Robust estimation of model with the fixed and random effects

Jan Ámos Víšek, *Charles University in Prague*, visek@fsv.cuni.cz

Abstract. Robustified estimation of the regression model is recalled, generalized for fixed and random effects and finally, the properties of this generalization are discussed. Patterns of the results of numerical study are presented and some conclusions implied by them given.

Keywords. Robust estimation, regression model, fixed and random effects, numerical study.

1 An introduction

It is nowadays commonly recognized that a few atypical observations can significantly distort the results of data analysis. That is why the statisticians have started to study the robust methods for the various frameworks. In the case of panel data with the fixed or with the random effects we can find several attempts to cope with the task, [5, 10, 12, 14, 15] and perhaps the best known [3]. They utilized mostly either the idea of M-estimators or the idea of highly robust methods as the Least Median of Squares or the Least Trimmed Squares. The former suffer by the lack of *scale-* and *regression-equivariance*. They require a studentization of residuals by a special estimate of standard deviation of the disturbances, see [2]. All these methods (may[1]) suffer by high sensitivity to the changes or to the deletion of data - even a very small change of one observation can turn the regression line about 90^o, see [16] or [23], compare also with [7]. We propose to utilize the method which rid of these disadvantages, i. e. the method is *scale-* and *regression-equivariant* and less sensitive to such a change of data. Moreover, it can be accommodated - by an appropriate selections of the weights, see the figures below - to the level and character of the contamination (see also a remark below).

2 Describing the framework

Let us denote by N the set of all positive integers and for any $p \in N$ let R^p be the p-dimensional Euclidean space. For any n, $T \in N$ and $\beta^0 \in R^p$ consider the linear regression model

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

where Y_{it}'s are the response variables, X_{it}'s are p-dimensional random explanatory variables, u_i's the effects and e_{it}'s the disturbances.

Definition 2.1. *If $cov(X_{itj}, u_i) = 0$ for all $i = 1, 2, ..., n$, $t = 1, 2, ..., T$ and $j = 1, 2, ..., p$, the model (1) is called the random effects model, otherwise we speak about the fixed effects model, see e. g. [8].*

[1] M-estimators in the case when ψ-function is discontinuous.

For any $\beta \in R^p$, $i \in N$ and $t \in \{1, 2, ..., T\}$ let us denote the residual of the (i,t)-th observation as

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

Denoting further $r^2_{(\ell)}(\beta)$ the ℓ-th order statistic among the squared residuals $r^2_{it}(\beta), i = 1, 2, ..., n$, $t = 1, 2, ..., T$, i. e. assuming that

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

let's recall the classical *Ordinary Least Squares* and their robust version, namely

Definition 2.2. *Let $w_\ell \in [0, 1]$, $\ell = 1, 2, ..., n \cdot T$ be weights. The estimators*

$$\hat{\beta}^{(OLS,n,T)} = \arg\min_{\beta \in R^p} \sum_{i=1}^{n} \sum_{t=1}^{T} r^2_{it}(\beta) \qquad \text{and} \qquad \hat{\beta}^{(LWS,n,T,w)} = \arg\min_{\beta \in R^p} \sum_{\ell=1}^{n \cdot T} w_\ell \, r^2_{(\ell)}(\beta) \tag{3}$$

are called the Ordinary Least Squares (OLS) and the Least Weighted Squares (LWS) estimator, respectively ([18]).

Remark 1 *Notice that in the definition of $\hat{\beta}^{(LWS,n,T,w)}$ the weights are assigned to the order statistics of squared residuals rather than to the squared residuals directly*[2]*. Further, let $h \in (\frac{n \cdot T}{2}, n \cdot T)$. Putting $w^{(1)}_h = 1$ and $w^{(1)}_\ell = 0$ otherwise and $w^{(2)}_\ell = 1, \ell = 1, 2, ..., h$ and $w^{(2)}_\ell = 0$ otherwise, we arrive at (to save the space write $\hat{\beta}^{(LWS,k)}$ instead of $\hat{\beta}^{(LWS,n,T,w^{(k)})}$, $k = 1, 2$)*

$$\hat{\beta}^{(LWS,1)} = \arg\min_{\beta \in R^p} r^2_{(h)}(\beta) = \hat{\beta}^{(LMS,n,T,h)} \text{ and } \hat{\beta}^{(LWS,2)} = \arg\min_{\beta \in R^p} \sum_{\ell=1}^{h} r^2_{(\ell)}(\beta) = \hat{\beta}^{(LTS,n,T,h)}$$

i. e. for such weights the Least Weighted Squares coincide with the Least Median of Squares ([13]) and with the Least Trimmed Squares ([6]), respectively. On the other hand, considering $w_\ell \equiv 1$ we have $\hat{\beta}^{(LWS,n,T,w)} = \hat{\beta}^{(OLS,n,T)}$. It hints that by selecting the appropriate weights we can reach high breakdown point (as $\hat{\beta}^{(LTS,n,T,h)}$), on one hand side, but also the efficiency of $\hat{\beta}^{(OLS,n,T)}$, on the other hand.

Similarly as $\hat{\beta}^{(LMS,n,T,h)}$ and as $\hat{\beta}^{(LTS,n,T,h)}$, $\hat{\beta}^{(LWS,n,T,w)}$ is scale- and regression-equivariant[3]*. Moreover, unlike $\hat{\beta}^{(LMS,n,T,h)}$, $\hat{\beta}^{(LTS,n,T,h)}$ or M-estimators (which can be extremely sensitive to even a slight shift or to the deletion of some observation(s), especially at the 'center of the main cloud of observations', see e. g. [23]), $\hat{\beta}^{(LWS,n,T,w)}$ has low sensitivity to the shift or to the deletion of an observation, see [17, 19, 21]. Last but not least, the flexibility of $\hat{\beta}^{(LWS,n,T,w)}$ (which is due to the flexibility of the weight function w) together with the speed of the program for its computing (see [20]) allows to accommodate the estimator to the level and also to the character of contamination of data and so to reach the maximal possible efficiency. The former, the adaption to the level of contamination, can be done as in [1] (see also[17]), in fact shifting the "rejection" point of the weight function to the right, see Figure 1 below. It increases simultaneously the efficiency of estimation. The latter, the maximal efficiency can be reached by changing the shape of the weight function.*

In what follows we shall restrict ourselves to the monotone weights $w_1 \geq w_2 \geq ... \geq w_{n \cdot T}$ and assume that the weights are generated by a monotone function w in the way $w_\ell = w\left(\frac{\ell - 1}{n \cdot T}\right)$.

