\sum_{t=1}^{T}\left[w\left(F_\beta^{(nT)}(|r_{it}(\beta)|)\right)X_{it}\left(Y_{it} - X_{it}'\beta\right)\right]\right\} = 0 \tag{4}$$

namely β^0.

Theorem 1 *Let* ***Conditions*** $C\,1$, $C\,2$ *and* $C\,3$ *be fulfilled. Then any sequence* $\left\{\hat{\beta}^{(LWS,n,T,w)}\right\}_{n=1}^{\infty}$ *of the solutions of normal* Eq. 3 *is weakly consistent.*

The proof is a direct reformulation of the result (Víšek 2011a) from the *cross-sectional-data framework* to the *panel-data framework* (possibly with fixed or random effects). The result can be - under a bit stronger conditions - strengthen to $\sqrt{n}$-consistency of $\hat{\beta}^{(LWS,n,T,w)}$, (Víšek 2010a).

Remark 1 Notice that we need no condition for T. It can be fixed or it can go to ∞, simultaneously with n (Tables 1, 2, 3, 4).

3 Fixed and Random Effects Estimation

Let's recall the classical solutions (assuming normality of disturbances). For *model with random effects* we have (see Eq. 1)

$$Y_{it} = X_{it}'\beta^0 + v_{it}, \quad i = 1,2,...,n, \quad t = 1,2,...,T \tag{5}$$

where $v_{it} = u_i + e_{it}$, $\mathbb{E}v_{it} = 0$, $\mathbb{E}v_{it}^2 = \sigma_u^2 + \sigma_e^2$, $\text{cov}(X_{itj}, v_{it}) = 0$ *and* $\mathbb{E}[v_{it}, v_{is}] = \text{var}(u_i) = \sigma_u^2$. It implies that $\hat{\beta}^{(OLS,n,T)}$ is not efficient due to the correlation between

Table 1 The disturbances are independent from explanatory variables but the effects are correlated with them. There is no contamination but we assume some hence we take measures against it

True coeffs β^0	**3**	**1**	**2**	**−4**	**5**
Contamination level is assumed high, $h = 0.2 \cdot n$ and $g = 0.8 \cdot n$.					
$\hat{\beta}^{OLS}_{(\text{var}(\hat{\beta}^{OLS}))}$	$3.00_{(0.011)}$	$.00_{(0.002)}$	$2.00_{(0.002)}$	$-4.00_{(0.002)}$	$5.00_{(0.002)}$
$\hat{\beta}^{FE}_{(\text{var}(\hat{\beta}^{FE}))}$	$0.00_{(0.000)}$	$1.00_{(0.001)}$	$2.00_{(0.001)}$	$-4.00_{(0.001)}$	$5.00_{(0.001)}$
$\hat{\beta}^{RE}_{(\text{var}(\hat{\beta}^{RE}))}$	$3.00_{(0.011)}$	$1.00_{(0.001)}$	$2.00_{(0.001)}$	$-4.00_{(0.001)}$	$5.00_{(0.001)}$
$\hat{\beta}^{LWS}_{(\text{var}(\hat{\beta}^{LWS}))}$	$3.00_{(0.015)}$	$1.00_{(0.006)}$	$2.00_{(0.006)}$	$-4.00_{(0.006)}$	$5.00_{(0.006)}$
$\hat{\beta}^{FWE}_{(\text{var}(\hat{\beta}^{FWE}))}$	$0.00_{(0.000)}$	$1.00_{(0.003)}$	$2.00_{(0.003)}$	$-3.99_{(0.003)}$	$5.00_{(0.003)}$
$\hat{\beta}^{RWE}_{(\text{var}(\hat{\beta}^{RWE}))}$	$3.00_{(0.106)}$	$1.00_{(0.003)}$	$2.00_{(0.003)}$	$-4.00_{(0.003)}$	$5.00_{(0.003)}$
From previous lines of Table 1 it follows that the contamination is not high (as the all estimates have nearly the same value), hence we accommodated the weight function a bit and we put $h = 0.6 \cdot n$ and $g = 0.9 \cdot n$. (In the next two part of Table 1 the first three lines would be the same as in the previous part, hence they were omitted.)					
$\hat{\beta}^{LWS}_{(\text{var}(\hat{\beta}^{LWS}))}$	$3.00_{(0.013)}$	$1.00_{(0.004)}$	$2.00_{(0.004)}$	$-4.00_{(0.005)}$	$5.00_{(0.004)}$
$\hat{\beta}^{FWE}_{(\text{var}(\hat{\beta}^{FWE}))}$	$0.00_{(0.000)}$	$1.00_{(0.002)}$	$1.99_{(0.002)}$	$-4.00_{(0.003)}$	$5.00_{(0.002)}$
$\hat{\beta}^{RWE}_{(\text{var}(\hat{\beta}^{RWE}))}$	$2.99_{(0.579)}$	$1.00_{(0.002)}$	$2.00_{(0.002)}$	$-4.00_{(0.003)}$	$5.00_{(0.002)}$
We learned that contamination is probably very low, if any (as again the values of all estimates are the same and differences of MSE are small), and hence we further accommodated the weight function putting $h = 0.8 \cdot n$ and $g = 0.98 \cdot n$.					
$\hat{\beta}^{LWS}_{(\text{var}(\hat{\beta}^{LWS}))}$	$2.99_{(0.011)}$	$1.00_{(0.002)}$	$2.00_{(0.003)}$	$-4.00_{(0.002)}$	$5.00_{(0.002)}$
$\hat{\beta}^{FWE}_{(\text{var}(\hat{\beta}^{FWE}))}$	$0.00_{(0.000)}$	$1.00_{(0.001)}$	$2.00_{(0.001)}$	$-4.00_{(0.001)}$	$5.00_{(0.001)}$
$\hat{\beta}^{RWE}_{(\text{var}(\hat{\beta}^{RWE}))}$	$2.99_{(0.014)}$	$1.00_{(0.001)}$	$2.00_{(0.001)}$	$-4.00_{(0.001)}$	$5.00_{(0.001)}$

disturbances. The estimation can be improved either by utilizing the *Generalized Least Squares* or - equivalently - by considering slightly modified data

$$\tilde{Y}_{it} = Y_{it} - \lambda \overline{Y}_i \text{ and } \tilde{X}_{it} = X_{it} - \lambda \overline{X}_i, \text{ with } \lambda = 1 - \sigma_e^2 \cdot \left(\sigma_e^2 + T \cdot \sigma_u^2\right)^{-1}, \quad (6)$$

(where $\overline{Y}_i = T^{-1} \sum_{t=1}^{T} Y_{it}$ and $\overline{X}_i = T^{-1} \sum_{t=1}^{T} X_{it}$) and applying $\hat{\beta}^{(OLS,n,T)}$ on $\tilde{Y}_{it}$'s and $\tilde{X}_{it}$'s, see e. g. Wooldridge (2006). The variances σ_e^2 and σ_u^2 can be estimated employing $r_{it}\left(\hat{\beta}^{(OLS,n,T)}\right)$ by the classical estimators for the variance of disturbances, say $\hat{\sigma}_v^2$ and $\hat{\sigma}_u^2$, taking into account that $\sigma_e^2 = \sigma_v^2 - \sigma_u^2$ and then we can use $\hat{\lambda}$ instead of λ. Applying then *OLS* on the transformed data we obtain *RE-estimate* (which is below in tables denoted as $\hat{\beta}^{RE}$).

