Bulletin of the Econometric Society vol. 17, no 27 (2010)

Heteroscedasticity Resistant Robust Covariance Matrix Estimator

Jan Ámos Víšek[1]

Department of Macroeconomics and Econometrics, Faculty of Social Sciences, Charles University
&
Department of Econometrics, Institute of Information Theory and Automation, Academy of Sciences of the Czech Republic

[1]Research was supported by grant of GA ČR number 402/09/0557.

Abstract

It is straightforward that breaking the *orthogonality condition* implies biased and inconsistent estimates by means of the *ordinary least squares*. If moreover, the data are contaminated it may significantly worsen the data processing, even if it is performed by *instrumental variables* or the *(scaled) total least squares*. That is why the method of *instrumental weighted variables* based of weighting down order statistics of squared residuals (rather than directly squared residuals) was proposed. The main underlying idea of this method is recalled and discussed. Then it is also recalled that *neglecting heteroscedasticity* may end up in *significantly wrong specification* and *identification* of regression model, just due to wrong evaluation of *significance of the explanatory variables*. So, if the test of heteroscedasticity (which is in the case when we use the instrumental weighted variables just robustified version of the classical White test for heteroscedasticity) rejects the hypothesis of homoscedasticity, we need an *estimator of covariance matrix (of the estimators of regression coefficients) resistant to heteroscedasticity*. The proposal of such an estimator is the main result of the paper. At the end of paper the *numerical study of the proposed estimator* (together with results offering comparison of model estimation by means of the ordinary least squares, the least weighted squares and by the instrumental weighted variables) is included.

Key words: Robustification of classical instrumental variables; estimating the covariance matrix of the estimators of regression coefficients under heteroscedasticity; the weighting of order statistics of squared residuals.

JEL classification: C13, C19

AMS classification: 62J05, 62G35.

INTRODUCTION

The linear regression analysis, classically performed by means of *the ordinary least squares (OLS)*, proved to be one of the most powerful tool for data processing, see e. g. [6]. On the other hand, couple of conditions are to be (nearly strictly) fulfilled for the *OLS* because even a slight departure from them may cause large deviations of the estimates from true values of parameters, see [13], [30] or [1]. One of important conditions is *normality of disturbances* because otherwise *OLS* are the best method *only* among the linear estimators, for particularly nice discussion see [11] (a bit forgotten results concerning normality - not in regression framework - were given in [8]). Important conditions are also *orthogonality of explanatory variables and disturbances* and/or *homoscedasticity of disturbances*, see [17] or [43].

It is easy to show that breaking the orthogonality condition implies that the *OLS*-estimates are biased (see e. g. [14]) and hence method of *the instrumental variables* or *the (scaled) total least squares* are to be employed (for the former see [2], [4], [25] or [26], for the latter then see [19] or [27])[2]. It was also demonstrated in literature that neglecting heteroscedasticity may lead to underestimation of variance of estimator of some regression coefficients and so to the wrong evaluation of significance of some explanatory variables. Consequently it implies a false specification of model and finally it brings wrong estimate of regression coefficients, even for those explanatory variables which should be included into model, see e. g. [17] or [43] and see also an example given below. Finally, it is nowadays commonly accepted that the influential points can seriously damage the results of data processing. Hence along with the classical methods some robust one(s) should be used and results compared (see [11], [23] or [24], among many others). The paper focuses on solution of these problems in the case when they appear simultaneously.
So the paper offers a possibility how to cope (simultaneously) with the situation when:

- disturbances are heteroscedastic,
- orthogonality condition is broken,

and

- data are contaminated (by some outliers and/or leverage points).

1 Notations and framework

Let $\mathcal{N}$ denote the set of all positive integers, R the real line and R^p the p-dimensional Euclidean space. In what follows it is assumed that all random variables (r. v.'s) are defined on a probability space $(\Omega, \mathcal{A}, P)$. The linear regression model given as

$$Y_i = X_i'\beta^0 + e_i = \sum_{j=1}^{p} X_{ij}\beta_j^0 + e_i, \quad i = 1, 2, ..., n$$

will be considered (all vectors will be assumed to be the column ones) with random explanatory variables. Throughout the paper we shall assume:

[2]Due to the construction of the methods *the instrumental variables* are (mainly) used in social sciences while *the total least squares* in the exact ones.

$\mathcal{C}1$ The sequence $\{(V_i', e_i)'\}_{i=1}^{\infty}$ is sequence of independent p-dimensional random variables distributed according to the distribution functions (d.f.) $F_{V,e_i}(v,r) = F_{V,e}(v, \sigma_i \cdot r)$, $i \in N$ where $F_{V,e}(v,r)$ is a parent d.f., $\mathbb{E}V_1 = 0$, the covariance matrix $\mathbb{E}\{V_1 V_1'\}$ is regular and $\sigma_i^2 = \text{var}(e_i|V_i)$ with $0 < \sigma_i^2 < K < \infty$. There is ℓ, $0 \le \ell < p$ and coordinates $V_{11}, V_{12}, ..., V_{1\ell}$ of the vector V_1 are discrete with the distribution given by $\{p_{1,v} = P(V_{11} = v_1, V_{12} = v_2, ..., V_{1\ell} = v_\ell)\}_{\{v \in \mathcal{U}\}}$ where $\mathcal{U} \subset \mathcal{T}$ and $\mathcal{T} \subset R^\ell$ is a compact. The d.f. of the vector $(V_{1,\ell+1}, V_{1,\ell+2}, ..., V_{1,p-1}, e_1)'$ is absolutely continuous, the density $f_{V_{1,\ell+1},V_{1,\ell+2},...,V_{1,p-1},e_1}(v,e)$ is bounded, say by B, and the marginal density $f_{V_{1,\ell+1},V_{1,\ell+2},...,V_{1,p-1}}(v)$ have a bounded support, i.e. putting $M = \sup\left\{\|v\| \,:\, f_{V_{1,\ell+1},V_{1,\ell+2},...,V_{1,p-1}}(v) > 0\right\}$, we have $M < \infty$. Finally, consider the sequence $\{(X_i', e_i)'\}_{i=1}^{\infty}$ where $X_{i1} = 1$ and $X_{ij} = V_{i,j-1}$, $j = 2, 3, ..., p$ for all $i \in N$. This sequence will be considered as the sequence of explanatory variables and of disturbances.

