8-2012

A Robust Hausman-Taylor Estimator

Badi Baltagi
Syracuse University, bbaltagi@maxwell.syr.edu

Georges Bresson
Universite Paris II

ISSN: 1525-3066

Center for Policy Research
Working Paper No. 140

A ROBUST HAUSMAN-TAYLOR ESTIMATOR

Badi H. Baltagi and Georges Bresson

Center for Policy Research
Maxwell School of Citizenship and Public Affairs
Syracuse University
426 Eggers Hall
Syracuse, New York 13244-1020
(315) 443-3114 | Fax (315) 443-1081
e-mail: ctrpol@syr.edu

August 2012

$5.00

CENTER FOR POLICY RESEARCH – Fall 2012

Leonard M. Lopoo, Director
Associate Professor of Public Administration and International Affairs (PAIA)

Associate Directors

Margaret Austin
Associate Director
Budget and Administration

Douglas Wolf
Gerald B. Cramer Professor of Aging Studies
Associate Director, Aging Studies Program

John Yinger
Professor of Economics and PAIA
Associate Director, Metropolitan Studies Program

SENIOR RESEARCH ASSOCIATES

Badi H. Baltagi .... Economics
Robert Bifulco .... PAIA
Leonard Burman .... PAIA/Economics
Thomas Dennison .... PAIA
William Duncombe .... PAIA
Gary Engelhardt .... Economics
Madonna Harrington Meyer .... Sociology
Christine Himes .... Sociology
William C. Horrace .... Economics
Duke Kao .... Economics
Eric Kingson .... Social Work
Sharon Kioko .... PAIA
Thomas Kniesner .... Economics
Jeffrey Kubik .... Economics
Andrew London .... Sociology
Amy Lutz .... Sociology
Yingyi Ma .... Sociology
Jerry Miner .... Economics
Jan Ondrich .... Economics
John Palmer .... PAIA
Eleonora Patacchini .... Economics
David Popp .... PAIA
Christopher Rohlfs .... Economics
Stuart Rosenthal .... Economics
Ross Rubenstein .... PAIA
Perry Singleton....................................... Economics
Michael Wasylenko .... Economics
Jeffrey Weinstein....................................... Economics
Peter Wilcoxen .... PAIA/Economics
Janet Wilmoth .... Sociology

GRADUATE ASSOCIATES

Douglas Abbott .... PAIA
Kanika Arora .... PAIA
Dana Balter .... PAIA
Christian Buerger .... PAIA
Gillian Cantor .... PAIA
Mary Doohovskoy .... PAIA
Alissa Dubnicki .... Economics
Pallab Ghosh .... Economics
Lincoln Groves .... PAIA
Clorise Harvey .... PAIA
Jessica Hausauer .... Sociology
Hee Seung Lee .... PAIA
Chun Chieh Hu .... Economics
Jiayu Li .... Sociology
Jing Li .... Economics
Shimeng Liu .... Economics
Allison Marier .... Economics
Qing Miao .... PAIA
Nuno Abreu Faro E Mota .... Economics
Judson Murchie .... PAIA
Mariylin Nyanteh .... PAIA
Kerri Raissian .... PAIA
Laura Rodriquez-Ortiz .... PAIA
Kelly Stevens .... PAIA
Tian Tang .... PAIA
Liu Tian .... Economics
Mallory Vachon .... Economics
Pengju Zhang .... PAIA

STAFF

Kelly Bogart....................Administrative Specialist
Karen Cimilluca....................Office Coordinator
Alison Kirsche....................Administrative Secretary
Kathleen Nasto....................Administrative Secretary
Candi Patterson....................Computer Consultant
Mary Santy....................Administrative Secretary

Abstract

This paper suggests a robust Hausman and Taylor (1981) estimator, here-after HT, that deals with the possible presence of outliers. This entails two modifications of the classical HT estimator. The first modification uses the Bramati and Croux (2007) robust Within MS estimator instead of the Within estimator in the first stage of the HT estimator. The second modification uses the robust Wagenvoort and Waldmann (2002) two stage generalized MS estimator instead of the 2SLS estimator in the second step of the HT estimator. Monte Carlo simulations show that, in the presence of vertical outliers or bad leverage points, the robust HT estimator yields large gains in MSE as compared to its classical Hausman-Taylor counterpart. We illustrate this robust version of the Hausman-Taylor estimator using an empirical application.

JEL No. C23, C26

Key Words: Bad leverage points, Hausman-Taylor, panel data, two stage generalized MS estimator, vertical outliers.