Assuming *framework with fixed effects*, we have in model (6) $\mathbb{E} v_{it} = 0$ with $\text{cov}(X_{itj}, v_{it}) \neq 0$. It implies the inconsistency and biasedness of $\hat{\beta}^{(OLS,n,T)}$. Among all possible remedies it seems the most attractive *fixed-effect-estimation* (or alternatively called the *within-the-groups transformation*). It considers data

$$\tilde{Y}_{it} = Y_{it} - \overline{Y}_i \text{ and } \tilde{X}_{it} = X_{it} - \overline{X}_i, \text{ with } \overline{Y}_i = \frac{1}{T}\sum_{t=1}^{T} Y_{it} \text{ and } \overline{X}_i = \frac{1}{T}\sum_{t=1}^{T} X_{it} \quad (7)$$

Table 2 Both, the disturbances and the effects, are independent from explanatory variables. Contamination is created by outliers, i. e. explanatory variables are not changed but values of contaminated response variables are equal to −2.5 multiple of original values of response variable

True coeffs β^0	3	1	2	−4	5
Contamination level is equal to 0.5 %.					
$\hat\beta^{OLS}_{\left(\text{var}(\hat\beta^{OLS})\right)}$	$2.95_{(0.016)}$	$0.93_{(0.040)}$	$1.87_{(0.042)}$	$-3.73_{(0.060)}$	$4.63_{(0.076)}$
$\hat\beta^{FE}_{\left(\text{var}(\hat\beta^{FE})\right)}$	$0.00_{(0.000)}$	$0.93_{(0.039)}$	$1.87_{(0.041)}$	$-3.73_{(0.059)}$	$4.63_{(0.074)}$
$\hat\beta^{RE}_{\left(\text{var}(\hat\beta^{RE})\right)}$	$2.95_{(0.016)}$	$0.93_{(0.039)}$	$1.87_{(0.041)}$	$-3.73_{(0.059)}$	$4.63_{(0.074)}$
$\hat\beta^{LWS}_{\left(\text{var}(\hat\beta^{LWS})\right)}$	$3.00_{(0.015)}$	$0.99_{(0.007)}$	$2.00_{(0.006)}$	$-4.00_{(0.006)}$	$4.99_{(0.006)}$
$\hat\beta^{FWE}_{\left(\text{var}(\hat\beta^{FWE})\right)}$	$0.00_{(0.000)}$	$1.00_{(0.003)}$	$2.00_{(0.003)}$	$-4.00_{(0.003)}$	$4.99_{(0.003)}$
$\hat\beta^{RWE}\left(\text{var}(\hat\beta^{RWE})\right)$	$3.01_{(0.158)}$	$1.00_{(0.003)}$	$2.00_{(0.003)}$	$-4.00_{(0.003)}$	$4.99_{(0.003)}$
Contamination level is equal to 1 %					
$\hat\beta^{OLS}_{\left(\text{var}(\hat\beta^{OLS})\right)}$	$2.90_{(0.041)}$	$0.87_{(0.055)}$	$1.76_{(0.068)}$	$-3.49_{(0.093)}$	$4.36_{(0.105)}$
$\hat\beta^{FE}_{\left(\text{var}(\hat\beta^{FE})\right)}$	$0.00_{(0.000)}$	$0.87_{(0.055)}$	$1.76_{(0.068)}$	$-3.49_{(0.096)}$	$4.37_{(0.105)}$
$\hat\beta^{RE}_{\left(\text{var}(\hat\beta^{RE})\right)}$	$2.06_{(0.131)}$	$0.87_{(0.055)}$	$1.76_{(0.069)}$	$-3.49_{(0.095)}$	$4.36_{(0.106)}$
$\hat\beta^{LWS}_{\left(\text{var}(\hat\beta^{LWS})\right)}$	$3.00_{(0.026)}$	$1.00_{(0.004)}$	$2.00_{(0.003)}$	$-4.00_{(0.004)}$	$5.00_{(0.004)}$
$\hat\beta^{FWE}_{\left(\text{var}(\hat\beta^{FWE})\right)}$	$0.00_{(0.000)}$	$1.00_{(0.002)}$	$1.99_{(0.002)}$	$-3.98_{(0.002)}$	$4.98_{(0.002)}$
$\hat\beta^{RWE}_{\left(\text{var}(\hat\beta^{RWE})\right)}$	$0.50_{(0.118)}$	$1.00_{(0.002)}$	$1.99_{(0.002)}$	$-3.98_{(0.002)}$	$4.99_{(0.002)}$
Contamination level is equal to 4 %					
$\hat\beta^{OLS}_{\left(\text{var}(\hat\beta^{OLS})\right)}$	$2.57_{(0.093)}$	$0.55_{(0.135)}$	$1.11_{(0.164)}$	$-2.22_{(0.243)}$	$2.81_{(0.292)}$
$\hat\beta^{FE}_{\left(\text{var}(\hat\beta^{FE})\right)}$	$0.00_{(0.000)}$	$0.55_{(0.135)}$	$1.11_{(0.164)}$	$-2.22_{(0.249)}$	$2.81_{(0.295)}$
$\hat\beta^{RE}_{\left(\text{var}(\hat\beta^{RE})\right)}$	$2.26_{(0.127)}$	$0.55_{(0.134)}$	$1.11_{(0.164)}$	$-2.21_{(0.243)}$	$2.80_{(0.293)}$
$\hat\beta^{LWS}_{\left(\text{var}(\hat\beta^{LWS})\right)}$	$2.99_{(0.023)}$	$1.00_{(0.003)}$	$2.00_{(0.003)}$	$-4.00_{(0.003)}$	$5.00_{(0.003)}$
$\hat\beta^{FWE}_{\left(\text{var}(\hat\beta^{FWE})\right)}$	$0.00_{(0.000)}$	$0.99_{(0.002)}$	$1.98_{(0.002)}$	$-3.96_{(0.002)}$	$4.95_{(0.002)}$
$\hat\beta^{RWE}_{\left(\text{var}(\hat\beta^{RWE})\right)}$	$0.78_{(0.292)}$	$0.99_{(0.002)}$	$1.99_{(0.002)}$	$-3.98_{(0.002)}$	$4.97_{(0.002)}$
Contamination level is equal to 8 %					
$\hat\beta^{OLS}_{\left(\text{var}(\hat\beta^{OLS})\right)}$	$2.06_{(0.135)}$	$0.16_{(0.184)}$	$0.37_{(0.210)}$	$-0.69_{(0.303)}$	$0.88_{(0.313)}$
$\hat\beta^{FE}_{\left(\text{var}(\hat\beta^{FE})\right)}$	$0.00_{(0.000)}$	$0.16_{(0.190)}$	$0.37_{(0.219)}$	$-0.69_{(0.314)}$	$0.89_{(0.322)}$
$\hat\beta^{RE}_{\left(\text{var}(\hat\beta^{RE})\right)}$	$1.91_{(0.190)}$	$0.16_{(0.184)}$	$0.37_{(0.211)}$	$-0.69_{(0.303)}$	$0.87_{(0.315)}$
$\hat\beta^{LWS}_{\left(\text{var}(\hat\beta^{LWS})\right)}$	$3.00_{(0.027)}$	$1.00_{(0.005)}$	$2.00_{(0.005)}$	$-4.00_{(0.005)}$	$5.00_{(0.004)}$
$\hat\beta^{FWE}_{\left(\text{var}(\hat\beta^{FWE})\right)}$	$0.00_{(0.000)}$	$0.98_{(0.004)}$	$1.96_{(0.005)}$	$-3.92_{(0.004)}$	$4.91_{(0.005)}$
$\hat\beta^{RWE}_{\left(\text{var}(\hat\beta^{RWE})\right)}$	$0.86_{(0.211)}$	$0.98_{(0.004)}$	$1.97_{(0.004)}$	$-3.95_{(0.004)}$	$4.94_{(0.004)}$

Table 2 (continued)