Notice please that the marginal d.f.'s $F_V(v)$ of vectors V_i's are the same for all $i \in N$. Also notice that we assume that the disturbances e_i's can be correlated with explanatory variables V_i's. Moreover, disturbances are assumed generally heteroscedastic. Finally, as $f_{V,e_i}(v,r) = \sigma_i \cdot f_{v,e}(v, \sigma_i \cdot r)$, we have $f_{V,e_i}(v,r) < \sup_{i \in N} \sigma_i \cdot B$. Notice also that the assumption $\mathbb{E}V_1 = 0$ (which of course implies that $\mathbb{E}X_1 = (1, 0, ..., 0)'$) does not restrict the generality because otherwise we would consider $\tilde{X}_{j1} = X_{1j} - \mathbb{E}X_{1j}$ and $\tilde{\beta}_j^0 = \beta_j^0 + \mathbb{E}X_{1j} \cdot \beta_j^0$ for $j = 2, 3, ..., p$.

The form of condition $\mathcal{C}1$ is a bit complicated because we need an upper bound on the density of those explanatory variables which are absolutely continuous and on the other side we would like to keep framework enough general to allow for discrete explanatory variables. In fact, the structure of explanatory variables X's is simple. The first coordinate represents in fact the intercept, the second one up to $\ell + 1$ are discrete r.v.'s and the rest of coordinates are absolutely continuous r. v.'s.

Further, for any $\beta \in R^p$ $\quad r_i(\beta) = Y_i - X_i'\beta$ denotes the i-th residual[3] and $r_{(h)}^2(\beta)$ stays for the h-th order statistic among the squared residuals, i.e. we have

$$r_{(1)}^2(\beta) \le r_{(2)}^2(\beta) \le \;\ldots\; \le r_{(n)}^2(\beta).$$

[3]Sometimes residuals are considered only with respect to an estimator $\hat{\beta}$ of β^0 as $r_i = Y_i - X_i'\hat{\beta}$. We shall consider them generally with respect to any $\beta \in R^p$ as $r_i(\beta) = Y_i - X_i'\beta$. The reasons for it will be evident e.g. from Definition 1.

2 The Least Weighted Squares

To be able to discus in the fourth section the robustified version of *the instrumental variables* (as given in [33] and called there *the instrumental weighted variables*) we need to remind (see [30] and [31]):

Definition 1 *Let for any* $n \in \mathcal{N}$ $\quad w_i \in [0, 1],\ i = 1, 2, ..., n$. *Then*

$$\hat{\beta}^{(LWS,n,w)} = \underset{\beta \in R^p}{arg\,min} \sum_{i=1}^{n} w_i r^2_{(i)}(\beta) \tag{1}$$

is called the least weighted squares (LWS) estimator.

Remark 1 *Let us fix an integer* $h \in [n/2, n]$. *Then for* $w_h = 1$ *and* $w_i = 0$ *for* $1 \leq i \leq n, i \neq h,$ $\hat{\beta}^{(LWS,n,w)}$ *turns to be the least median of squares*

$$\hat{\beta}^{(LMS,n,h)} = \underset{\beta \in R^p}{arg\,min}\ r^2_{(h)}(\beta)$$

and for $w_i = 1, i = 1, 2, ..., h$ *and* $w = 0$ *otherwise,* $\hat{\beta}^{(LWS,n,w)}$ *is the least trimmed squares (for both see [11] or [24])*

$$\hat{\beta}^{(LTS,n,h)} = \underset{\beta \in R^p}{arg\,min} \sum_{i=1}^{h} r^2_{(i)}(\beta).$$

Remark 2 *The idea of downweighting observations or residuals was employed from the very beginning of robust statistics. After all, generalized* M*-estimators do it also as well as recalled the least median of squares and the least trimmed squares. Nevertheless, the least median of squares and the least trimmed squares have advantage in comparison with (generalized)* M*-estimat- ors, namely they are "automatically" scale-adjusted. In other words, both of them - as well as the least weighted squares - are scale- and regression-equivariant. To achieve scale- and regression-equivariance for* M*-estimators we have to studentize the residuals by scale-equivariant and regression-invariant estimator of standard deviation, see [3] or [15]. The only example of such an estimator was given in [15] (to my knowledge), being based on regression quantiles (or regression ranks, see [15]). However, it is not very easy to compute it. Hence it is preferable to have estimators which do not need scale adjustment.*

We usually generate the weights w_i's by a non-increasing (absolutely continuous) weight function $w(v) : [0,1] \to [0,1]$, putting $w_i = w\left(\frac{i-1}{n}\right)$. Then (1) is changed to

$$\hat{\beta}^{(LWS,n,w)} = \underset{\beta \in R^p}{arg\,min} \sum_{i=1}^{n} w\left(\frac{i-1}{n}\right) r^2_{(i)}(\beta). \qquad (2)$$

Of course, monotonicity of w restricts a bit generality as e. g. LMS cannot be then assumed as special case of LWS. Similarly, continuity hamper to see LTS as special case of LWS. The reason for restricting ourselves on continuous function w is the fact that for the discontinuous function the subsample stability of the estimator may be low. In fact, when w has a point of discontinuity, as in the case of the least median of squares and the least trimmed squares, the values of estimates may change dramatically when excluding one observations (at the first glance surprisingly not at the outskirts of a "cloud" of data but usually at the center of it), see [28], [29] or [34].
So we may conclude that $\hat{\beta}^{(LWS,n,w)}$ has the same advantage as $\hat{\beta}^{(LMS,n,h)}$ and $\hat{\beta}^{(LTS,n,h)}$ (being scale- and regression-equivariant) and simultaneously (if the weight function w is continuous) having acceptable subsample stability.

In what follows we will need (2) in a bit modified form. So following Hájek and Šidák [10] for any $i \in \{1, 2, ..., n\}$ let us define the rank of i-th residual by

$$\pi(\beta, i) = j \in \{1, 2, ..., n\} \qquad \Leftrightarrow \qquad r^2_i(\beta) = r^2_{(j)}(\beta)$$

(notice that $\pi(\beta, i)$ is r. v. since it depends on $X_i(\omega)$'s and $e_i(\omega)$'s). Then we have

$$\hat{\beta}^{(LWS,n,w)} = \underset{\beta \in R^p}{\arg\min} \sum_{i=1}^{n} w\left(\frac{\pi(\beta, i) - 1}{n}\right) r^2_i(\beta). \qquad (3)$$

Now, realize please, that having fixed $\beta \in R^p$ and denoting $a_i(\beta) = |r_i(\beta)|$, the order statistics $a_{(i)}(\beta)$'s and the order statistics of the squared residuals $r^2_{(i)}(\beta)$'s assign to given fix observation i_0 (say) the same rank, i. e. if the squared residual of the observation i_0 is on the ℓ-th position (say) in the sequence