[2]We can speak about an *implicit weighting*, hence the order of words in *LWS*. Similar phenomenon is indicated by the order of words in the title of robust estimator of the *fixed* and of the *random effects models* below.

[3]It is advantage when comparing the *Least Weighted Squares* with M-estimators. The latter need - to reach *scale-* and *regression-equivariance* - studentization of residuals by an estimate of standard deviation which is *scale-equivariant* and *regression-invariant*, see [2], [4], [9] or [22].

Remark 2 *Denoting the empirical distribution function of the absolute values of residuals $r_{it}(\beta)$'s*

$$F_\beta^{(n)}(r) = \frac{1}{n \cdot T} \sum_{i=1}^{n} \sum_{t=1}^{T} I\left\{|r_{it}(\beta)| < r\right\},$$

it is only a technical task to show that $\hat{\beta}^{(LWS,n,T,w)}$ is one of solutions of the normal equations

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

Conditions C 1 *The sequence $\left\{\left\{(X_{it}', e_{it})'\right\}_{t=1}^{T}\right\}_{i=1}^{\infty}$ is sequence T-tuples of the mutually independent $(p+1)$-dimensional random variables (r.v.'s) distributed according to the distribution functions (d.f.) $F_{X,e_{it}}(x,r) = F_X(x) \cdot F_{e_{it}}(r)$ where $F_{e_{it}}(r) = F_e(r\sigma_{it}^{-1})$ with $I\!\!E e_{it} = 0$, $var(e_{it}) = \sigma_{it}^2$ and $0 < \liminf_{i\to\infty} \min_{1\le t\le T} \sigma_{it} \le \limsup_{i\to\infty} \max_{1\le t\le T} \sigma_{it} < \infty$[4]. Moreover, $F_e(r)$ is absolutely continuous with bounded density $f_e(r)$. Further, there is $q > 1$ so that $I\!\!E \left\|X_1\right\|^{2q} < \infty$. Finally, $\{u_t\}_{t=1}^{\infty}$ is a sequence of independent and identically distributed r. v.'s with d. f. $F_u(u)$ with finite σ_u^2. This sequence is independent from the sequence $\left\{\{e_{it}\}_{t=1}^{T}\right\}_{i=1}^{\infty}$.*

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

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

$$I\!\!E\left\{\sum_{i=1}^{n} \sum_{t=1}^{T} \left[w\left(F_\beta^{(it)}(|r_{it}(\beta)|)\right) X_{it}\left(Y_{it} - X_{it}'\beta\right)\right]\right\} = 0 \tag{5}$$

namely β^0. Moreover $\lim_{n\to\infty} \frac{1}{n \cdot T} \sum_{i=1}^{n} \sum_{t=1}^{T} \sigma_{it} = 1$.

Theorem 2.3. *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 equations (4) is weakly consistent.*

The proof is a direct reformulation of the result [23] from the *cross-sectional-data framework* to the *panel-data framework* (possibly with the fixed or with the random effects, see also (6) below). The result in [23] is based on a generalization of the classical Kolmogorov-Smirnov result on uniform convergence of empirical d. f. for the *regression model framework* (see [24]) which claims that under some mild conditions[5]

$$\sup_{-\infty<r<\infty} \sqrt{n} \left| F_\beta^{(n)}(r) - (n \cdot T)^{-1} \sum_{i=1}^{n} \sum_{t=1}^{T} F_\beta^{(it)}(r) \right| = O_p(1).$$

The applicability of Theorem 2.3 for the *fixed effect model* requires to transform data by the *within-group transformation*, see the discussion on the robustified version of (8) below.

3 Estimating the coefficients of model with the effects

The estimation of the coefficients of regression model with the fixed or with the random effects is a straightforward task. Let's start with the random effects and assume that the sequence

[4] Notice that $F_X(x)$ does not depend on i and t, i. e. the sequence $\left\{\{X_{it}\}_{t=1}^{T}\right\}_{i=1}^{\infty}$ is sequence of independent and identically distributed (i.i.d.) T-tuples of $p+1$-dimensional r.v.'s.

[5] The main technical tool for the proof was the Skorohod embedding into Wiener process, see [11].

$\{u_t\}_{t=1}^{\infty}$ is also independent from the sequence $\left\{\{X_{it}\}_{t=1}^{T}\right\}_{i=1}^{\infty}$. We can write the model (1) in the form

$$Y_{it} = X_{it}'\beta^0 + v_{it}, \quad i = 1, 2, ..., n, \quad t = 1, 2, ..., T \tag{6}$$

where $v_{it} = u_i + e_{it}$, $i = 1, 2, ..., n$, $t = 1, 2, ..., T$. Then we have

$$I\!\!E v_{it} = 0, \quad \text{cov}(X_{itj}, v_{it}) = 0 \quad \text{and} \quad I\!\!E\,[v_{it}, v_{is}] = \text{var}(u_i) = \sigma_u^2$$

and hence $\hat{\beta}^{(OLS,n,T)}$ (even under normality of disturbances) is not efficient due to the correlation between disturbances. The performance of analysis can be improved either by utilizing the *Generalized Least Squares* ($\hat{\beta}^{(GLS,n,T)}$, say) or - equivalently - by considering slightly modified data

$$\tilde{Y}_{it} = Y_{it} - \lambda \overline{Y}_i \quad \text{and} \quad \tilde{X}_{it} = X_{it} - \lambda \overline{X}_i, \quad \text{with} \quad \lambda = 1 - \sigma_e^2 \cdot \left(\sigma_e^2 + T \cdot \sigma_u^2\right)^{-1}, \tag{7}$$

(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. [25]. The variances σ_e^2 and σ_u^2 can be estimated employing $r_{it}(\hat{\beta}^{(OLS,n,T)})$, taking into account that $\sigma_e^2 = \sigma_v^2 - \sigma_u^2$ and then we can utilize $\hat{\lambda}$ instead of λ. Applying then *OLS* on the transformed data we obtain *RE-estimate* (which is below in tables denoted as $\hat{\beta}^{RE}$).