Badi Baltagi-Department of Economics and Center for Policy Research 426 Eggers Hall, Syracuse University, Syracuse, NY 13244-1020, USA
Georges Bresson-Department of Economics, Université Paris II / Sorbonne Université 12 place du Panthéon 75 230 Paris Cedex 05, France
Corresponding author. Tel.: + 1 315 443 1630; Fax: + 1 315 443 1081.
E-mail: bbaltagi@maxwell.syr.edu (B.H. Baltagi), bresson-georges@orange.fr (G. Bresson).

We would like to thank three referees for their helpful comments and suggestions. Many thanks to the participants of the 11th Advances in Econometrics Conference: Essays in Honor of Jerry Hausman, Louisiana State University, Baton Rouge, February 17-19, 2012. We thank Christophe Croux for providing his Matlab codes and for helpful discussion on the Within MS estimator. We dedicate this paper in memory of Halbert L. White Jr.

1 Introduction

It is well known in the statistical literature that the presence of outlying observations can strongly distort the classical least squares estimator and lead to unreliable inference. Three types of outliers that influence the least squares estimator are vertical outliers, good leverage points and bad leverage points (see Rousseeuw and Leroy (2003)). Vertical outliers are observations that have outlying values for the corresponding error term (the y-dimension) but are not outlying in the design space (the X-dimension). They contaminate the estimation of the intercept but only mildly influence that of the regression coefficients. Good leverage points are observations that are outlying in the design space but are located close to the regression line. They marginally influence the estimation of both the intercept and the regression coefficients but they affect inference. In contrast, bad leverage points are observations located far away from the regression line. They contaminate the least squares estimation for both the intercept and the slopes (see Dehon, Gassner and Verardi (2009)).

The focus of this paper is on panel data regression methods based on estimators such as fixed effects or random effects least squares that control for heterogeneity of the individuals, but are sensitive to data contamination and outliers like any least squares procedure (see Ronchetti and Trojani (2001)). This sensitivity can be characterized by measures of robustness such as the breakdown point, which evaluates the smallest contaminated fraction of a sample that can arbitrarily change the estimates (see Huber (1981), Donoho and Huber (1983), and Huber and Ronchetti (2009) to mention a few).[1] Since the breakdown point of linear estimators such as least squares is asymptotically zero, the statistical literature has stressed the importance of robust and positive breakdown-point methods (*e.g.*, Wagenvoort and Waldmann (2002), Maronna et al. (2006), and Čížek (2008) to mention a few).

Panel data also suffer from data contamination and outliers. Besides paying attention to vertical outliers or bad leverage points, one has to pay attention to block-concentrated outliers (block-concentrated bad leverage points). In the latter case, most of the vertical outliers (bad leverage points) are concentrated on few individuals, but for most of the time period we observe these individuals. There are only few studies dealing with these problems using panel data. Wagenvoort and Waldmann (2002) and Lucas *et al.* (2007)

[1] See appendix 1.

studied the bounded-influence estimation of static and dynamic panel data models, respectively. Wagenvoort and Waldmann (2002) developed two estimation procedures: the two stage generalized M (2SGM) estimator and the robust generalized method of moments (RGMM) estimator. Both estimators are B-robust, *i.e.* their associated influence function is bounded, consistent and asymptotic normally distributed. For dynamic panel data models, Lucas *et al.* (2007) proposed a variant of the GMM estimator which is less sensitive to anomalous observations. Positive breakdown-point methods for static and dynamic panel models were proposed by Bramati and Croux (2007), Dhaene and Zhu (2009) and Aquaro and Čížek (2010). The Within MS (WMS) estimator proposed by Bramati and Croux (2007) is the robust counterpart of the least squares dummy variables representation of the Within group estimator[2]. Using Monte Carlo simulations, they observe that, without outliers, the efficiency of the robust estimator is very close to that of the Within group estimator. However, the Within estimator performs badly when there are vertical outliers and even worse in the presence of bad leverage points. In contrast, the WMS estimator performs well and yields stable results over different sampling schemes. The Bramati and Croux (2007) WMS estimator yields large gains in MSE with respect to the classical Within estimator in the presence of outliers, and leads to very small efficiency loss in the absence of outliers.