True coeffs β^0	**3**	**1**	**2**	**−4**	**5**
Contamination level is equal to 16 %					
$\hat{\beta}^{OLS}_{\left(\text{var}(\hat{\beta}^{OLS})\right)}$	$1.27_{(0.174)}$	$-0.30_{(0.201)}$	$-0.54_{(0.230)}$	$1.13_{(0.294)}$	$-1.42_{(0.375)}$
$\hat{\beta}^{FE}_{\left(\text{var}(\hat{\beta}^{FE})\right)}$	$0.00_{(0.000)}$	$-0.30_{(0.212)}$	$-0.54_{(0.234)}$	$1.12_{(0.298)}$	$-1.41_{(0.379)}$
$\hat{\beta}^{RE}_{\left(\text{var}(\hat{\beta}^{RE})\right)}$	$1.19_{(0.243)}$	$-0.30_{(0.201)}$	$-0.54_{(0.229)}$	$1.13_{(0.295)}$	$-1.43_{(0.377)}$
$\hat{\beta}^{LWS}_{\left(\text{var}(\hat{\beta}^{LWS})\right)}$	$2.99_{(0.022)}$	$1.00_{(0.004)}$	$2.01_{(0.004)}$	$-3.99_{(0.004)}$	$5.00_{(0.004)}$
$\hat{\beta}^{FWE}_{\left(\text{var}(\hat{\beta}^{FWE})\right)}$	$0.00_{(0.000)}$	$0.96_{(0.006)}$	$1.92_{(0.005)}$	$-3.84_{(0.007)}$	$4.81_{(0.007)}$
$\hat{\beta}^{RWE}_{\left(\text{var}(\hat{\beta}^{RWE})\right)}$	$0.96_{(0.132)}$	$0.97_{(0.004)}$	$1.95_{(0.004)}$	$-3.89_{(0.005)}$	$4.87_{(0.005)}$

$(\tilde{X}_{it1} \equiv 0)$. Hence, let's put $\tilde{V}_{itj} = \tilde{X}_{itj+1}$ and $\gamma_j^0 = \beta_{j+1}^0$ for $i = 1, 2, ..., n, t = 1, 2, ..., T$ and $j = 1, 2, ...p - 1$,, we have

$$\tilde{Y}_{it} = \tilde{V}_{it}'\gamma^0 + e_{it}, \quad \text{with} \;\; \mathbb{E}\left[\tilde{V}_{itj} \cdot e_{it}\right] = 0 \tag{8}$$

and hence $\hat{\gamma}^{(OLS,n,T)}$ is unbiased and consistent but even under the normality of e_{it}'s it is not efficient as the covariance matrix of "new" disturbances $\tilde{e}_{it} = e_{it} - \bar{e}_i$ is not diagonal (of course, the efficiency can be reached by $\hat{\beta}^{(EGLS,n,T)}$ as we can estimate the covariance matrix in question). The intercept, if it was included in the original model, can be additionally estimated by $\hat{\beta}_1^{(OLS,n,T)} = \frac{1}{n}\sum_{i=1}^n \left[\overline{Y}_i - \overline{X}_i'\hat{\gamma}^{(OLS,n,T)}\right]$. An alternative way is to make differences of successive observation within the groups. It is evident that the differences don't contain the fixed effects, so we can apply the ordinary least squares. But this method is not suitable for robustification - it is impossible (at least form the application point of view) to rid of the contamination in this way.

On the very first lines of this paper we have reminded that the classical estimators (typically) suffer by vulnerability to outliers and/or leverage points. That is why we propose below the robustified versions of the both just recalled estimating methods. Prior to it let's briefly mention what was already done. There are only a few papers devoted to the topic. Some of them are case studies ((Amos 1994; Rocke 1991) - analysis of variance, Jones (1992) - trying to cope with publication bias, paper from meta-analysis, (Cameron et al. 2011) - clustering) considering the situations which need not assume the leverage points. Therefore they cope with the problem by quantile regression of Koenker and Bassett (1978) - the former even addressed the topic but he restricted himself also only on quantile regression[3], see Koenker (2004). Some papers try to solve the problem with outliers by employing the distribution of disturbances with heavy tails (Gill 2000; Pinheiro and Liu 2001) or by the bayesian approach (Schall 1991).

[3]One can notice that on the economic conferences when people referring results of case studies consider more and more robust approach. Unfortunately they employ (nearly) exclusively quantile regression, probably mostly due to the fact that it is available in commercially supplied packages. Sometimes they are not even aware that the quantile regression can cope (reliably) only with outliers but (distant) leverage points can cause problems. Moreover, in economic papers we can meet with the robustness with respect to heteroscedaticity in the sense of paper by White (1980).

Table 3 The disturbances are independent from explanatory variables but the effects are correlated with them. Contamination is created by outliers, i. e. explanatory variables are not changed but values of contaminated response variables are equal to −2.5 multiple of original values of response variable

True coeffs β^0	**3**	**1**	**2**	**−4**	**5**
Contamination level is equal to 0.5 %					
$\hat{\beta}^{OLS}_{(\mathrm{var}(\hat{\beta}^{OLS}))}$	$2.94_{(0.088)}$	$0.59_{(0.062)}$	$1.58_{(0.059)}$	$-4.25_{(0.023)}$	$4.46_{(0.131)}$
$\hat{\beta}^{FE}_{(\mathrm{var}(\hat{\beta}^{FE}))}$	$0.00_{(0.000)}$	$0.77_{(0.049)}$	$1.76_{(0.047)}$	$-4.07_{(0.021)}$	$4.64_{(0.104)}$
$\hat{\beta}^{RE}_{(\mathrm{var}(\hat{\beta}^{RE}))}$	$2.03_{(0.559)}$	$0.63_{(0.045)}$	$1.62_{(0.043)}$	$-4.21_{(0.018)}$	$4.50_{(0.102)}$
$\hat{\beta}^{LWS}_{(\mathrm{var}(\hat{\beta}^{LWS}))}$	$3.00_{(0.004)}$	$0.80_{(0.003)}$	$1.80_{(0.003)}$	$-4.20_{(0.003)}$	$4.80_{(0.003)}$
$\hat{\beta}^{FWE}_{(\mathrm{var}(\hat{\beta}^{FWE}))}$	$0.00_{(0.000)}$	$0.99_{(0.003)}$	$1.99_{(0.003)}$	$-3.99_{(0.003)}$	$4.98_{(0.003)}$
$\hat{\beta}^{RWE}_{(\mathrm{var}(\hat{\beta}^{RWE}))}$	$0.90_{(0.351)}$	$0.95_{(0.006)}$	$1.94_{(0.006)}$	$-4.04_{(0.005)}$	$4.94_{(0.005)}$
Contamination level is equal to 1 %					
$\hat{\beta}^{OLS}_{(\mathrm{var}(\hat{\beta}^{OLS}))}$	$2.87_{(0.196)}$	$0.47_{(0.104)}$	$1.40_{(0.129)}$	$-4.24_{(0.043)}$	$4.23_{(0.230)}$
$\hat{\beta}^{FE}_{(\mathrm{var}(\hat{\beta}^{FE}))}$	$0.00_{(0.000)}$	$0.67_{(0.070)}$	$1.60_{(0.086)}$	$-4.04_{(0.042)}$	$4.43_{(0.162)}$
$\hat{\beta}^{RE}_{(\mathrm{var}(\hat{\beta}^{RE}))}$	$1.84_{(0.756)}$	$0.53_{(0.073)}$	$1.46_{(0.091)}$	$-4.18_{(0.036)}$	$4.29_{(0.173)}$
$\hat{\beta}^{LWS}_{(\mathrm{var}(\hat{\beta}^{LWS}))}$	$3.00_{(0.004)}$	$0.80_{(0.002)}$	$1.80_{(0.002)}$	$-4.20_{(0.002)}$	$4.80_{(0.002)}$
$\hat{\beta}^{FWE}_{(\mathrm{var}(\hat{\beta}^{FWE}))}$	$0.00_{(0.000)}$	$0.99_{(0.003)}$	$1.99_{(0.003)}$	$-3.99_{(0.003)}$	$4.98_{(0.003)}$
$\hat{\beta}^{RWE}_{(\mathrm{var}(\hat{\beta}^{RWE}))}$	$0.77_{(0.198)}$	$0.96_{(0.004)}$	$1.96_{(0.004)}$	$-4.03_{(0.005)}$	$4.95_{(0.004)}$
Contamination level is equal to 4 %					
$\hat{\beta}^{OLS}_{(\mathrm{var}(\hat{\beta}^{OLS}))}$	$2.62_{(0.320)}$	$0.04_{(0.179)}$	$0.88_{(0.185)}$	$-4.09_{(0.102)}$	$3.34_{(0.453)}$
$\hat{\beta}^{FE}_{(\mathrm{var}(\hat{\beta}^{FE}))}$	$0.00_{(0.000)}$	$0.21_{(0.152)}$	$1.05_{(0.180)}$	$-3.91_{(0.096)}$	$3.51_{(0.425)}$
$\hat{\beta}^{RE}_{(\mathrm{var}(\hat{\beta}^{RE}))}$	$1.33_{(0.503)}$	$0.08_{(0.156)}$	$0.92_{(0.180)}$	$-4.04_{(0.080)}$	$3.37_{(0.443)}$
$\hat{\beta}^{LWS}_{(\mathrm{var}(\hat{\beta}^{LWS}))}$	$3.00_{(0.004)}$	$0.80_{(0.002)}$	$1.81_{(0.002)}$	$-4.20_{(0.002)}$	$4.80_{(0.003)}$
$\hat{\beta}^{FWE}_{(\mathrm{var}(\hat{\beta}^{FWE}))}$	$0.00_{(0.000)}$	$0.99_{(0.002)}$	$1.98_{(0.003)}$	$-3.99_{(0.003)}$	$4.98_{(0.003)}$
$\hat{\beta}^{RWE}_{(\mathrm{var}(\hat{\beta}^{RWE}))}$	$0.65_{(0.017)}$	$0.99_{(0.003)}$	$1.98_{(0.003)}$	$-4.00_{(0.003)}$	$4.98_{(0.003)}$
Contamination level is equal to 8 %					
$\hat{\beta}^{OLS}_{(\mathrm{var}(\hat{\beta}^{OLS}))}$	$2.20_{(0.580)}$	$-0.45_{(0.211)}$	$0.14_{(0.284)}$	$-3.60_{(0.220)}$	$2.05_{(0.632)}$
$\hat{\beta}^{FE}_{(\mathrm{var}(\hat{\beta}^{FE}))}$	$0.00_{(0.000)}$	$-0.27_{(0.208)}$	$0.32_{(0.278)}$	$-3.42_{(0.216)}$	$2.22_{(0.614)}$
$\hat{\beta}^{RE}_{(\mathrm{var}(\hat{\beta}^{RE}))}$	$1.00_{(0.361)}$	$-0.43_{(0.191)}$	$0.15_{(0.268)}$	$-3.57_{(0.192)}$	$2.05_{(0.614)}$
$\hat{\beta}^{LWS}_{(\mathrm{var}(\hat{\beta}^{LWS}))}$	$3.00_{(0.003)}$	$0.80_{(0.002)}$	$1.81_{(0.002)}$	$-4.20_{(0.002)}$	$4.80_{(0.002)}$
$\hat{\beta}^{FWE}_{(\mathrm{var}(\hat{\beta}^{FWE}))}$	$0.00_{(0.000)}$	$0.97_{(0.002)}$	$1.97_{(0.003)}$	$-4.00_{(0.002)}$	$4.96_{(0.003)}$
$\hat{\beta}^{RWE}_{(\mathrm{var}(\hat{\beta}^{RWE}))}$	$0.67_{(0.008)}$	$1.00_{(0.002)}$	$2.00_{(0.002)}$	$-3.98_{(0.002)}$	$4.99_{(0.002)}$
Contamination level is equal to 16 %					
$\hat{\beta}^{OLS}_{(\mathrm{var}(\hat{\beta}^{OLS}))}$	$1.37_{(0.687)}$	$-1.12_{(0.249)}$	$-0.91_{(0.328)}$	$-2.23_{(0.472)}$	$-0.23_{(0.824)}$
$\hat{\beta}^{FE}_{(\mathrm{var}(\hat{\beta}^{FE}))}$	$0.00_{(0.000)}$	$-0.93_{(0.301)}$	$-0.73_{(0.388)}$	$-2.05_{(0.410)}$	$-0.06_{(0.911)}$
$\hat{\beta}^{RE}_{(\mathrm{var}(\hat{\beta}^{RE}))}$	$0.53_{(0.306)}$	$-1.16_{(0.244)}$	$-0.96_{(0.332)}$	$-2.26_{(0.415)}$	$-0.31_{(0.841)}$