For $\hat{\beta}^{(LWS,n,T,w)}$ we are not able to demonstrate analytically an improvement when considering transformed data (7) but a numerical study (patterns of which are given bellow) indicates an extent of improvement. The estimates of variances σ_v^2 and σ_u^2 were obtained also by means of the *Least Weighted Squares* (for one-dimensional case), $\hat{\sigma}^2_{LWS,e}$ and $\hat{\sigma}^2_{LWS,u}$ (say), employing the same weight function as for $\hat{\beta}^{(LWS,n,T,w)}$ (its optimality is discussed below). Such estimators are scale-equivariant and regression-invariant. Then we established $\hat{\lambda}^{(LWS)}$ and transformed data as in (7). Finally, $\hat{\beta}^{(LWS,n,T,w)}$ was applied. Such an estimator is denoted in the tables below as the *Random Weighted Effects* estimator $\hat{\beta}^{RWE}$. The resulting estimator is then scale- and regression-invariant (see **Remark 1)**. So we can conclude that the robustification of estimation of the random effects model was straightforward.

Now, let's turn to the estimation of model with fixed effects. It will require a bit more explanation. We can again utilize model (6) but then we have

$$I\!\!E v_{it} = 0 \quad \text{but} \quad \text{cov}(X_{itj}, v_{it}) \neq 0$$

and hence $\hat{\beta}^{(OLS,n,T)}$ is generally biased and inconsistent. There are several ways how to cope with this situation, e. g. by differencing data. However *fixed-effect-estimation* (or alternatively called the *within-group transformation*) seems to be more attractive. It considers data

$$\tilde{Y}_{it} = Y_{it} - \overline{Y}_i \;\text{ and }\; \tilde{X}_{it} = X_{it} - \overline{X}_i, \;\text{ where again } \overline{Y}_i = \frac{1}{T}\sum_{t=1}^{T} Y_{it} \;\text{ and }\; \overline{X}_i = \frac{1}{T}\sum_{t=1}^{T} X_{it}. \tag{8}$$

Notice that the first coordinate of the vectors $\tilde{X}_{it}$ are equal to zero. Hence putting $V_{itj} = \tilde{X}_{itj+1}$ 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

$$\tilde{Y}_{it} = V_{it}'\gamma^0 + \tilde{e}_{it}, \quad i = 1, 2, ..., n, \quad t = 1, 2, ..., T, \tag{9}$$

and since $I\!\!E[V_{itj} \cdot \tilde{e}_{it}] = 0$, the least squares estimator $\hat{\gamma}^{(OLS,n,T)}$ is (under some standard conditions) unbiased, consistent, etc., under normality of e_{it}'s (and hence also normality of $\tilde{e}_{it}$'s) even reaching Rao-Cramer lower bound for covariance matrix (if we consider ordering by means of positive definiteness). A minor disadvantages is that we cannot estimate (directly) the intercept of the model (1) (but sometimes the estimate $\hat{\beta}_1^{(OLS,n,T)} = \frac{1}{n}\sum_{i=1}^{n}\left[\overline{Y}_i - \overline{X}_i'\hat{\gamma}^{(OLS,n,T)}\right]$ is recommended, being supported by $\frac{1}{n \cdot T}\sum_{i=1}^{n}\sum_{t=1}^{T} r_{it}(\hat{\gamma}^{(OLS,n,T)}) \to 0$ a.s.).

So, an idea for a robustification of the estimation seems to be again straightforward - having estimated robustly mean values of response and explanatory variables (utilizing again the idea of *LWS* again for one-dimensional case), then employing transformation (8) with these estimates instead of with $\overline{Y}_i$'s and $\overline{X}_i$'s and finally utilizing $\hat{\gamma}^{(LWS,n,T,w)}$ (i. e. analogy of $\hat{\beta}^{(LWS,n,T,w)}$ for the model (9)) we compute the *Fixed Weighted Effects Estimator* (denoted below as $\hat{\beta}^{FWE}$). The resulting estimator is again - due to the fact that we have utilized in all steps LWS - scale- and regression-invariant (see again **Remark 1**).

4 Numerical study

Selecting $n = 50$, $T = 20$ and $p = 5$, we have generated two groups, containing 1000 sets, as follows

$$\left\{ \left\{ \left\{ X_{it}^{(k)}, u_i^{(k)} \right\}_{t=1}^{20} \right\}_{i=1}^{50}, \left\{ \varepsilon_{it}^{(k)} \right\}_{i=1}^{50} \right\}_{k=1}^{1000} \quad \text{and} \quad \left\{ \left\{ \left\{ \sigma_{it}^{(k)} \right\}_{t=1}^{20} \right\}_{i=1}^{50} \right\}_{k=1}^{1000}$$

where all variables in the first group were distributed according to the standard normal distribution while in the second group according to the uniform distribution on $[0.5, 1.5]$. Then, putting for $\tilde{X}_{itj}^{(k)} = X_{itj}^{(k)} + u_i^{(k)}, e_{it}^{(k)} = \varepsilon_{it}^{(k)} \cdot \sigma_{it}^{(k)}$ for $i = 1, 2, ..., 50$ $t = 1, 2, ..., 20$ and $j = 1, 2, ..., 5$, we created according to (1) two groups (of course, for the second one we employed $\tilde{X}_{it}^{(k)}$'s instead of $X_{it}^{(k)}$'s)

$$\left\{ \left\{ \left\{ Y_{it}^{(k)}, X_{it}^{(k)} \right\}_{t=1}^{20} \right\}_{i=1}^{50} \right\}_{k=1}^{1000} \quad \text{and} \quad \left\{ \left\{ \left\{ \tilde{Y}_{it}^{(k)}, \tilde{X}_{it}^{(k)} \right\}_{t=1}^{20} \right\}_{i=1}^{50} \right\}_{k=1}^{1000}.$$

It is clear that the first group represents the model with random effects, the second one the model with the fixed effects[6]. Then the estimates of regression coefficients were computed[7] - 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. the estimator proposed for the model with the fixed effects works when estimating the coefficients of the model with random effects etc. . So, we obtained, say

$$\left\{ \hat{\beta}^{(index,k)} = (\hat{\beta}_1, \hat{\beta}_2, ..., \hat{\beta}_5)' \right\}_{k=1}^{1000}$$

where the abbreviations *OLS, RE, FE, LWS, RWE* and *FWE* at the position of *"index"* indicate the method employed for the computation, namely *OLS* for the *Ordinary Least Squares*, *RE* - for the *Random Effects*, *FE* - for the *Fixed Effects*, *LWS* - for the *Least Weighted Squares*, *RWE* - for the *Random Weighted Effects* and finally *FWE* for the *Fixed Weighted Effects.*

The weight function $w(r) : [0, 1] \to [0, 1]$ was constructed so to generate the weights as follows:

$$w_\ell = 1 \quad \text{for} \quad \ell = 1, 2, ..., h, \quad h \in \{1, 2, ..., n \cdot T\},$$
$$w_\ell = 0 \quad \text{for} \quad \ell = g, g + 1, ..., n \cdot T, \quad g \in \{1, 2, ..., n \cdot T\}, \ g > h$$

and for $h \le \ell \le g$ the weights w_ℓ's decreased linearly from 1 to 0 (so that $w_h = 1$ and $w_g = 0$).