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

then the absolute value of residual of the i_0-th observation is in the sequence

$$a_{(1)}(\beta) \leq a_{(2)}(\beta) \leq ...a_{(n)}(\beta)$$

also on the ℓ-th position. Further, let us denote for any $\beta \in R^p$ and any $v \in R$ the empirical d. f. of the absolute value of residuals $|r_i(\beta)| = |Y_i - X_i'\beta|$ by $F_n^{(\beta)}(v)$, i. e.

$$F_\beta^{(n)}(v) = \frac{1}{n}\sum_{i=1}^{n} I\left\{|r_i(\beta)| < v\right\} = \frac{1}{n}\sum_{i=1}^{n} I\left\{|Y_i - X_i'\beta| < v\right\} \tag{4}$$

where $I\{A\}$ denotes the indicator of the set A. Now, let's realize that the empirical d. f. $F_\beta^{(n)}(v)$ has the first jump just "after" $a_{(1)}(\beta)$ (as due to the sharp inequality in (4), we have $F_\beta^{(n)}(a_{(1)}(\beta)) = 0$ and $\lim_{v \to [a_{(1)}(\beta)]+} = \frac{1}{n}$), the second jump at $a_{(2)}(\beta)$ (but $F_\beta^{(n)}(a_{(2)}(\beta)) = \frac{1}{n}$), the third at $a_{(3)}(\beta)$ (but $F_\beta^{(n)}(a_{(3)}(\beta)) = \frac{2}{n}$), etc. Hence it has the $\pi(\beta, i)$-th jump at $a_{(\pi(\beta,i))}(\beta)$, i. e.

$$F_\beta^{(n)}(a_{(\pi(\beta,i))}(\beta)) = F_\beta^{(n)}(|r_i(\beta)|) = \frac{\pi(\beta, i) - 1}{n}.$$

So, we can rewrite (3) as

$$\hat{\beta}^{(LWS,n,w)} = \arg\min_{\beta \in R^p} \sum_{i=1}^{n} w\left(F_\beta^{(n)}(|r_i(\beta)|)\right) r_i^2(\beta). \tag{5}$$

The equation (5) represents the extremal problem which defines the classical *weighted least squares* with "randomized" weights, namely with the weights $w\left(F_\beta^{(n)}(|r_i(\beta)|)\right)$. The solution of such problem is given as

$$\hat{\beta}^{(LWS,n,w)} = \left(X'W(F_\beta^{(n)})X\right)^{-1} X'W(F_\beta^{(n)})Y \tag{6}$$

where $X = (X_1, X_2, ..., X_n)'$ is the random design matrix (where the assumptions about the random vectors X_i's were given in conditions $\mathcal{C}1$), $Y = (Y_1, Y_2, ..., Y_n)'$ is the vector of response variables and $W(F_\beta^{(n)})$ is the diagonal matrix with the i-th diagonal element equal to $w\left(F_\beta^{(n)}(|r_i(\beta)|)\right)$. It is straightforward that (6) is the solution of *the normal equations*

$$\sum_{i=1}^{n} w\left(F_\beta^{(n)}(|r_i(\beta)|)\right) X_i\left(Y_i - X_i'\beta\right) = 0. \tag{7}$$

Finally, (7) is just the equation which will allow us to define in the next section the robustified version of *the instrumental variables.*

3 Potential problems with neglecting heteroscedasticity

The consequences of not taking into account of heteroscedasticity of error terms are frequently underestimated due to the fact that it is usually stressed that under heteroscedasticity the *OLS*-estimates are still unbiased, consistent and asymptotically normal, so that the only deterioration of estimates is a decrease of efficiency. In fact however, neglecting heteroscedasticity may lead to the false evaluation of significance of explanatory variables due to the fact that it is established by means of an estimate of nonexisting common variance of all error terms. Consequently it produce the wrong specification of model. Finally, it may end up in significantly biased estimates of coefficients of regression model. Let's give an example.

We looked for a model of export from the Czech republic into EU for the period from 1993 to 2001. The Czech economy was divided into 61 industries and export (X), import (M), export prices (PI), import prices (ME), value added (VA), capital (K), labor (L), debts (DE), foreigner direct investment (FDI) and gross domestic production per capita in EU (GDPeu) were employed as response and explanatory variables, respectively. We arrived at the model ($t = 1994, 1995, ..., 2001, \quad i = 1, 2, ..., 61$)

$$ln(X_{i,t}) = \beta_1 + \beta_2 \cdot ln(X_{i,t-1}) + ... + \beta_{14} \cdot ln(GDPeu_{i,t-1}) \tag{8}$$

with the estimate of coefficients given as (explanatory variable with the subindex t-1 stays for the lagged value of it)[4]:

Explanat. Variable	Estim. Coeffs	Stand. Errors	p-values
intercept	9.643	5.921	0.104
$ln(X_{t-1})$	0.827	0.033	0.000
$ln(PE_t)$	-0.164	0.06	0.007
$ln(PE_{t-1})$	0.2	0.062	0.001
$ln(VA_t)$	0.337	0.077	0.000
$ln(VA_{t-1})$	-0.228	0.079	0.004
$ln(K_t/L_t)$	-0.625	0.159	0.000

[4] As Durbin-Watson statistics indicated autocorrelation of error terms, we employed proposal of James Durbin to include into the model also lagged values of both - response as well as explanatory variables, see e. g. [7] or [14]. This approach is much safer then application of Cochrane-Orcutt (see [5]) or Prais-Winston (see [22]) transformation. For a nice discussion of it see [18]. Grayham Mizon there in fact says: *"If we consider wrong model, we have nearly no chance to estimate correctly the parameters of the true one."*

Explanat. Variable	Estim. Coeffs	Stand. Errors	p-values
$ln(K_{t-1}/L_{t-1})$	0.518	0.157	0.001
$ln(DE_t/VA_t)$	0.296	0.122	0.016
$ln(DE_{t-1}/VA_{t-1})$	-0.292	0.119	0.015
$ln(FDI_t)$	0.147	0.056	0.009
$ln(FDI_{t-1})$	-0.151	0.056	0.007
$ln(GDPeu_t)$	1.126	0.629	0.045
$ln(GDPeu_{t-1})$	-1.966	0.623	0.002