Table 3 (continued)

True coeffs β^0	3	1	2	−4	5
$\hat{\beta}^{LWS}_{(\text{var}(\hat{\beta}^{LWS}))}$	$3.00_{(0.004)}$	$0.80_{(0.002)}$	$1.81_{(0.003)}$	$-4.20_{(0.002)}$	$4.80_{(0.003)}$
$\hat{\beta}^{FWE}_{(\text{var}(\hat{\beta}^{FWE}))}$	$0.00_{(0.000)}$	$0.93_{(0.006)}$	$1.92_{(0.006)}$	$-4.01_{(0.005)}$	$4.89_{(0.006)}$
$\hat{\beta}^{RWE}_{(\text{var}(\hat{\beta}^{RWE}))}$	$0.71_{(0.018)}$	$1.00_{(0.003)}$	$2.01_{(0.003)}$	$-3.95_{(0.003)}$	$4.98_{(0.004)}$

Robustification of the classical estimation of *model with random effects* and of *model with fixed effects* consists of substituting $\hat{\beta}^{(OLS,n,T)}, \overline{Y}_i, \overline{X}_i, \hat{\sigma}_v^2$ and $\hat{\sigma}_u^2$ by $\hat{\beta}^{(LWS,n,T,w)}, \overline{Y}_i^{(LWS,n,T,w)}, \overline{X}_i^{(LWS,n,T,w)}, \hat{\sigma}^2_{LWS,v}$ and $\hat{\sigma}^2_{LWS,u}$, respectively. Such estimator is denoted in the tables below as the *Random Weighted Effects* estimator $\hat{\beta}^{RWE}$ and the *Fixed Weighted Effects* estimator $\hat{\beta}^{FWE}$, respectively. However, we are not able to derive analytically an improvement (if any) of $\hat{\beta}^{RWE}$ compared to $\hat{\beta}^{LWS}$ (of course the same is true for mutual relation of $\hat{\beta}^{FWE}$ and $\hat{\beta}^{LWS}$). That is why we performed the numerical study and we discuss the results in Section 5.

A few lines above we mentioned that in the robustification of $\hat{\beta}^{RE}$ (to obtain $\hat{\beta}^{RWE}$) we need to employ the robust estimates of scale of the disturbaces v's and u's and we proposed to utilize $\hat{\sigma}^2_{LWS,v}$ and $\hat{\sigma}^2_{LWS,u}$. Let's be more specific and give the first of it, $\hat{\sigma}^2_{LWS,v}$, explicitly. We estimate the naive $\hat{\beta}^{(LWS,n,T,w)}$ for the model given in Eq. 5, say $\hat{\beta}^{(LWS,n,T,w)}(Y, X)$. Employing the residuals from this model, say

$$r_{it}\left(\hat{\beta}^{(LWS,n,T,w)}\right) = Y_{it} - X'_{it} \cdot \hat{\beta}^{(LWS,n,T,w)},$$

we calculate instead of the classical estimates of scale of disturbances v's

$$\hat{\sigma}_v^2 = \frac{1}{n \cdot T} \sum_{i=1}^{n} \sum_{t=1}^{T} \left(r_{it}(\hat{\beta}^{(LWS,n,T,w)} - \bar{r}(\hat{\beta}^{(LWS,n,T,w)}\right)^2$$

with

$$\bar{r}(\hat{\beta}^{(LWS,n,T,w)}) = \frac{1}{n \cdot T} \sum_{i=1}^{n} \sum_{t=1}^{T} r_{it}(\hat{\beta}^{(LWS,n,T,w)}),$$

the robust LWS-estimates of location and variance as follows. Let us denote

$$a_{(it)}\left(\hat{\beta}^{(LWS,n,T,w)}\right) = \left|r_{it}\left(\hat{\beta}^{(LWS,n,T,w)}\right)\right|$$

and by $a_{(\ell)}\left(\hat{\beta}^{(LWS,n,T,w)}\right)$ the corresponding order statistics, i. e. we have

$$a_{(1)}\left(\hat{\beta}^{(LWS,n,T,w)}\right) \leq ... \leq a_{(\ell)}\left(\hat{\beta}^{(LWS,n,T,w)}\right) \leq ... \leq a_{(n \cdot T)}\left(\hat{\beta}^{(LWS,n,T,w)}\right).$$