Further, the empirical means and empirical variances of estimates of coefficients (over these 1000 repetitions) were computed, i. e. we report values (for $j = 1, 2, 3, 4$ and 5)

$$\hat{\beta}_j^{(index)} = \frac{1}{1000} \sum_{k=1}^{1000} \hat{\beta}_j^{(index,k)} \quad \text{and} \quad \widehat{\text{var}} \left(\hat{\beta}_j^{(index)} \right) = \frac{1}{1000} \sum_{k=1}^{1000} \left[\hat{\beta}_j^{(index,k)} \right]^2 - \left[\hat{\beta}_j^{(index)} \right]^2.$$

[6]Please, realize that the second model, based on $\tilde{Y}_{it}^{(k)}$'s and $\tilde{X}_{it}^{(k)}$'s, "contains" $u_i^{(k)}$'s "two times". $u_i^{(k)}$'s are utilized when creating $\tilde{X}_{it}^{(k)}$'s and they are second times employed when generating $\tilde{Y}_{it}^{(k)}$'s according to (1).

[7]We have started with 2.5% of contamination created by outliers - see below the discussion about the optimality of the weight function w.

Now the optimality of the weights w_ℓ's was searched. It was done for the level of contamination equal to 2.5%. The following sum of two statistics, namely cumulative mean absolute bias of estimates and their cumulative empirical variances

$$\frac{1}{5000}\sum_{k=1}^{1000}\sum_{j=1}^{5}|\hat{\beta}_j^{(index,k)} - \hat{\beta}_j^0| \;+\; \frac{1}{5}\sum_{j=1}^{5}\widehat{\text{var}}\left(\hat{\beta}_j^{(index)}\right),$$

were employed and we searched for its minimality. Although we would probably expect - due to the level of contamination - the weight function of the form given in the left-hand-side of Figure 1, the optimal is that one on the right-hand-side. As the empirically found optimal weight function weights down the contamination very gradually, we can hope that it is approximately optimal for other levels of contamination as well (at least for higher levels). Hence the estimates of regression coefficients for other levels of contamination were computed employing the same weight function.

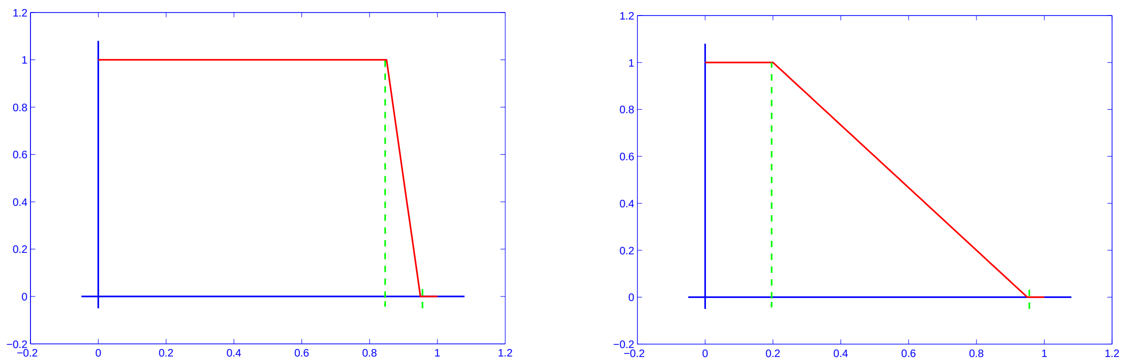

Figure 1. The intuitively optimal (left) and really optimal (right) weight function

Let us give now the (patterns of) results of this numerical study[8].

TABLE 1

True coeffs β^0	**3.1**	**1.1**	**2.3**	**-3.4**	**6.5**
These coefficients were employed in the whole numerical study.					

The disturbances are heteroscedastic. Both, the disturbances and the effects, are independent from explanatory variables. Contamination is created by outliers, values of which are equal to the 10th multiple of original values of response variable.					
Contamination level is equal to 1%.					
$\hat{\beta}^{OLS}(\mathbf{var}(\hat{\beta}^{OLS}))$	$3.41_{(0.120)}$	$1.55_{(0.570)}$	$3.18_{(0.608)}$	$\mathbf{-4.72}_{(0.761)}$	$9.03_{(1.271)}$
$\hat{\beta}^{FE}$ $(\mathbf{var}(\hat{\beta}^{FE}))$		$1.08_{(0.238)}$	$2.23_{(0.248)}$	$-3.29_{(0.227)}$	$6.30_{(0.250)}$
$\hat{\beta}^{RE}$ $(\mathbf{var}(\hat{\beta}^{RE}))$	$3.41_{(0.120)}$	$1.55_{(0.571)}$	$3.18_{(0.609)}$	$-4.72_{(0.763)}$	$9.03_{(1.272)}$
$\hat{\beta}^{LWS}(\mathbf{var}(\hat{\beta}^{LWS}))$	$3.09_{(0.018)}$	$1.10_{(0.010)}$	$2.30_{(0.011)}$	$-3.40_{(0.011)}$	$6.50_{(0.011)}$
$\hat{\beta}^{FWE}(\mathbf{var}(\hat{\beta}^{FWE}))$		$1.10_{(0.017)}$	$2.28_{(0.018)}$	$-3.38_{(0.018)}$	$6.50_{(0.018)}$
$\hat{\beta}^{RWE}(\mathbf{var}(\hat{\beta}^{RWE}))$	$3.09_{(0.086)}$	$1.10_{(0.004)}$	$2.30_{(0.004)}$	$-3.40_{(0.004)}$	$6.50_{(0.004)}$
Contamination level is equal to 2.5%.					
$\hat{\beta}^{OLS}(\mathbf{var}(\hat{\beta}^{OLS}))$	$3.85_{(0.209)}$	$1.95_{(0.971)}$	$4.05_{(1.119)}$	$\mathbf{-6.09}_{(1.303)}$	$11.91_{(2.056)}$
$\hat{\beta}^{FE}$ $(\mathbf{var}(\hat{\beta}^{FE}))$		$1.07_{(0.143)}$	$2.21_{(0.134)}$	$-3.29_{(0.129)}$	$6.28_{(0.145)}$
$\hat{\beta}^{RE}$ $(\mathbf{var}(\hat{\beta}^{RE}))$	$3.85_{(0.208)}$	$1.95_{(0.971)}$	$4.05_{(1.120)}$	$-6.09_{(1.303)}$	$11.61_{(2.051)}$

[8] The limited space, we could devote to the results of numerical study, does not allow for presenting more results of numerical study but they can be found on http://samba.fsv.cuni.cz/~visek/FixedRandomEffects/.