The model seemed to be rather complicated. Nevertheless, it was well determined with approximately normally distributed residuals and, as the p-values indicate (see table above), all explanatory variables (except of intercept) were significant. Nevertheless, White's test of homoscedasticity gave value 244.1 with corresponding p-value equal to 0.0000. Clearly the error terms in model were heteroscedastic. So, it was necessary for appropriate judgement about the significance of explanatory variables to employ *White's estimator of covariance matrix of the estimates of regression coefficients* which is given as

$$\widehat{\Sigma} = \frac{1}{8 \cdot 61} \sum_{t=1994}^{2001} \sum_{i=1}^{61} r_{i,t}^2(\hat{\beta}^{(OLS,n)}) \cdot X_{i,t} \cdot X_{i,t}'$$

where $r_{i,t}(\beta) = Y_{i,t} - X_{i,t}'\beta$, for details see White (1980). We obtained then

Explanat. Variable	Estim. Coeffs	Stand. Errors	p-values
intercept	9.643	4.128	0.200
$ln(X_{t-1})$	0.827	0.046	0.000
$ln(PE_t)$	-0.164	0.107	0.127
$ln(PE_{t-1})$	0.2	0.107	0.062
$ln(VA_t)$	0.337	0.203	0.098
$ln(VA_{t-1})$	-0.228	0.192	0.235
$ln(K_t/L_t)$	-0.625	0.257	0.016
$ln(K_{t-1}/L_{t-1})$	0.518	0.301	0.087
$ln(DE_t/VA_t)$	0.296	0.292	0.312
$ln(DE_{t-1}/VA_{t-1})$	-0.292	0.282	0.302
$ln(FDI_t)$	0.147	0.141	0.300
$ln(FDI_{t-1})$	-0.151	0.123	0.222
$ln(GDPeu_t)$	1.126	1.097	0.305
$ln(GDPeu_{t-1})$	-1.966	0.995	0.049

Excluding successively insignificant explanatory variables (insignificant on the level $p > 0.05$), we finally arrived at the model

Explanatory Variable	Estimated Coefficients	Standard Errors	t-statistics	p-value
intercept	9.643	4.128	2.336	0.200
$ln(X_{t-1})$	0.804	0.05	16.125	0.000
$ln(VA_t)$	0.149	0.039	3.784	0.000
$ln(K_t/L_t)$	-0.214	0.063	-3.38	0.001
$ln(GDPeu_t)$	1.896	0.782	2.425	0.016
$ln(GDPeu_{t-1})$	-2.538	0.778	-3.261	0.001

which is much simpler than (8)[5]. Of course, neglecting heteroscedasticity leads to the estimates of regression coefficients which are still unbiased, consistent and asymptotically normal but which can have rather large variances and hence they are less reliable. Moreover, due to the wrong p-values, the model may include some insignificant explanatory variables which can "contest" with those really significant. The consequence may be that the "explanation" of the response variable is "divided" among more explanatory variables than it would be appropriate, so that identification of regression model may be incorrect. In other words, when applying White's estimate of covariance matrix of the estimates of regression coefficients and correcting significance judgement, excluding successively the "most insignificant" explanatory variable, we may meet with situation when the estimates of other coefficients may dramatically change. So, the conclusion is that *the heteroscedasticity of error terms should be taken into account seriously.*

4 Instrumental weighted variables

In this section we discus the possibilities of identifying regression model when the orthogonality condition is broken and recall the *instrumental weighted variables* (IWV) - a robustified version of the classical instrumental variables. We also recall conditions for consistency of IWV as a basic results in the study of this robust method of estimating regression model.
The violation of orthogonality condition $\mathbb{E}\{e_i|X_i\} = 0$ implies that

$$\lim_{n\to\infty} \frac{1}{n} \sum_{i=1}^{n} X_i e_i \neq 0 \quad \text{in probability} \tag{9}$$

[5]The model passed all commonly applied tests, e. g. for normality of residuals, homoscedasticity, specification of model, etc.

and hence also inconsistency of

$$\hat{\beta}^{(OLS,n)} = \beta^0 + \left(\frac{1}{n}\sum_{i=1}^{n} X_i X_i'\right)^{-1} \frac{1}{n}\sum_{i=1}^{n} X_i e_i. \tag{10}$$

The most frequently given examples of failure of the orthogonality condition are the measurement of explanatory variables with a random error (error-in-variable model) or the (dynamic) regression model with lagged response in the role of explanatory variable, see e.g. [9] or [14].
One possibility how to cope with this situation, usually employed in natural sciences is the method of *the (scaled) total least squares*, see [27]. This method can be applied even in the case when we assume that only explanatory variables were measured with random error while response variable without it (for details as well as discussion see [20]). Econometricians offer as a remedy the method of the *instrumental variables* which defines the estimator as (any) solution of the normal equations

$$\sum_{i=1}^{n} Z_i(Y_i - X_i'\beta) = 0 \tag{11}$$

where the sequence $\{Z_i\}_{i=1}^{\infty}$ is a sequence of instruments for explanatory variables X_i's given as follows: Let $\{U_i\}_{i=1}^{\infty}$ be a sequence of (p - 1)-dimensional r. v.'s such that

$$I\!E\,(e_1|U_1) = 0, \tag{12}$$

so that putting

$$Z_{i1} = 1 \quad \text{and} \quad Z_{ij} = U_{i,j-1} \;\; j = 2, 3, ..., p \tag{13}$$

for all $i \in N$ the orthogonality condition $I\!E\,(e_1|Z_1) = 0$ holds. The analogy of (10)

$$\hat{\beta}^{(IV,n)} = \beta^0 + \left(\frac{1}{n}\sum_{i=1}^{n} Z_i X_i'\right)^{-1} \frac{1}{n}\sum_{i=1}^{n} Z_i e_i. \tag{14}$$

hints that the estimator evaluated by means of method of the *instrumental variables* is (under some conditions) consistent provided (e. g.)

$$I\!E Z_1 X_1' = Q \;\text{ is regular.} \tag{15}$$

Then, following (7), we can define a robustified version of *instrumental variables* (see [33], [35] or [36])

Definition 2 *Any solution of*

$$\sum_{i=1}^{n} w\left(F_{\beta}^{(n)}(|r_i(\beta)|)\right) Z_i \left(Y_i - X_i'\beta\right) = 0. \tag{16}$$

will be called instrumental weighted variables, (IWV) and denoted by $\hat{\beta}^{(IWV,n,w)}$.