Moreover, let s_ℓ is equal 1 if the residual $r_{it}\left(\hat{\beta}^{(LWS,n,T,w)}\right)$ corresponding to $a_{(\ell)}\left(\hat{\beta}^{(LWS,n,T,w)}\right)$ has the nonnegative sign and equal to -1 otherwise. Then put

$$\bar{r}^{(LWS,n,T,w)}(\hat{\beta}^{(LWS,n,T,w)}) = \frac{1}{n \cdot T} \sum_{i=1}^{n} \sum_{t=1}^{T} s_\ell \cdot w_\ell \cdot a_{(\ell)}\left(\hat{\beta}^{(LWS,n,T,w)}\right). \tag{9}$$

Table 4 The disturbances are independent from explanatory variables but the effects are correlated with them. Contamination is created by leverage points, i. e. values of explanatory variables are equal to the 10 times the original values of them and response is taken with minus sign

True coeffs $\boldsymbol{\beta}^0$	**3**	**1**	**2**	**−4**	**5**
Contamination level is equal to 0.5 %					
$\hat{\beta}^{OLS}_{(\mathrm{var}(\hat{\beta}^{OLS}))}$	$2.36_{(1.761)}$	$-0.00_{(1.922)}$	$0.66_{(2.852)}$	$-2.91_{(3.543)}$	$2.35_{(6.582)}$
$\hat{\beta}^{FE}_{(\mathrm{var}(\hat{\beta}^{FE}))}$	$0.00_{(0.000)}$	$-0.42_{(1.555)}$	$0.24_{(2.073)}$	$-3.39_{(4.343)}$	$1.97_{(5.629)}$
$\hat{\beta}^{RE}_{(\mathrm{var}(\hat{\beta}^{RE}))}$	$0.88_{(15.404)}$	$-0.25_{(1.514)}$	$0.41_{(2.185)}$	$-3.30_{(3.625)}$	$2.22_{(5.867)}$
$\hat{\beta}^{LWS}_{(\mathrm{var}(\hat{\beta}^{LWS}))}$	$3.00_{(0.002)}$	$1.20_{(0.002)}$	$2.20_{(0.002)}$	$-3.81_{(0.002)}$	$5.20_{(0.002)}$
$\hat{\beta}^{FWE}_{(\mathrm{var}(\hat{\beta}^{FWE}))}$	$0.00_{(0.000)}$	$1.00_{(0.003)}$	$1.99_{(0.003)}$	$-4.01_{(0.003)}$	$4.99_{(0.003)}$
$\hat{\beta}^{RWE}_{(\mathrm{var}(\hat{\beta}^{RWE}))}$	$3.00_{(0.018)}$	$1.07_{(0.006)}$	$2.06_{(0.006)}$	$-3.94_{(0.006)}$	$5.07_{(0.005)}$
Contamination level is equal to 1 %					
$\hat{\beta}^{OLS}_{(\mathrm{var}(\hat{\beta}^{OLS}))}$	$-1.15_{(1.833)}$	$-0.39_{(1.168)}$	$-0.80_{(1.546)}$	$1.62_{(1.781)}$	$-1.98_{(2.575)}$
$\hat{\beta}^{FE}_{(\mathrm{var}(\hat{\beta}^{FE}))}$	$0.00_{(0.000)}$	$-0.41_{(1.368)}$	$-0.86_{(1.673)}$	$1.75_{(1.911)}$	$-2.17_{(2.586)}$
$\hat{\beta}^{RE}_{(\mathrm{var}(\hat{\beta}^{RE}))}$	$0.01_{(0.179)}$	$-0.42_{(1.368)}$	$-0.86_{(1.672)}$	$1.75_{(1.899)}$	$-2.17_{(2.580)}$
$\hat{\beta}^{LWS}_{(\mathrm{var}(\hat{\beta}^{LWS}))}$	$3.00_{(0.003)}$	$1.00_{(0.003)}$	$2.00_{(0.003)}$	$-4.00_{(0.003)}$	$5.00_{(0.003)}$
$\hat{\beta}^{FWE}_{(\mathrm{var}(\hat{\beta}^{FWE}))}$	$0.00_{(0.000)}$	$0.96_{(0.025)}$	$1.94_{(0.023)}$	$-3.88_{(0.024)}$	$4.85_{(0.025)}$
$\hat{\beta}^{RWE}_{(\mathrm{var}(\hat{\beta}^{RWE}))}$	$0.01_{(0.032)}$	$0.97_{(0.022)}$	$1.96_{(0.018)}$	$-3.91_{(0.019)}$	$4.89_{(0.020)}$
Contamination level is equal to 4 %					
$\hat{\beta}^{OLS}_{(\mathrm{var}(\hat{\beta}^{OLS}))}$	$-1.35_{(1.153)}$	$-0.93_{(0.200)}$	$-1.57_{(0.285)}$	$2.21_{(1.030)}$	$-3.44_{(0.857)}$
$\hat{\beta}^{FE}_{(\mathrm{var}(\hat{\beta}^{FE}))}$	$0.00_{(0.000)}$	$-1.10_{(0.229)}$	$-1.74_{(0.272)}$	$2.08_{(1.055)}$	$-3.65_{(0.822)}$
$\hat{\beta}^{RE}_{(\mathrm{var}(\hat{\beta}^{RE}))}$	$0.50_{(0.294)}$	$-1.05_{(0.235)}$	$-1.70_{(0.287)}$	$2.15_{(1.021)}$	$-3.62_{(0.831)}$
$\hat{\beta}^{LWS}_{(\mathrm{var}(\hat{\beta}^{LWS}))}$	$3.00_{(0.005)}$	$0.80_{(0.003)}$	$1.80_{(0.004)}$	$-4.20_{(0.003)}$	$4.80_{(0.003)}$
$\hat{\beta}^{FWE}_{(\mathrm{var}(\hat{\beta}^{FWE}))}$	$0.00_{(0.000)}$	$0.98_{(0.004)}$	$1.98_{(0.004)}$	$-3.99_{(0.005)}$	$4.96_{(0.004)}$
$\hat{\beta}^{RWE}_{(\mathrm{var}(\hat{\beta}^{RWE}))}$	$0.67_{(0.009)}$	$0.99_{(0.004)}$	$1.98_{(0.004)}$	$-4.00_{(0.004)}$	$4.97_{(0.004)}$
Contamination level is equal to 8 %					
$\hat{\beta}^{OLS}_{(\mathrm{var}(\hat{\beta}^{OLS}))}$	$-2.75_{(0.006)}$	$-0.92_{(0.004)}$	$-1.83_{(0.005)}$	$3.66_{(0.007)}$	$-4.58_{(0.008)}$
$\hat{\beta}^{FE}_{(\mathrm{var}(\hat{\beta}^{FE}))}$	$0.00_{(0.000)}$	$-0.92_{(0.088)}$	$-1.85_{(0.086)}$	$3.66_{(0.090)}$	$-4.58_{(0.095)}$
G $\hat{\beta}^{RE}_{(\mathrm{var}(\hat{\beta}^{RE}))}$	$0.03_{(0.309)}$	$-0.92_{(0.087)}$	$-1.85_{(0.085)}$	$3.66_{(0.088)}$	$-4.58_{(0.092)}$
$\hat{\beta}^{LWS}_{(\mathrm{var}(\hat{\beta}^{LWS}))}$	$3.00_{(0.008)}$	$1.00_{(0.008)}$	$2.00_{(0.008)}$	$-4.00_{(0.008)}$	$4.99_{(0.007)}$
$\hat{\beta}^{FWE}_{(\mathrm{var}(\hat{\beta}^{FWE}))}$	$0.00_{(0.000)}$	$0.96_{(0.045)}$	$1.94_{(0.046)}$	$-3.86_{(0.047)}$	$4.83_{(0.045)}$
$\hat{\beta}^{RWE}_{(\mathrm{var}(\hat{\beta}^{RWE}))}$	$-0.00_{(0.041)}$	$0.97_{(0.042)}$	$1.96_{(0.044)}$	$-3.91_{(0.043)}$	$4.88_{(0.040)}$
Contamination level is equal to 16 %					
$\hat{\beta}^{OLS}_{(\mathrm{var}(\hat{\beta}^{OLS}))}$	$-2.65_{(0.016)}$	$-1.00_{(0.002)}$	$-1.93_{(0.003)}$	$3.64_{(0.010)}$	$-4.72_{(0.008)}$
$\hat{\beta}^{FE}_{(\mathrm{var}(\hat{\beta}^{FE}))}$	$0.00_{(0.000)}$	$-1.09_{(0.029)}$	$-2.03_{(0.029)}$	$3.54_{(0.036)}$	$-4.81_{(0.029)}$
$\hat{\beta}^{RE}_{(\mathrm{var}(\hat{\beta}^{RE}))}$	$-1.85_{(0.466)}$	$-1.06_{(0.029)}$	$-2.00_{(0.028)}$	$3.57_{(0.035)}$	$-4.79_{(0.028)}$
$\hat{\beta}^{LWS}_{(\mathrm{var}(\hat{\beta}^{LWS}))}$	$3.01_{(0.009)}$	$0.81_{(0.005)}$	$1.80_{(0.006)}$	$-4.20_{(0.006)}$	$4.80_{(0.005)}$
$\hat{\beta}^{FWE}_{(\mathrm{var}(\hat{\beta}^{FWE}))}$	$0.00_{(0.000)}$	$0.92_{(0.021)}$	$1.89_{(0.034)}$	$-4.04_{(0.034)}$	$4.87_{(0.036)}$
$\hat{\beta}^{RWE}_{(\mathrm{var}(\hat{\beta}^{RWE}))}$	$0.71_{(0.019)}$	$0.99_{(0.011)}$	$1.98_{(0.011)}$	$-3.98_{(0.017)}$	$4.96_{(0.010)}$