$\hat\beta^{LWS}_{(\mathrm{var}(\hat\beta^{LWS}))}$	$3.10_{(0.016)}$	$1.10_{(0.011)}$	$2.30_{(0.011)}$	$-3.40_{(0.012)}$	$6.50_{(0.011)}$
$\hat\beta^{FWE}_{(\mathrm{var}(\hat\beta^{FWE}))}$		$1.01_{(0.016)}$	$2.27_{(0.018)}$	$-3.37_{(0.019)}$	$6.48_{(0.018)}$
$\hat\beta^{RWE}_{(\mathrm{var}(\hat\beta^{RWE}))}$	$3.06_{(3.071)}$	$1.10_{(0.003)}$	$2.30_{(0.003)}$	$-3.40_{(0.003)}$	$6.51_{(0.003)}$
Contamination level is equal to 6%.					
$\hat\beta^{OLS}_{(\mathrm{var}(\hat\beta^{OLS}))}$	$4.82_{(0.406)}$	$2.79_{(1.453)}$	$5.84_{(1.756)}$	$-8.67_{(1.919)}$	$16.6_{(3.409)}$
$\hat\beta^{FE}_{(\mathrm{var}(\hat\beta^{FE}))}$		$1.06_{(0.073)}$	$2.21_{(0.073)}$	$-3.27_{(0.069)}$	$6.26_{(0.076)}$
$\hat\beta^{RE}_{(\mathrm{var}(\hat\beta^{RE}))}$	$4.82_{(0.406)}$	$2.79_{(1.499)}$	$5.84_{(1.759)}$	$-8.67_{(1.924)}$	$16.6_{(3.400)}$
$\hat\beta^{LWS}_{(\mathrm{var}(\hat\beta^{LWS}))}$	$3.10_{(0.016)}$	$1.10_{(0.011)}$	$2.30_{(0.010)}$	$-3.40_{(0.011)}$	$6.50_{(0.011)}$
$\hat\beta^{FWE}_{(\mathrm{var}(\hat\beta^{FWE}))}$		$1.09_{(0.015)}$	$2.27_{(0.016)}$	$-3.38_{(0.018)}$	$6.48_{(0.017)}$
$\hat\beta^{RWE}_{(\mathrm{var}(\hat\beta^{RWE}))}$	$1.87_{(2.705)}$	$1.12_{(0.014)}$	$2.34_{(0.015)}$	$-3.46_{(0.015)}$	$6.62_{(0.016)}$

TABLE 2

The disturbances are again heteroscedastic. Both, the disturbances and the effects, are independent from explanatory variables. Contamination is created by leverage points, values of which are equal to the 10th multiple of original values of explanatory variable.					
Contamination level is equal to 1%					
$\hat\beta^{OLS}_{(\mathbf{var}(\hat\beta^{OLS}))}$	$1.75_{(0.266)}$	$0.28_{(0.389)}$	$0.62_{(0.383)}$	$-0.85_{(0.487)}$	$1.65_{(0.724)}$
$\hat\beta^{FE}_{(\mathbf{var}(\hat\beta^{FE}))}$		$0.27_{(0.248)}$	$0.57_{(0.253)}$	$-0.82_{(0.333)}$	$1.55_{(0.614)}$
$\hat\beta^{RE}_{(\mathbf{var}(\hat\beta^{RE}))}$	$0.98_{(0.817)}$	$0.29_{(0.339)}$	$0.63_{(0.345)}$	$-0.88_{(0.446)}$	$1.70_{(0.707)}$
$\hat\beta^{LWS}_{(\mathbf{var}(\hat\beta^{LWS}))}$	$3.10_{(0.016)}$	$1.10_{(0.020)}$	$2.30_{(0.010)}$	$-.40_{(0.011)}$	$6.50_{(0.010)}$
$\hat\beta^{FWE}_{(\mathbf{var}(\hat\beta^{FWE}))}$		$1.09_{(0.019)}$	$2.28_{(0.020)}$	$-3.37_{(0.020)}$	$6.47_{(0.020)}$
$\hat\beta^{RWE}_{(\mathbf{var}(\hat\beta^{RWE}))}$	$3.08_{(0.261)}$	$1.10_{(0.004)}$	$2.30_{(0.004)}$	$-3.39_{(0.004)}$	$6.49_{(0.004)}$
Contamination level is equal to 2.5%					
$\hat\beta^{OLS}_{(\mathbf{var}(\hat\beta^{OLS}))}$	$0.94_{(0.056)}$	$0.12_{(0.048)}$	$0.27_{(0.053)}$	$-0.37_{(0.058)}$	$0.72_{(0.077)}$
$\hat\beta^{FE}_{(\mathbf{var}(\hat\beta^{FE}))}$		$0.12_{(0.024)}$	$0.25_{(0.029)}$	$-0.35_{(0.032)}$	$0.68_{(0.057)}$
$\hat\beta^{RE}_{(\mathbf{var}(\hat\beta^{RE}))}$	$0.16_{(0.093)}$	$0.13_{(0.031)}$	$0.27_{(0.037)}$	$-0.38_{(0.040)}$	$0.73_{(0.069)}$
$\hat\beta^{LWS}_{(\mathbf{var}(\hat\beta^{LWS}))}$	$3.10_{(0.016)}$	$1.10_{(0.011)}$	$2.30_{(0.011)}$	$-3.40_{(0.012)}$	$6.50_{(0.010)}$
$\hat\beta^{FWE}_{(\mathbf{var}(\hat\beta^{FWE}))}$		$1.09_{(0.034)}$	$2.26_{(0.036)}$	$-3.37_{(0.037)}$	$6.47_{(0.032)}$
$\hat\beta^{RWE}_{(\mathbf{var}(\hat\beta^{RWE}))}$	$2.95_{(0.728)}$	$1.09_{(0.015)}$	$2.25_{(0.015)}$	$-3.33_{(0.016)}$	$6.37_{(0.022)}$
Contamination level is equal to 6%					
$\hat\beta^{OLS}_{(\mathbf{var}(\hat\beta^{OLS}))}$	$0.44_{(0.008)}$	$0.06_{(0.005)}$	$0.11_{(0.005)}$	$-0.17_{(0.006)}$	$0.32_{(0.008)}$
$\hat\beta^{FE}_{(\mathbf{var}(\hat\beta^{FE}))}$		$0.52_{(0.003)}$	$0.11_{(0.003)}$	$-0.16_{(0.004)}$	$0.30_{(0.006)}$
$\hat\beta^{RE}_{(\mathbf{var}(\hat\beta^{RE}))}$	$-0.13_{(0.006)}$	$0.05_{(0.002)}$	$0.12_{(0.003)}$	$-0.17_{(0.003)}$	$0.32_{(0.006)}$
$\hat\beta^{LWS}_{(\mathbf{var}(\hat\beta^{LWS}))}$	$3.10_{(0.015)}$	$1.10_{(0.013)}$	$2.30_{(0.012)}$	$-3.40_{(0.013)}$	$6.50_{(0.01)}$
$\hat\beta^{FWE}_{(\mathbf{var}(\hat\beta^{FWE}))}$		$1.07_{(0.155)}$	$2.25_{(0.143)}$	$-3.33_{(0.163)}$	$6.41_{(0.128)}$
$\hat\beta^{RWE}_{(\mathbf{var}(\hat\beta^{RWE}))}$	$0.73_{(4.565)}$	$0.87_{(0.152)}$	$1.86_{(0.214)}$	$-2.72_{(0.260)}$	$5.22_{(0.548)}$