In the case of classical *instrumental variables* (13) and (14) indicated that we don't need any "qualitative relation" between explanatory variables and instruments[6]. However for robust version of the method we need some assumption about the mutual behaviour of X_i's and Z_i's even for consistency. To be able to formulate it let's enlarge notations.
Let's recall that we assume heteroscedasticity of the disturbances (see $\mathcal{C}1$) and define a "mean" d.f.

$$\overline{F}_{n,\beta}(v) = \frac{1}{n}\sum_{i=1}^{n} P\left(|Y_i - X_i'\beta| < v\right). \tag{17}$$

The possibility to approximate the empirical distribution $F_{\beta}^{(n)}(v)$ - see (4) - by $\overline{F}_{n,\beta}(v)$ uniformly in $v \in R$ as well as in $\beta \in R^p$ opened in fact the way for results given below, see [41]. Further define

$$F_{\beta' Z X' \beta}(u) = P(\beta' Z_1 X_1' \beta < u)$$

and put for any $\lambda \in R^+$ and any $a \in R$

$$\gamma_{\lambda,a} = \sup_{\|\beta\|=\lambda} F_{\beta' Z X' \beta}(a). \tag{18}$$

Finally, for any $\lambda \in R^+$ let us denote

$$\tau_\lambda = -\inf_{\|\beta\|\le\lambda} \beta' \mathbb{E}\left[Z_1 X_1' \cdot I\{\beta' Z_1 X_1' \beta < 0\}\right] \beta. \tag{19}$$

$\mathcal{C}3$ *The* (p - 1)*-dimensional r. v.'s* $\{U_i\}_{i=1}^{\infty}$ *are independent and identically distributed with distribution function* $F_U(u)$*, the covariance matrix* $\mathbb{E}\,\{U_1 U_1'\}$ *is regular and positive definite,* $\mathbb{E}\,(e_i|U_i) = 0$ *and* $\sigma_i^2 = \operatorname{var}(e_i|U_i)$ *with* $0 <$

[6] (12) is sufficient for unbiasedness and consistency. It is however easy to see that the variances of $\hat{\beta}_j^{(IV,n)}$'s depend on the mutual relation between X_i's and Z_i's. In other words, if there are not "natural" instrument, e.g. lagged values, the method can work poorly.

$\sigma_i^2 < K < \infty$. There is ℓ, $0 \le \ell < p$ and coordinates $U_{11}, U_{12}, ..., U_{1\ell}$ of the vector U_1 are discrete with the distribution given by $\{p_{1,v} = P(U_{11} = v_1, U_{12} = v_2, ..., U_{1\ell} = v_\ell)\}_{\{v \in \mathcal{S}\}}$ where $\mathcal{S} \subset \mathcal{G}$ and $\mathcal{G} \subset R^\ell$ is a compact. The d.f. of the vector $(U_{1,\ell+1}, U_{1,\ell+2}, ..., U_{1p})'$ is absolutely continuous, the density $f_{U_{1,\ell+1},V_{1,\ell+2},...,V_{1p}}(v)$ is bounded, say by C, and have a bounded support, i.e. putting $M^* = \sup\left\{\|v\| \, : \, f_{U_{1,\ell+1},U_{1,\ell+2},...,U_{1p}}(v) > 0\right\}$, we have $M^* < \infty$. Further, consider the sequence $\{(Z_i', e_i)'\}_{i=1}^\infty$ where $Z_{i1} = 1$ and $Z_{ij} = U_{i,j-1}$, $j = 2, 3, ..., p$ for all $i \in N$. Moreover, there is $q > 1$ so that $I\!\!E\{\|Z_1\| \cdot \|X_1\|\}^q < \infty$ and $n_0 \in N$ so that for all $n > n_0$

$$I\!\!E\left\{\frac{1}{n}\sum_{i=1}^{n}\left[w(\overline{F}_{n,\beta}(|e_i|))Z_iX_i'\right]\right\}$$

is regular. Finally, there is $a > 0$, $b \in (0,1)$ and $\lambda > 0$ so that

$$a \cdot (b - \gamma_{\lambda,a}) \cdot w(b) > \tau_\lambda. \tag{20}$$

For discussion of **C3** see [35] or [37].

$\mathcal{C}4$ There is $n_0 \in N$ so that for all $n > n_0$ the vector equation

$$\beta' I\!\!E\left\{\frac{1}{n}\sum_{i=1}^{n}\left[w(\overline{F}_{n,\beta}(|r_i(\beta)|))Z_i\left(e_i - X_i'\beta\right)\right]\right\} = 0 \tag{21}$$

in the variable $\beta \in R^p$ has unique solution $\beta^0 = 0$.

Lemma 1 *Let Conditions $\mathcal{C}1$, $\mathcal{C}2$, $\mathcal{C}3$ and $\mathcal{C}4$ be fulfilled. Then any sequence $\left\{\hat{\beta}^{(IWV,n,w)}\right\}_{n=1}^\infty$ of the solutions of normal equations (16) is weakly consistent.*

Proof is given in [37] where also a simulation study demonstrates that the algorithm, firstly presented in [36], works very well.

5 Estimating covariance matrix of the IWV-estimators under heteroscedasticity

When processing data we can arrive at a suspicion that data are contaminated and then we may use except of OLS also LWS. If the results are similar, our suspicion is probably wrong. In the opposite case we may still have a suspicion that the orthogonality condition was broken. Then we may use also IWV and a robustified version of Hausman test (see [21]). Finally, we should test for homoscedasticity/heteroscedasticity by the robustified version of White test (see [40]). Of course, a couple of other tests as Durbin-Watson (in the case of *panel data*) or test for normality of residuals are to be applied. A few of them are already available for some robust methods, see [16] or [32], others are under progress.

If the homoscedasticity is rejected, but we ignore it, the conclusions about significance of the explanatory variables may be misleading - because they are based on the wrong assumption that the standard deviations of the disturbances are the same for all of them. Hence we need an estimator of variance matrix of the estimates of regression coefficients which is resistant against heteroscedasticity. Such estimator of covariance matrix for the case when we process data by means of the *ordinary least squares* was proposed by Halbert White [42]. To be able to present its robustified version (designed for the IWV) we need some conditions (which allowed to prove $\sqrt{n}$-consistency and derive the asymptotic representation of the respective estimators of regression coefficients).

$\mathcal{NC}1$ *Denote by* $f_{e|V}(r|V_1 = x)$ *the conditional density corresponding to the d.f.* $F_{V,e}(v,r)$*). The density* $f_{e|V}(r|V_1 = x)$ *is uniformly with respect to* x *Lipschitz of the first order (with the corresponding constant equal to* B_e*). Moreover,* $f'_e(r)$ *exists and is bounded in absolute value by* U'_e*.*

$\mathcal{NC}2$ *The derivative* $w'(\alpha)$ *of the weight function is Lipschitz of the first order (with the corresponding constant* J_w*).*

Denote by $g(r)$ the density of the d.f. $G(r) = P(e_1^2 < r)$ (notice that under $\mathcal{C}1$ density $g(r)$ always exists). Moreover, for any $\alpha \in (0,1)$ denote by u_α^2 the upper α-quantile of d.f. G, i.e. we have $P(e_1^2 > u_\alpha^2) = \alpha$.