Further denote

$$b_{it}^2(\hat{\beta}^{(LWS,n,T,w)}) = \left(r_{it}(\hat{\beta}^{(LWS,n,T,w)}) - \bar{r}^{(LWS,n,T,w)}(\hat{\beta}^{(LWS,n,T,w)})\right)^2$$

and (similarly as above) $b_{(\ell)}^2(\hat{\beta}^{(LWS,n,T,w)})$ the corresponding order statistics, i. e. we have

$$b_{(1)}^2(\hat{\beta}^{(LWS,n,T,w)}) \leq ... \leq b_{(\ell)}^2(\hat{\beta}^{(LWS,n,T,w)}) \leq ... \leq b_{(n\cdot T)}^2(\hat{\beta}^{(LWS,n,T,w)}).$$

Finally, put

$$\hat{\sigma}_{LWS,v}^2 = \frac{1}{n \cdot T} \sum_{i=1}^{n} \sum_{t=1}^{T} w_\ell \cdot b_{(\ell)}^2(\hat{\beta}^{(LWS,n,T,w)}).$$

In the similar way we estimate $\hat{\sigma}_{LWS,u}^2$ (of course, we use the classical "trick" over empirical covariances of $r_{it}(\hat{\beta}^{(LWS,n,T,w)})$, see e. g. Judge et al. (1985)). It allows to compute

$$\hat{\sigma}_{LWS,e}^2 = \hat{\sigma}_{LWS,v}^2 - \hat{\sigma}_{LWS,u}^2.$$

Employing formula for λ given in Eq. 6 but plugging in the robust estimates $\hat{\sigma}_{LWS,e}^2$ and $\hat{\sigma}_{LWS,u}^2$, we obtain a robust estimate $\hat{\lambda}^{(LWS,n,T,w)}$. Finally, estimating by LWS-approach also mean values of Y_{it}'s and X_{it}'s, we conclude transformation of original data by formulas given in the left-hand-side of Eq. 6 and obtain (say) $\tilde{Y}_{it}^{(LWS,n,T,w)}$ and $\tilde{X}_{it}^{(LWS,n,T,w)}$. Then we compute

$$\hat{\beta}^{(LWS,n,t)}\left(\tilde{Y}_{it}^{(LWS,n,T,w)}, \tilde{X}_{it}^{(LWS,n,T,w)}\right).$$

For robustification of the classical estimation of the *model with fixed effects* we need robust estimators of the mean values of response $\overline{Y}_i$ and of explanatory variables $\overline{X}_i$ by $\overline{Y}_i^{(LWS,n,T,w)}$ and $\overline{X}_i^{(LWS,n,T,w)}$, respectively. These estimator we calculate in the same way as it was indicated for location and scale estimators of the residulas $r_{it}\left(\hat{\beta}^{(LWS,n,T,w)}\right)$, i. e. employing the order statistics of absolute values of Y_{it}'s and X_{itj}'s, their original signs and weights w_ℓ - see Eq. 9. Then employing transformation (7) with these estimates instead of with $\overline{Y}_i$'s and $\overline{X}_i$'s, we can finally employ $\hat{\gamma}^{(LWS,n,T,w)}$ (i. e. analogy of $\hat{\beta}^{(LWS,n,T,w)}$ for the model (8)) we compute the *Fixed Weighted Effects Estimator* (denoted below as $\hat{\beta}^{FWE}$).

4 Numerical Study

First of all, we specified the model given in Eq. 1 as

$$Y_{it} = 3+1\cdot X_{it1}+2\cdot X_{it2}-4\cdot X_{it3}+5\cdot X_{it5}+u_i+e_{it}, \quad i=1,2,...,50, \;\; t=1,2,...,20. \quad (10)$$

Having generated data

$$\left\{ \left\{\left\{X_{it}^{(k)}, e_{it}^{(k)}\right\}_{t=1}^{20}\right\}_{i=1}^{50}, \quad \left\{u_i^{(k)}\right\}_{i=1}^{50}\right\}_{k=1}^{500}$$

where all variables were distributed according to the standard normal distribution[4], we put $\tilde{X}_{itj}^{(k)} = X_{itj}^{(k)} + u_i^{(k)}$ for $i = 1, 2, ..., 50 \;\; t = 1, 2, ..., 20$ and $j = 1, 2, ..., 5$. Then employing

[4] We have used also data with larger or smaller variance of effects - it is described in the heads of tables below.

(10) we computed response variables for two groups (for the second one we employed $\tilde{X}_{it}^{(k)}$'s instead of $X_{it}^{(k)}$'s)

$$\left\{ \left\{ \left\{ Y_{it}^{(k)}, X_{it}^{(k)} \right\}_{t=1}^{20} \right\}_{i=1}^{50} \right\}_{k=1}^{500} \quad \text{and} \quad \left\{ \left\{ \left\{ \tilde{Y}_{it}^{(k)}, \tilde{X}_{it}^{(k)} \right\}_{t=1}^{20} \right\}_{i=1}^{50} \right\}_{k=1}^{500}.$$

The first group represents data for the model with random effects, the second one the model with fixed effects. Then the estimates of regression coefficients were computed - for both groups all estimators, i. e. $\hat{\beta}^{(OLS,n,T)}$, $\hat{\beta}^{FE}$, $\hat{\beta}^{RE}$, $\hat{\beta}^{(LWS,n,T,w)}$, $\hat{\beta}^{FWE}$ and $\hat{\beta}^{RWE}$ were computed, to offer the reader a possibility to create an idea how e. g. estimator proposed for the model with the fixed effects works when estimating coefficients of the model with random effects etc.. So, we obtained, say

$$\left\{ \hat{\beta}^{(index,k)} = (\hat{\beta}_1, \hat{\beta}_2, ..., \hat{\beta}_5)' \right\}_{k=1}^{500}$$

(where index "attains values" *OLS, RE, FE, LWS, RWE* and *FWE* indicating the method employed for the computation).