TABLE 3

The disturbances are again heteroscedastic. The disturbances are independent from explanatory variables but the effects are correlated with them. Contamination is created by outliers, values of which are equal to the 10th of values of response variable.					
Contamination level is equal to 1 %					
$\hat\beta^{OLS}_{(\mathbf{var}(\hat\beta^{OLS}))}$	$3.62_{(1.623)}$	$2.72_{(1.560)}$	$4.16_{(2.018)}$	$-2.52_{(0.811)}$	$9.06_{(3.141)}$

$\hat{\beta}^{FE}$ $(\mathrm{var}(\hat{\beta}^{FE}))$		$2.56_{(1.117)}$	$3.99_{(1.490)}$	$-2.72_{(0.516)}$	$8.89_{(2.615)}$
$\hat{\beta}^{RE}$ $(\mathrm{var}(\hat{\beta}^{RE}))$	$3.62_{(1.601)}$	$2.52_{(1.008)}$	$3.95_{(1.309)}$	$-2.75_{(0.448)}$	$8.85_{(2.433)}$
$\hat{\beta}^{LWS}$ $(\mathrm{var}(\hat{\beta}^{LWS}))$	$3.10_{(0.002)}$	$1.30_{(0.003)}$	$2.50_{(0.003)}$	$-3.20_{(0.003)}$	$6.70_{(0.002)}$
$\hat{\beta}^{FWE}$ $(\mathrm{var}(\hat{\beta}^{FWE}))$		$1.10_{(0.003)}$	$2.30_{(0.003)}$	$-3.40_{(0.003)}$	$6.50_{(0.003)}$
$\hat{\beta}^{RWE}$ $(\mathrm{var}(\hat{\beta}^{RWE}))$	$3.11_{(1936.753)}$	$1.13_{(0.003)}$	$2.33_{(0.003)}$	$-3.37_{(0.003)}$	$6.53_{(0.003)}$
Contamination level is equal to 2.5%					
$\hat{\beta}^{OLS}$ $(\mathrm{var}(\hat{\beta}^{OLS}))$	$3.78_{(2.233)}$	$3.50_{(1.973)}$	$5.15_{(2.601)}$	$-2.39_{(1.169)}$	$10.75_{(5.112)}$
$\hat{\beta}^{FE}$ $(\mathrm{var}(\hat{\beta}^{FE}))$		$3.48_{(1.639)}$	$5.13_{(2.143)}$	$-2.41_{(0.901)}$	$10.76_{(4.501)}$
$\hat{\beta}^{RE}$ $(\mathrm{var}(\hat{\beta}^{RE}))$	$3.81_{(2.316)}$	$3.34_{(1.471)}$	$4.99_{(2.093)}$	$-2.54_{(0.771)}$	$10.62_{(4.341)}$
$\hat{\beta}^{LWS}$ $(\mathrm{var}(\hat{\beta}^{LWS}))$	$3.11_{(0.002)}$	$1.30_{(0.003)}$	$2.50_{(0.003)}$	$-3.20_{(0.003)}$	$6.69_{(0.003)}$
$\hat{\beta}^{FWE}$ $(\mathrm{var}(\hat{\beta}^{FWE}))$		$1.10_{(0.002)}$	$2.31_{(0.002)}$	$-3.40_{(0.002)}$	$6.49_{(0.002)}$
$\hat{\beta}^{RWE}$ $(\mathrm{var}(\hat{\beta}^{RWE}))$	$3.14_{(1974.513)}$	$1.11_{(0.003)}$	$2.31_{(0.003)}$	$-3.39_{(0.003)}$	$6.50_{(0.002)}$
Contamination level is equal to 6 %					
$\hat{\beta}^{OLS}$ $(\mathrm{var}(\hat{\beta}^{OLS}))$	$5.04_{(2.898)}$	$5.67_{(2.352)}$	$7.81_{(3.316)}$	$-2.67_{(1.904)}$	$15.27_{(7.250)}$
$\hat{\beta}^{FE}$ $(\mathrm{var}(\hat{\beta}^{FE}))$		$5.60_{(2.229)}$	$7.74_{(3.038)}$	$-2.65_{(1.560)}$	$15.25_{(7.038)}$
$\hat{\beta}^{RE}$ $(\mathrm{var}(\hat{\beta}^{RE}))$	$5.04_{(2.887)}$	$5.49_{(1.960)}$	$7.62_{(2.790)}$	$-2.77_{(1.383)}$	$15.12_{(6.770)}$
$\hat{\beta}^{LWS}$ $(\mathrm{var}(\hat{\beta}^{LWS}))$	$3.10_{(0.002)}$	$1.29_{(0.002)}$	$2.50_{(0.003)}$	$-3.20_{(0.003)}$	$6.70_{(0.002)}$
$\hat{\beta}^{FWE}$ $(\mathrm{var}(\hat{\beta}^{FWE}))$		$1.16_{(0.006)}$	$2.37_{(0.008)}$	$-3.34_{(0.006)}$	$6.57_{(0.005)}$
$\hat{\beta}^{RWE}$ $(\mathrm{var}(\hat{\beta}^{RWE}))$	$3.06_{(1957.953)}$	$1.02_{(0.007)}$	$2.23_{(0.007)}$	$-3.48_{(0.007)}$	$6.44_{(0.006)}$