$\mathcal{AC}1$ *For any* $\alpha \in (0,1)$ *there is* $\delta(\alpha) > 0$ *so that*

$$\inf_{r\in(0,u_\alpha^2+\delta(\alpha))} g(r) > L_{g,\alpha} > 0 \quad \text{and} \quad \inf_{|r|\in(0,\sqrt{u_\alpha^2+\delta(\alpha)})} f(r) > L_{f,\alpha} > 0. \tag{22}$$

Similarly as above (see text under $\mathcal{C}1$*) the condition* $\mathcal{AC}1$ *implies in fact that (22) holds for all densities* $g_{e_i}(r)$ *and* $f_{e_i}(r)$*, i.e. for all* $i \in N$*.*

$\mathcal{AC}2$ *There is* $q > 1$ *so that* $\sup_{i\in N} I\!\!E\,|e_i|^{2q} < \infty$*.*

Following Halbert White ([42]), we may prove:

Lemma 2 *Let the conditions $\mathcal{C}1$, $\mathcal{C}2$, $\mathcal{C}3$, $\mathcal{C}4$, $\mathcal{NC}1$, $\mathcal{NC}2$, $\mathcal{AC}1$ and $\mathcal{AC}2$ hold. Then*

$$\left[\frac{1}{n}\sum_{i=1}^{n} Z_i X_i'\right]^{-1} \left[\sum_{i=1}^{n} r_i^2(\hat{\beta}^{(IWV,n,w)}) Z_i X_i'\right] \left[\frac{1}{n}\sum_{i=1}^{n} Z_i X_i'\right]^{-1} \tag{23}$$

is weakly consistent estimator of covariance matrix of $\hat{\beta}^{(IWV,n,w)}$.

The proof follows in a straightforward way from the lemma 3.

6 Simulation study of robustified White's estimator

In the simulations we employed 3 different frameworks as follows:
In every framework we fixed T (number of observations) and p (dimension of model) and considered regression model

$$Y_t = \beta_1 \cdot X_{t1} + \beta_2 \cdot X_{t2} + ... + \beta_p \cdot X_{tp} + e_t, \quad t = 1, 2, ..., T \tag{24}$$

(of course, T, p and values of regression coefficients β_j's are specified at the beginning of reports of the results for each framework).
Further, for every framework we generated one fix sequence $\{\sigma_t\}_{t=1}^{T}$ of i.i.d. r.v.'s uniformly distributed on the interval $[0.5, 1.5]$. Finally, we selected $\alpha \in (0, 1)$ (α is also specified at the start of every framework). Then we repeated 100 experiments, each of them containing the following 4 steps.
Step 1 In each experiment we started with generating two, mutually independent i.i.d. sequences $\{v_t\}_{t=1}^{T}$ and $\{u_t\}_{t=1}^{T}$ of p-dimensional standard normal r.v.'s. which were independent from the sequence $\{\sigma_t\}_{t=1}^{T}$.
Step 2 Then we put

$$X_t = \alpha \cdot v_t + (1 - \alpha) \cdot u_t, \quad Z_t = u_t \ \text{ and } \ e_t = \sigma_t \sum_{j=1}^{p} v_{tj}. \tag{25}$$

It follows from (25) that $\{X_t\}_{t=1}^{\infty}$ is a sequences of i.i.d. r.v.'s, similarly $\{Z_t\}_{t=1}^{\infty}$ while $\{e_t\}_{t=1}^{\infty}$ is a sequences of independent r.v.'s but e_t's are heteroscedastic. (25) also implies that X_t's and e_t's are mutually correlated while Z_t's and e_t's are mutually independent. In what follows we have employed X_t's as explanatory variables, Z_t's as instrumental variables and finally, e_t's as disturbances.

Step 3 We evaluated response variables employing model (24) (without intercept, as estimating intercept - which is always independent from disturbances - is not interesting). Although we know d.f. of X_t's,Z_t's and of e_t's, due to the fact that we can find only approximative numerical solution of the extremal problem (16), we are not able to evaluate var$(\hat{\hat{\beta}}_j^{(IWV,n,w)})$, $j = 1, 2, .., p$. That is why we simulated values these variances, i.e. we repeat 100 times the experiments and evaluated empirical value of them, see (26) below.

Step 4 Finally we contaminated data. Five last explanatory vectors of explanatory variables as well as instrumental variables were shifted about ± 8 with randomly selected sign (to create leverage points). First two response variables were increased about 5 (to generate outliers).

As already mentioned, each experiment was 100 times repeated (in given fixed framework) and estimates of regression coefficients by means of OLS, LWS and IWV were evaluated (the OLS and LWS estimates were evaluated to demonstrate that they are not able to cope with the fact that the orthogonal condition is broken; of course, OLS estimates suffer also due to the contamination of data). We denoted them for the k-th experiment by

$$\hat{\beta}_{(k)}^{(OLS,T)}, \quad \hat{\beta}_{(k)}^{(LWS,T,w)} \quad \text{and} \quad \hat{\beta}_{(k)}^{(IWV,T,w)},$$

respectively. Then we calculated mean values (over these 100 experiments) of the estimates

$$\hat{\beta}_{(mean)}^{(OLS,T)} = \frac{1}{100}\sum_{k=1}^{100} \hat{\beta}_{(k)}^{(OLS,T)}, \qquad \hat{\beta}_{(mean)}^{(LWS,T,w)} = \frac{1}{100}\sum_{k=1}^{100} \hat{\beta}_{(k)}^{(LWS,T,w)}$$

and

$$\hat{\beta}_{(mean)}^{(IWV,T,w)} = \frac{1}{100}\sum_{k=1}^{100} \hat{\beta}_{(k)}^{(IWV,T,w)}.$$

At the end of realization of given framework we considered only the IWV-estimates. As for each framework 100 experiments were performed, we obtained 100 estimates $\hat{\beta}_j^{IWV}$, $j = 1, 2, .., p$ of regression coefficients and hence we may evaluate the "empirical variance" of these estimates and compare them with the estimate of variances given by (23). In other words, we evaluated

$$\widehat{\text{var}}(\hat{\beta}_j) = \frac{1}{100}\sum_{k=1}^{100} \left([\hat{\beta}_{(k)}^{(IWV,T,w)}]_j - [\hat{\beta}_{(mean)}^{(IWV,T,w)}]_j \right)^2, \quad j = 1, 2, ..., p. \tag{26}$$

All calculations including generating random numbers were performed employing MATLAB. The software is available on request. The subroutine for generating weights is given, as an example in the Appendix.
The results are collected below, successively for 3 frameworks.