The weight function $w(r) : [0, 1] \rightarrow [0, 1]$ was selected in an optimal way - see a brief discussion in Conclusions. We used the weights:

$$w_i = 1 \quad \text{for } i = 1, 2, ..., h, \quad h \in \{1, 2, ..., n\},$$

$$w_i = 0 \quad \text{for } i = g, g+1, ..., n, \quad g \in \{1, 2, ..., n\}, \quad g > h$$

and for $h \leq i \leq g$ the weights w_i's decreased linearly from 1 to 0. The values of h and g was selected according to our long years experiences to be approximately optimal for given situation, the details can be found in (Víšek 2011c; 2013; 2014a) or (Víšek 2014b). We can generally say that g has to be selected so that the weight function assigns weights equal to zero all leverage points and outliers it is able to cope with small number of outliers even in the interval between h and g. The selection of h is approximately optimal in (surprisingly) wide range of values - even rather small values of h don't decrease the efficiency of estimation too much. In the case of processing the real data we can employ the "forward search", up to the moment when the estimated model significantly changes (usually by a jump), see Atkinson and Riani (2000).

The empirical means and empirical variances of estimates of coefficients (over these 500 repetitions indicated above)

$$\hat{\beta}_j^{(index)} = \frac{1}{500} \sum_{k=1}^{500} \hat{\beta}_j^{(index,k)} \quad \text{and} \quad \widehat{\text{var}} \left(\hat{\beta}_j^{(index)} \right) = \frac{1}{500} \sum_{k=1}^{500} \left[\hat{\beta}_j^{(index,k)} \right]^2 - \left[\hat{\beta}_j^{(index)} \right]^2$$

are reported below in tables.

We should also explain how the observations (which we contaminated) were selected and then contaminated. There are principally two possible approaches. The first one (which could be called "Huber's" one) can select randomly an apriori given percent of observations and from them create either outliers and/or leverage points. We obtain a sample of observation generated by some distribution function of "Huber's" type, namely

$$Q(x) = (1 - \delta) \cdot F(x) + \delta \cdot H(x), \quad \delta \in (0, 1.)$$

Consider for a while a creation of outliers. We can generate them in such a way that we multiply the response variable by some constant (see the heads of tables below). Then it can happen (and it do happen) that among the randomly selected observations some of them can have $X'_{it}\beta^0$ nearly equal to zero (in the case when the norm $\|X_{it}\|$ is small) and also the disturbance e_{it} can be small. But it results in the fact that the corresponding outlier does not represent (nearly) no problem for the estimator in question (because the observation is similar as the noncontaminated observations).

That is why we adopted an alternative way. We selected the observations from the generated sample according to their magnitude of $\|X_{it}\|$. If it overcome some threshold, the observation was contaminated. We believe that the contamination prepared in this way can cause the estimator really the problems. The snag however is that than the distribution of noncontaminated observations would be completely without the tails. Hence we adjusted the threshold a bit lower and part (approximately half of them) of observations with large $\| X_{it} \|$ were not contaminated. Then the threshold was adjusted so that in the mean we obtained the required level of contamination. The way how we ten created outliers and/or leverage points is given at the heads of tables.

Prior to collecting results for the contaminated data we offer a possibility of creating an idea how the robust procedures (here $\hat{\beta}^{(LWS,n,T,w)}$, $\hat{\beta}^{FWE}$ and $\hat{\beta}^{RWE}$) work for noncontaminated data. So, in other words, Table 1 addresses the problem of efficiency of robust estimation. In other words, we had assumed a contamination of data and hence we employed, except of the classical procedures, also $\hat{\beta}^{(LWS,n,T,w)}$, $\hat{\beta}^{FWE}$ and $\hat{\beta}^{RWE}$. We have started with rather high robustness of $\hat{\beta}^{(LWS,n,T,w)}$, $\hat{\beta}^{FWE}$ and $\hat{\beta}^{RWE}$ (i. e. the weight function had h and g near to $n/2$). But we found that the results of $\hat{\beta}^{(OLS,n,T)}$, $\hat{\beta}^{FE}$, $\hat{\beta}^{RE}$, $\hat{\beta}^{(LWS,n,T,w)}$, $\hat{\beta}^{FE}$ and $\hat{\beta}^{RE}$ practically coincide. Our study - due to the repetition of generating the data sets - allows to estimate the variance of all estimators and we found that the classical estimators were more efficient than the robust one. So we accommodated the weights a bit - for a lower level of contamination. But than we learned that again the estimated values are the same for the all estimators. So we assigned to the robust estimators weights only slightly depressing influence of observations with large residuals. For such an adjustment of weights the efficiency of all estimators are nearly the same (for the sake of space we present only results of this final experiment, with h and g near to n). (Due to the speed of the algorithm we can afford to do the same also for the real data without any problems.)

Now we are going to present results of simulations for the different situations - all necessary items are given at the heads of tables.

5 Conclusions

It is evident that even very low level of contamination (in the form of outliers as well as of leverage points) causes problems to $\hat{\beta}^{(OLS,n)}$ in efficiency and increasing level of contamination brings problems also in bias. Robustfication by $\hat{\beta}^{(LWS,n,T,w)}$ helps a lot and $\hat{\beta}^{FWE}$ and $\hat{\beta}^{RWE}$ attain an improvement even in efficiency of robust estimation. Notice however that improvement is not extremely significant. Moreover it is true only up to 12 % for contamination in the case of outliers and up to 8 % of contamination in the case of leverage points (please, see also the results given in the file on the address http://samba.fsv.cuni.cz/~visek/asdma2013/). For higher levels of contamination $\hat{\beta}^{FWE}$ and $\hat{\beta}^{RWE}$ even worsen the estimation and it is better to employ only $\hat{\beta}^{(LWS,n,T,w)}$. Further, the deviating from normal distribution of disturbances would probably worsen the situation - similarly as in the case of well-known results by (Fisher 1920). But it has to be confirmed by much more extended simulations.

We have also briefly addressed the problem of efficiency of robust estimation and optimality of the selection of weight function[5]. Firstly, we assumed high contamination and hence we tried to depress the influential points very resolutely. But the data were generated without any contamination. Our selection of weights resulted in the large variances of $\hat{\beta}^{(LWS,n,T,w)}$, $\hat{\beta}^{FWE}$ and $\hat{\beta}^{RWE}$ but the values of estimators were the same as values $\hat{\beta}^{(OLS,n,T)}$, $\hat{\beta}^{FE}$ and $\hat{\beta}^{RE}$. The fact that the values of non-robust estimators were not different from the values of robust estimators indicated the absence or at least low level of contamination. That was why we modified weights so that they depressed less a possible contamination but the situation repeated. So, we modified the weights once again, adjusting them for a very low contamination. For such an adjustment of weights the efficiency of both estimators is the same. Due to the speed of the algorithm we can afford it for the real data without any problems as well as in the simulation study. The complete collection of results of the above described numerical study can be seen in file which is on the address http://samba.fsv.cuni.cz/~visek/asdma2013/. One can find there the results for much more levels of contamination and hence it offers a possibility to create an idea about the scale of deviations of the estimator as a function of level of contamination.

As we have already mentioned some additional studies are to be made to map the deviations of the estimator from the assumed behavior when the underlying distribution of disturbances (and other variables) is not normal one, we should follow (Fisher 1920) or (Huber 1981). Moreover, the topic which was not probably addressed at all (even for the classical estimation) is the behavior of the estimators under heteroscedasticity (maybe some studies of a similar type as was performed in the pioneering paper by (White 1980) or in the paper by (Cragg 1983) showing how to employ the heteroscedasticity for an improvement of estimation). Similarly, it is with Hausman's test, see (Hausman 1978). A very first attempt - when considering the possibility of utilization of instrumental variables when the orthogonality condition is broken - was made in (Víšek 1998). Nevertheless, any consideration about robustification of Hausman's test in the sense in which it was treated in the paper by (Hausman 1978) or in (Hausman and Taylor 1981) (the latter one written with William Taylor) is still missing.