TABLE 4

The disturbances are again heteroscedastic. The disturbances are independent from explanatory variables but the effects are correlated with them. Contamination by leverage points, values of which are equal to the 10th of original values of explanatory variables.					
Contamination level is equal to 1 %					
$\hat{\beta}^{OLS}$ $(\mathbf{var}(\hat{\beta}^{OLS}))$	$1.99_{(0.998)}$	$0.06_{(0.594)}$	$0.64_{(0.742)}$	$-2.12_{(1.030)}$	$2.77_{(2.018)}$
$\hat{\beta}^{FE}$ $(\mathbf{var}(\hat{\beta}^{FE}))$		$-0.29_{(0.357)}$	$0.35_{(0.372)}$	$-2.40_{(1.130)}$	$2.40_{(1.619)}$
$\hat{\beta}^{RE}$ $(\mathbf{var}(\hat{\beta}^{RE}))$	$0.50_{(5.587)}$	$-0.18_{(0.393)}$	$0.51_{(0.464)}$	$-2.40_{(1.092)}$	$2.64_{(1.845)}$
$\hat{\beta}^{LWS}$ $(\mathbf{var}(\hat{\beta}^{LWS}))$	$3.10_{(0.002)}$	$1.29_{(0.003)}$	$2.50_{(0.003)}$	$-3.21_{(0.003)}$	$6.70_{(0.003)}$
$\hat{\beta}^{FWE}$ $(\mathbf{var}(\hat{\beta}^{FWE}))$		$1.09_{(0.003)}$	$2.29_{(0.003)}$	$-3.41_{(0.003)}$	$6.49_{(0.003)}$
$\hat{\beta}^{RWE}$ $(\mathbf{var}(\hat{\beta}^{RWE}))$	$3.14_{(2049.653)}$	$1.12_{(0.003)}$	$2.33_{(0.003)}$	$-3.37_{(0.003)}$	$6.53_{(0.003)}$
Contamination level is equal to 2.5%					
$\hat{\beta}^{OLS}$ $(\mathbf{var}(\hat{\beta}^{OLS}))$	$1.50_{(0.792)}$	$0.09_{(0.304)}$	$0.42_{(0.384)}$	$-1.51_{(0.733)}$	$1.81_{(1.023)}$
$\hat{\beta}^{FE}$ $(\mathbf{var}(\hat{\beta}^{FE}))$		$-0.17_{(0.184)}$	$0.19_{(0.192)}$	$-1.70_{(0.810)}$	$1.52_{(0.797)}$
$\hat{\beta}^{RE}$ $(\mathbf{var}(\hat{\beta}^{RE}))$	$0.21_{(2.571)}$	$-0.10_{(0.210)}$	$0.27_{(0.246)}$	$-1.74_{(0.800)}$	$1.67_{(0.980)}$
$\hat{\beta}^{LWS}$ $(\mathbf{var}(\hat{\beta}^{LWS}))$	$3.10_{(0.002)}$	$1.29_{(0.002)}$	$2.50_{(0.003)}$	$-3.20_{(0.002)}$	$6.70_{(0.003)}$
$\hat{\beta}^{FWE}$ $(\mathbf{var}(\hat{\beta}^{FWE}))$		$1.07_{(0.003)}$	$2.28_{(0.004)}$	$-3.42_{(0.004)}$	$6.47_{(0.005)}$
$\hat{\beta}^{RWE}$ $(\mathbf{var}(\hat{\beta}^{RWE}))$	$3.13_{(1967.989)}$	$1.10_{(0.005)}$	$2.31_{(0.004)}$	$-3.40_{(0.005)}$	$6.50_{(0.006)}$
Contamination level is equal to 6 %					
$\hat{\beta}^{OLS}$ $(\mathbf{var}(\hat{\beta}^{OLS}))$	$0.65_{(0.131)}$	$0.00_{(0.038)}$	$0.18_{(0.042)}$	$-0.65_{(0.120)}$	$0.78_{(0.155)}$
$\hat{\beta}^{FE}$ $(\mathbf{var}(\hat{\beta}^{FE}))$		$-0.12_{(0.025)}$	$0.05_{(0.032)}$	$-0.74_{(0.132)}$	$0.63_{(0.133)}$

$\hat{\beta}^{RE}{}_{(\mathrm{var}(\hat{\beta}^{RE}))}$	$-0.16_{(0.073)}$	$-0.13_{(0.025)}$	$0.05_{(0.038)}$	$-0.78_{(0.135)}$	$0.65_{(0.152)}$
$\hat{\beta}^{LWS}{}_{(\mathrm{var}(\hat{\beta}^{LWS}))}$	$3.10_{(0.003)}$	$1.30_{(0.004)}$	$2.49_{(0.003)}$	$-3.21_{(0.006)}$	$6.70_{(0.003)}$
$\hat{\beta}^{FWE}{}_{(\mathrm{var}(\hat{\beta}^{FWE}))}$		$0.72_{(0.151)}$	$1.90_{(0.154)}$	$-3.77_{(0.122)}$	$6.08_{(0.181)}$
$\hat{\beta}^{RWE}{}_{(\mathrm{var}(\hat{\beta}^{RWE}))}$	$3.46_{(2778.762)}$	$0.73_{(0.134)}$	$1.92_{(0.138)}$	$-3.75_{(0.109)}$	$6.07_{(0.182)}$

5 Conclusions

Prior to a discussion of the results let me mention that one thing can seem to be a bit strange (or even, at the first glance, stricking?). The results for OLS and RE are very similar. But it is O.K. - the theory confirms it. Under the model (6) both estimators are unbiased and consistent, but the latter is not efficient (but the empirical variances exhibit very small improvement, if any).