1. framework $T = 50, \;\; p = 2$ and $\alpha = 0.3$.

True values of regression coefficients					
$\beta_1^0 = -3.7$			$\beta_2^0 = 4.8$		
Ordinary Least Squares		*Least Weighted Squares*		*Instrumental Weighted Vars*	
$\hat{\beta}_1$	$\hat{\beta}_2$	$\hat{\beta}_1$	$\hat{\beta}_2$	$\hat{\beta}_1$	$\hat{\beta}_2$
-4.51	2.86	-4.28	2.63	-3.58	4.67

Instrumental Weighted Variables - variances of the estimators of regression coeffs			
Empirical values		*Estimates by (23)*	
$\widehat{\text{var}}(\hat{\beta}_1)$	$\widehat{\text{var}}(\hat{\beta}_2)$	$\widehat{\text{var}}(\hat{\beta}_1)$	$\widehat{\text{var}}(\hat{\beta}_2)$
1.350	2.764	1.304	2.647

2. framework $T = 80, \;\; p = 3$ and $\alpha = 0.2$.

True values of regression coefficients								
$\beta_1^0 = -4.1$			$\beta_2^0 = 1.9$			$\beta_3^0 = -3.2$		
Ordinary Least Squares			*Least Weighted Squares*			*Instrumental Weighted Vars*		
$\hat{\beta}_1$	$\hat{\beta}_2$	$\hat{\beta}_3$	$\hat{\beta}_1$	$\hat{\beta}_2$	$\hat{\beta}_3$	$\hat{\beta}_1$	$\hat{\beta}_2$	$\hat{\beta}_3$
-3.51	2.75	-4.33	-3.78	2.98	-4.59	-4.02	2.04	-3.36

Instrumental Weighted Variables - variances of the estimators of regression coeffs					
Empirical values			*Estimates by (23)*		
$\widehat{\text{var}}(\hat{\beta}_1)$	$\widehat{\text{var}}(\hat{\beta}_2)$	$\widehat{\text{var}}(\hat{\beta}_3)$	$\widehat{\text{var}}(\hat{\beta}_1)$	$\widehat{\text{var}}(\hat{\beta}_2)$	$\widehat{\text{var}}(\hat{\beta}_3)$
2.621	0.967	2.055	2.227	0.971	2.414

3. framework $T = 100, \;\; p = 4$ and $\alpha = 0.4$.

True values of regression coefficients											
$\beta_1^0 = 3.2 \quad \beta_2^0 = -1.1 \quad \beta_3^0 = -2.6 \quad \beta_4^0 = 4.8$											
Estimates of regression coefficients											
Ordinary Least Squares				*Least Weighted Squares*				*Instrumental Weighted Vars*			
$\hat{\beta}_1$	$\hat{\beta}_2$	$\hat{\beta}_3$	$\hat{\beta}_4$	$\hat{\beta}_1$	$\hat{\beta}_2$	$\hat{\beta}_3$	$\hat{\beta}_4$	$\hat{\beta}_1$	$\hat{\beta}_2$	$\hat{\beta}_3$	$\hat{\beta}_4$
4.57	-2.75	-2.34	3.48	3.92	-2.01	-3.02	2.94	3.36	-1.18	-2.48	4.67

Instrumental Weighted Variables - variances of the estimators of regression coeffs							
Empirical values				*Estimates by (23)*			
$\widehat{\text{var}}(\hat{\beta}_1)$	$\widehat{\text{var}}(\hat{\beta}_2)$	$\widehat{\text{var}}(\hat{\beta}_3)$	$\widehat{\text{var}}(\hat{\beta}_4)$	$\widehat{\text{var}}(\hat{\beta}_1)$	$\widehat{\text{var}}(\hat{\beta}_2)$	$\widehat{\text{var}}(\hat{\beta}_3)$	$\widehat{\text{var}}(\hat{\beta}_4)$
2.905	2.001	2.071	3.022	3.006	1.999	1.968	3.204

7 Acknowledgement

We would like to express our gratitude to the anonymous referees for carefully reading the manuscript. We also thank for the suggestions which led to the general improvement of the paper, in order that this, as we hope, is easier to understand.

8 Appendix

The asymptotic representation (27) under homoscedasticity was already proved in [35] under homoscedasticity. Nevertheless, we need it in a bit more general way, under heteroscedasticity. The way for proving it was in fact opened by [41].

Lemma 3 *Let the conditions $\mathcal{C}1$, $\mathcal{C}2$, $\mathcal{C}3$, $\mathcal{C}4$, $\mathcal{NC}1$, $\mathcal{NC}2$, $\mathcal{AC}1$ and $\mathcal{AC}2$ hold. Then*

$$\sqrt{n}\left(\hat{\beta}^{(IWV,n,w)} - \beta^0\right) =$$

$$\left[\frac{1}{n}\sum_{i=1}^{n} w\left(\overline{F}_{n,\beta}(^0|e_i|)\right) \cdot Z_i X_i'\right]^{-1} \cdot \frac{1}{\sqrt{n}}\sum_{i=1}^{n} w\left(\overline{F}_{n,\beta^0}(|e_i|)\right) \cdot Z_i e_i + o_p(1) \qquad (27)$$

as $n \to \infty$.

For the proof see [39].

8.1 Routine for generating data

In this section we offer a pattern of software, namely routine which indicates how the weights were created. They were generated by means of subroutine.

```
% Generating weights
% T - number of observations
% h - number of observations with large weights from 1 to 1-h/k
% h-g - number of observations with rapidly decreasing weights
% T-h-g - observations with zero weights
% k - level to go down with large weights

function [w]=Weights(h,g,T,k)
L=(1-h/k)/(g-h+1);
for i=1:T; w(i,1)=0; end f=g-h; for i=1:f; w(i+h,1)=1-h/k-i*L; end
for i=1:h; w(i,1)=1-i/k; end
```

For all frameworks we put $h = T - 7$, $h - g = 5$ and $k = 10 \cdot h$. Other software is available on request.