References

Amos CI (1994) Robust variance-components approach for assessing genetic linkage, in pedigrees. Am J Hum Genet 54:535–543

Atkinson AC, Riani M (2000) Robust diagnostic regresion analysis. Springer, New York

Bickel PJ (1975) One-step Huber estimates in the linear model. JASA 70:428–433

Bramati MC, Croux C (2007) Robust estimators for the fixed effects panel data model. Econ J 10:521–540

Boček P, Lachout P (1993) Linear programming approach to LMS-estimation. Memorial vol Comput Statist & Data Anal 19(1995):129–134

[5]We have even started with it, see Table 1, to attract also those readers who still believe in the statistical folklore that the sacrifice for robustness is a significant decrease of efficiency. In fact, the tax, we pay for the robustness, is typically more complicated proofs of ($\sqrt{n}$-)consistency, and asymptotic representations requiring more sophisticated tools. Moreover, the additional tax is a necessity to invent a new, quick an reliable algorithm for the computation of the respective estimator. Even the implementation need not be very simple. However, when we pay once these taxes and when the "product" is evaluated, what concerns of the quality by sufficiently reliable journals, the results are usually easy to use by anybody who is interested in.

Balestra P, Nerlove M (1966) Source pooling cross section and time series data in the estimation of a dynamic model: The demand for natural gas. Econometrica 34:585–612
Cameron AC, Gelbach JB, Miller DL (2011) Robust inference with multiway clustering. J Bus & Econ Stat 29:238–249
Cragg JG (1983) More efficient estimation in the presence of heteroscedasticity of unknown form. Econometrica 51:751–763
Croux C, Rousseeuw PJ (1992) A class of high-breakdown scale estimators based on subranges. Comm Statist- Theory Methods 21:1935–1951
Davies L (1990) The asymptotics of S-estimators in the linear regression model. Ann Statistics 18:1651–1675
Dehon C, Gassner M, Veradi V (2009) Beware of 'good' outliers and overoptimistic conclussios. Oxford Bull Econ Statist 71:437–452
Fisher RA (1920) A mathematical examination of the methods of determining the accuracy of an observation by the mean error and by the mean squares error. Mon Not Roy Astr Soc 80:758–770
Galton F (1886) Regression towards mediocrity in hereditary stature. J Anthropol Inst 15:246–263
Gill PS (2000) A robust mixed linear model analysis for longitudinal data. Stat Med 19:975?987
Hampel FR, Ronchetti EM, Rousseeuw PJ, Stahel WA (1986) Robust statistics – the approach based on influence functions. Wiley, New York
Hausman J (1978) Specification test in econometrics. Econometrica 46(1978):1251–1271
Hausman JA, Taylor WE (1981) Panel data and unobservable individual effects. Econometrica 49:1377–1398
Hettmansperger TP, Sheather SJ (1992) A cautionary note on the method of least median squares. Am Stat 46:79–83
Huber PJ (1981) Robust statistics. Wiley, New York
Jones DR (1992) Meta-analysis of observational epidemiological studies: A review. J R Soc Med 85:165–168
Judge GG, Griffiths WE, Hill RC, Lutkepohl H, Lee TC (1985) The Theory and Practice of Econometrics. Wiley, New York
Jurečková J, Sen PK (1984) On adaptive scale-equivariant M-estimators in linear models Statistics and Decisions 2 (1984), Suppl. Issue No. 1.
Koenker R, Bassett G (1978) Regression quantiles. Econometrica 46:33–50
Koenker R (2004) Quantile regression for longitudinal data. J Multivar Anal 91:74–89
Kott P (1989) Robust small domain estimation using random effects modelling. Survey Methodology 15:1–12
Maddala GS (1971) The use of variance components models in pooling cross section and time series data. Econometrica 39:341–358
Pinheiro JC, Liu C (2001) Efficient Algorithms for Robust Estimation, in Linear mixed-Effects models using the multivariate t-Distribution. J Comput Graph Stat 10:249–276
Portnoy S (1983) Tightness of the sequence of empiric c.d.f. processes defined from regression fractiles. In: Franke J, Hạ̈rdle W, Martin D (eds) Robust and nonlinear time-series analysis. Springer-Verlag, NY 1983, pp 231–246
Rocke DM (1991) Robustness and balance in the mixed models. Biometrics 47:303–309
Rousseeuw PJ (1984) Least median of square regression. JASA 79:871–880
Schall R (1991) Estimation in generalized linear models with random effects. Biometrika 72:593–9
Šindelář J (1994) On L-estimators viewed as M-estimators. Kybernetika 30(1994):5, 551–562
Telser LG (1964) Iterative Simultaneous Estimation of Sets of Linear Regressions. J Am Stat Assoc 59:845–62
Veradi V, Croux C (2009) Robust regression in Stata. Stat J 9(3):439–453
Veradi V, Wagner J (2010) Robust estimation of of linear fixed effects panel data model with an application to the exporter productivity premium Discussion paper No 4928, IZA, Bonn
Víšek JÁ (1996) A cautionary note on the method of Least Median of Squares reconsidered Trans Twelfth Prague Conf on Inform Theory, Statist Dec Functions and Rand Proc, Academy of Sciences of the Czech Republic, pp 254–259
Víšek JÁ (1996) Sensitivity analysis of M-estimates. Ann Inst Stat Math 48:469–495
Víšek JÁ (1998) Robust specification test. In: Marie Hušková, Petr Lachout &, Jan Ámos Víšek (eds) Proceedings of Prague Stochastics'98. Published by Union of Czechoslovak Mathematicians and Physicists, 1998, pp 581–586
Víšek JÁ (2000) Regression with high breakdown point. In: Antoch J &, Dohnal G (eds) Proceeding of the conference ROBUST 2000. Published by The Union of the Czech Mathematicians and Physicists 2001, pp 324–356
Víšek JÁ (2002) Sensitivity analysis of M-estimates of nonlinear regression model: Influence of data subsets. Ann Inst Stat Math 54:261–290

Víšek JÁ (2006) Instrumental Weighted Variables - algorithm. In: Rizzi A, Vichi M (eds) Proceedings of COMPSTAT 2006. Physica-Verlag (Springer Company) Heidelberg 2006, pp 777-786
Víšek JÁ (2006) The least trimmed squares. Sensitivity study. In: Hušková M. &, Janžura M (eds) Proc Prague Stochastics 2006. matfyzpress, pp 728–738
Víšek JÁ (2010) Weak $\sqrt{n}$-consistency of the least weighted squares under heteroscedasticity. Acta Universit Carolinae, Math 2/51:71–82
Víšek JÁ (2010) Robust error-term-scale estimate. IMS Collections. Nonparametrics and Robustness in Modern Statistical Inference and Time Series Analysis: Festschrift for Jana Jurečková, vol 7 (2010), pp 254–267
Víšek JÁ (2011) Consistency of the least weighted squares under heteroscedasticity. Kybernetika 47:179–206
Víšek JÁ (2011) Empirical distribution function under heteroscedasticity. Stat 45:497–508
Víšek JÁ (2011) Robust estimation under broken orthogonality condition. The lecture given in the Nuffield College, Oxford, November 25, 2011 (the paper is still in a form of preprint, presentation is available on http://samba.fsv.cuni.cz/visek/Oxford/)
Víšek JÁ (2013) Least weighted squares with constraints and under heteroscedasticity. Bulletin Czech Econ Soc 20(31):21–54
Víšek JÁ (2014) Asymptotic representation of the instrumental weighted variables - theory and practice. Part II - numerical study. Bulletin Czech Econ Soc 21(32):48–72
Víšek JÁ (2014) Instrumental weighted variables under heteroscedasticity. Part II. Numerical study. Submitted to Kybernetika
Wallace TD, Hussain A (1969) The Use of error components models in combining cross section with time series data. Econometrica 37:55–72
White H (1980) A heteroskedasticity-consistent covariance matrix estimator and a direct test for heteroscedasticity. Econometrica 48:817–838
Wooldridge JM (2006) Introductory Econometrics. A Modern Approach MIT Press Cambridge Massachusetts, second edition 2009
Zellner A (1962) An Efficient Method of Estimating Seemingly Unrelated Regressions and Tests for AggregationBias. J Am Stat Assoc 57:348–68
Zellner A (1965) Joint Estimation of relationships involving discrete random variables. Econ 33:382–394