The results indicate without any doubts that even a mild contamination causes problems for the classical estimators. The proposed robustification of the estimation of model with fixed or random effects appeared to cope with contamination - up to level of 4% or even 6% (please open http://samba.fsv.cuni.cz/~visek/FixedRandomEffects/ where the results for much more levels of contamination are given - also the results for data without contamination) - rather well. Moreover, the loss of efficiency is also relatively mild. For a higher contamination it is preferable to utilize directly $\hat{\beta}^{(LWS,n,T,w)}$ or $\hat{\beta}^{(IWV,n,T,w)}$ (*Instrumental Weighted Variables*, see [20]). Naturally, a question appears how to learn about the level of contamination. Taking into account the fact that the computational times are (due to the algorithm, we have employed - see [20], as well as due to the level of computational means) more than acceptable. So we can start - as we have already mentioned above - with the estimation of model assuming (very) high contamination of data and we may repeat computations with decreasing level of robustness of the estimator while the changes of the estimates of regression coefficients are smooth. If data are (significantly) contaminated, we typically arrive to a break in the development of the estimates of regression coefficients (similarly as it is described in [1]). Then we should stop and we can (usually) consider the results obtained immediately (or a bit earlier) before this break as the correct ones.

Let me to return to two facts mentioned a few lines above - the small decrease of efficiency and the speed of algorithm. Both these facts indicate that the (still overliving) idea that the robust methods suffer by a (significant) loss of efficiency and/or that they require extensive computations need not be right. Of course, we feel intuitively that some *"law of preservation of mass"* should hold, i. e. we feel that the insurance against the contamination is to be paid. And it is really paid - by the intricacy of proofs of the consistency, the asymptotic normality, etc. and similarly by the requirement of invention of new (complicated) algorithms, of their implementation and the verification of their quality. On the other hand, when these objections are once overcome, the user can profit from new possibilities a lot.

Acknowledgement First of all, we would like to thank two anonymous reviewers for a lot of recommendations and proposals which led (we hope) to an improvement of the paper. Moreover, their ideas may lead in future to the substantial enlargement of the study and to a discussion of some other topics then opened here. Finally, this paper was written with the support of the Czech Science Foundation project GA CR P402/12/0982 *Trade Flows in Times of Economic Boom and Slump: Modifying the Gravity Model for Country, Time and Product-Specific Decision-Making.*

Bibliography

[1] Atkinson, A. C., M. Riani (2000): *Robust Diagnostic Regresion Analysis.* New York: Springer Verlag.

[2] Bickel, P. J. (1975): *One-step Huber estimates in the linear model.* JASA 70, 428 - 433.

[3] Bramati, M. C., Croux, C. (2007): *Robust estimators for the fixed effects panel data model.* Econometric Journal 10, 521 - 540.

[4] Croux, C., Rousseeuw, P. J. (1992): *A class of high-breakdown scale estimators based on subranges.* Communications in Statistics - Theory and Methods 21, 1935 - 1951.

[5] Dehon, C., Gassner, M., Veradi, V. (2009): *Beware of 'good' outliers and overoptimistic conclussios.* Oxford Bulletin of Economics and Statistics 71 (3), 437 - 452.

[6] Hampel, F. R., E. M. Ronchetti, P. J. Rousseeuw, W. A. Stahel (1986): *Robust Statistics – The Approach Based on Influence Functions.* New York: J.Wiley *&* Son.

[7] Hettmansperger, T. P., S. J. Sheather (1992): *A Cautionary Note on the Method of Least Median Squares.* The American Statistician 46, 79 - 83.

[8] Judge, G. G., Griffiths, W. E.,Hill,R. C., Lutkepohl, H., Lee, T. C. (1985): *The Theory and Practice of Econometrics.* New York: J.Wiley *&* Sons (second edition).

[9] Jurečková, J., Sen, P. K. (1984): *On adaptive scale-equivariant M-estimators in linear models.* Statistics and Decisions, vol. 2 (1984), Suppl. Issue No. 1.

[10] Kott,P. (1989): *Robust small domain estimation using random effects modelling.* Survey Methodology 15, 1 - 12.

[11] Portnoy, S. (1983): *Tightness of the sequence of empiric c.d.f. processes defined from regression fractiles.* In Robust and Nonlinear Time-Series Analysis (J. Franke, W. Hărdle, D. Martin, eds.), 231 - 246. Springer-Verlag, New York, 1983.

[12] Rocke, D. M. (1991): *Robustness and balance in the mixed models.* Biometrics 47, 303 - 309.

[13] Rousseeuw, P.J. (1984): *Least median of square regression.* JASA 79, 871 - 880.

[14] Veradi,V., Croux, C. (2009): *Robust regression in Stata.* The Stata Journal 9 (3), 439 - 453.

[15] Veradi,V., Wagner, J. (2010): *Robust estimation of linear fixed effects panel data model with an application to the exporter productivity premium.* Discussion paper No 4928, IZA, Bonn.

[16] Víšek, J. Á. (1996): *A cautionary note on the method of Least Median of Squares reconsidered.* Trans. Twelfth Prague Conf. on Inf. Theory, Statist. Decision Functions and Random Processes, Academy of Sciences of the Czech Republic, 254 - 259.

[17] Víšek, J. Á. (1996): *Sensitivity analysis of M-estimates.* Annals of the Institute of Statistical Mathematics 48, 469 - 495.

[18] Víšek, J. Á. (2000): *Regression with high breakdown point.* Proc. ROBUST 2000, eds. J. Antoch & G. Dohnal, The Union of the Czech Mathematicians, 324 - 356.

[19] Víšek, J. Á. (2002): *Sensitivity analysis of M-estimates of nonlinear regression model: Influence of data subsets.* Annals of the Institute of Statistical Mathematics 54, 261 - 290.

[20] Víšek, J. Á. (2006): *Instrumental Weighted Variables - algorithm.* Proc. COMPSTAT 2006, eds. A. Rizzi & M. Vichi, Physica-Verlag (Springer Company) Heidelberg 2006, 777 - 786.

[21] Víšek, J. Á. (2006): *The least trimmed squares. Sensitivity study.* Proc. Prague Stochastics 2006, eds. M. Hušková & M. Janžura, matfyzpress, 728 - 738.

[22] Víšek, J. Á. (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), 254 - 267.

[23] Víšek, J. Á. (2011): *Consistency of the least weighted squares under heteroscedasticity.* Kybernetika 47 , 179 - 206.

[24] Víšek, J. Á. (2011): *Empirical distribution function under heteroscedasticity.* Statistics 45, 497 - 508.

[25] Wooldridge, J. M. (2006): *Introductory Econometrics. A Modern Approach.* MIT Press, Cambridge, Massachusetts, second edition 2009.