References

[1] Anděl, J., J. Zichová (2002): A method foe estimating parameter in nonnegative $MA(1)$ models. *Communications in Satistics - Theory and Methods, 31, 2101 - 2111.*

[2] Arellano, M., O. Bover (1995): Another look at the instrumental variables estimation of error components models. *Journal of Econometrics, 68, 1, 29 - 52.*

[3] Bickel, P. J. (1975): One-step Huber estimates in the linear model. *J. Amer. Statist. Assoc. 70, 428–433.*

[4] Bowden, R. J., D. A. Turkington (1984): *Instrumental Variables.* Cambridge: Cambridge University Press.

[5] Cochrane, D., G. H. Orcutt (1949): Application of least squares regression to relationhips containing autocorrelated disturbances. *Journal of the American Statistical Association, 44, 32 - 61.*

[6] Draper, N. R., H. Smith (1998): *Applied Regression Analysis.* J.Wiley *&* Sons, New York, 3rd edition.

[7] Durbin, J. (1960): Estimation of parameters in time-series regression models. *J. of Royal Statistical Society, Series B, 22, 139 - 153.*

[8] Fisher, R. A. (1922): On the mathematical foundations of theoretical statistics. *Philos. Trans. Roy. Soc. London Ser. A 222, 309–368.*

[9] Greene, W. H. (1993): *Econometric Analysis*, Macmillam Press, New York.

[10] Hájek, J., Z. Šidák (1967): *Theory of Rank Test.* Academic Press, New York.

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

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

[13] Huber, P.J.(1981): *Robust Statistics.* J.Wiley *&* Sons, New York.

[14] Judge, G. G., W. E. Griffiths, R. C. Hill, H. Lűtkepohl, T. C. Lee (1982): *Introduction to the Theory and Practice of Econometrics.* New York: J.Wiley *&* Sons.

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

[16] Kalina, J. (2004): Durbin-Watson Test for Least Weighted Squares. *Proceedings of COMPSTAT 2004, Physica-Verlag/Springer, 1287 - 1294.*

[17] Maddala, G.S. (2001): *Introduction to Econometrics*, Prentice Hall International, London.

[18] Mizon, G. E. (1995): A simple message for autocorrelation correctors: Don't. *Journal of Econometrics 69, 267 - 288.*

[19] Nievergelt, Y. (1994) Total least squares: State-of-the-art regression in numerical analysis. *SIAM Review 36, 258 - 264.*

[20] Paige, C. C., Strakoš, Z. (2002): Scaled total least squares fundamentals. *Numerische Mathematik 91, 117 - 146.*

[21] Plát, P. (2008): Robustified Hausman's test. *Preprint.*

[22] Prais, S. J., C. B. Winsten (1954): Trend estimators and serial correlation. *Cowless Commission Discussion Paper No 383, Chicago.*

[23] Ronchetti, E. M. (1982): *Robust Testing in Linear Models: The Infinitesimal approach.* Dissertation.

[24] Rousseeuw, P. J., A. M. Leroy (1987): *Robust Regression and Outlier Detection.* J.Wiley *&* Sons, New York.

[25] Sargan, J. D. (1988): Testing for misspecification after estimating using instrumental variables. In: *Massouumi, E. (ed.) Contribution to Econometrics: John Denis Sargan, Vol. 1, Cambridge University Press.*

[26] Stock, J. H., Trebbi, F. (2003): Who invented instrumental variable regression? *Journal of Economic Perspectives 17, 177 - 194.*

[27] Van Huffel, S. (2004): Total least squares and error-in-variables modelling: Bridging the gap between statistics, computational mathematics and enginnering. *Proceedings in Computational Statistics, COMPSTAT 2004, ed. J. Antoch, Physica-Verlag, Springer, 539 - 555.*

[28] Víšek, J.Á. (1994): A cautionary note on the method of Least Median of Squares reconsidered, *Transactions of the Twelfth Prague Conference on Information Theory, Statistical Decision Functions and Random Processes, Lachout, P., Víšek, J.Á. (eds), Academy of Sciences of the Czech Republic, Prague, 1994, 254 - 259.*

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

[30] Víšek, J.Á. (2000): Regression with high breakdown point. In: *Antoch, J., Dohnal, G. (eds) Robust 2000, 324 - 356. Union of Czechoslovak Mathematicians and Physicists, Prague (2001).*

[31] Víšek, J.Á. (2002): LWS, asymptotic linearity, consistency, asymptotic normality. *Bulletin of the Czech Econometric Soc., vol.* ***9****, no.15, 31 - 58, no.16, 1 - 28 (2002).*

[32] Víšek, J.Á. (2003): Durbin-Watson statistic in robust regression. *Probability and Mathematical Statistics, vol. 23., Fasc. 2(2003), 435 - 483.*

[33] Víšek, J.Á. (2004): Robustifying instrumental variables. *Proceedings of COMPSTAT'2004. Physica-Verlag/Springer, 1947 - 1954.*

[34] Víšek, J.Á. (2006): The least trimmed squares. Sensitivity study. *Proceedings of the Prague Stochastics 2006, eds. Marie Hušková & Martin Janžura, matfyzpress, 728-738.*

[35] Víšek, J.Á. (2006): Instrumental weighted variables. *Austrian Journal of Statistics, vol 35(2006), No. 2& 3, 379 - 387.*

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

[37] Víšek, J. Á. (2009): Consistency of the instrumental weighted variables. *Annals of the Institute of Statistical Mathematics, (2009) 61, 543 - 578.*

[38] Víšek, J. Á. (2009): $\sqrt{n}$-consistency of the instrumental weighted variables. *Proceedings of the conference Statistics: Investment in the Future 2, CD-book, 1-15.*

[39] Víšek, J. Á. (2010): Asymptotic representation of the instrumental weighted variables. *Discussion paper.*

[40] Víšek, J. Á. (2010): White test for the instrumental weighted variables. *Discussion paper*, presented on ICORS 2006.

[41] Víšek, J. Á. (2010): Empirical distribution function under heteroscedasticity. To appear in *Statistics*.

[42] White, H. (1980): A heteroskedasticity-consistent covariance matrix estimator and a direct test for heteroscedasticity. *Econometrica, 48, 817 - 838.*

[43] Wooldridge, J. M. (2001): *Econometric Analysis of Cross Section and Panel Data.* MIT Press, Cambridge, Massachusetts.

[44] Zichová, J. (1996): On a method of estimating parameters in non-negative ARMA models. *Kybernetika 32, 409 - 424.*

Address: FSV UK, Opletalova ul. 26, 110 01 Praha 1
e-mail: visek@fsv.cuni.cz