This paper proposes a robust version of the Hausman and Taylor (1981) estimator, hereafter HT. Briefly, the HT panel data estimator deals with the common empirical fact that some of our explanatory variables are time varying, while others are time invariant. In addition, some are correlated with the individual effects and some are not. HT proposed a two-step instrumental variable procedure that is (i) more efficient than the within estimator and (ii) recaptures the effects of time invariant variables which are wiped out by the within transformation. The structure of the paper is as follows: in section 2, we present the HT estimator and in section 3, we briefly review the M, MS and GM robust estimators. Section 4 proposes a robust HT estimator that deals with the possible presence of outliers. This entails two modifications of the classical HT estimator. The first modification uses the Bramati and Croux (2007) robust WMS estimator instead of the Within estimator in the

[2] M, S and MS estimators are discussed in details in the robust statistics literature, see Huber (1964), Rousseeuw and Yohai (1987) and Maronna and Yohai (2006) to mention a few. The corresponding WMS estimator for panel data is discussed in details in section 4.1.

first stage of the HT estimator[3].The second modification uses the robust Wagenvoort and Waldmann (2002) two stage generalized MS-estimate instead of the 2SLS estimate in the second step of the HT estimator. In section 5, we run Monte Carlo simulations to study the effects of vertical outliers, bad leverage points, block-concentrated outliers or block-concentrated bad leverage points on the classical and robust HT estimators. We show that the robust HT yields large gains in MSE as compared to its classical Hausman-Taylor counterpart. In section 6, we apply our robust Hausman-Taylor estimator to the Cornwell and Rupert (1988) estimation of a Mincer wage equation. Finally, section 7 concludes.

2 The Hausman-Taylor estimator

Hausman and Taylor (1981), hereafter HT, considered the following model where some of the explanatory variables are time varying (X_{it}), while others are time invariant (Z_i):

$$y_{it} = X_{it}'\beta + Z_i'\gamma + \mu_i + \nu_{it} \ , \ i = 1, ..., N, \ t = 1, ..., T \tag{1}$$

μ_i is IID($0,\sigma_\mu^2$), ν_{it} is IID($0,\sigma_\nu^2$) independent of each other and among themselves. HT allowed *some* of the X and Z variables to be correlated with the individual effects μ_i. This is in contrast to the fixed effects estimator where *all* the regressors are correlated with the individual effects, and the random effects estimator where *none* of the regressors are correlated with the individual effects. Using the HT notation: $X = [X_1, X_2]$ and $Z = [Z_1, Z_2]$ where X_1 is $(NT \times k_1)$, X_2 is $(NT \times k_2)$, Z_1 is $(NT \times g_1)$ and Z_2 is $(NT \times g_2)$. X_1 and Z_1 are assumed exogenous in that they are not correlated with μ_i and ν_{it}, while X_2 and Z_2 are endogenous because they are correlated with μ_i, but not ν_{it}.

HT proposed the following two-step consistent estimator[4] of β and γ:

[3] Aquaro and Čížek (2010) use a first difference rather than a Within transformation. Their simulations reveal superior performance over the median difference estimator. However, differencing eliminates the first wave, and in micro-panels that is a loss of N observations. Differencing is usually not employed in panel data unless the model is dynamic. In keeping with the spirit of the Hausman-Taylor approach that uses the Within estimator in the first stage, and in order not to waist N observations, we use a robust Within approach rather than an approach based on first differences or pairwise-differences.

[4] See Cornwell and Rupert (1988), Egger and Pfaffermayr (2004) and Serlenga and Shin (2007), to mention a few applications of the HT estimator.

1. Perform the fixed effects (FE) or Within estimator obtained by regressing $\widetilde{y}_{it} = (y_{it} - \overline{y}_{i.})$, where $\overline{y}_{i.} = \sum_{t=1}^{T} y_{it}/T$, on a similar within transformation on the regressors. Note that the Within transformation wipes out the Z_i variables since they are time invariant, and we only obtain an estimate of β which we denote by $\widetilde{\beta}_W$.

   - Then, HT average the within residuals over time

   $$\widehat{d}_i = \bar{y}_{i.} - \bar{X}_{i.}'\widetilde{\beta}_W \tag{2}$$

   - To get an estimate of γ, HT suggest running 2SLS of $\widehat{d}_i$ on Z_i with the set of instruments $A = [X_1, Z_1]$. This yields

   $$\widehat{\gamma}_{2SLS} = (Z'P_A Z)^{-1} Z'P_A \widehat{d} \tag{3}$$

   where $P_A = A(A'A)^{-1}A'$. It is clear that the order condition has to hold ($k_1 \geqslant g_2$) for $(Z'P_A Z)$ to be nonsingular. In fact, if $k_1 = g_2$, then the model is just-identified and one stops here.

2. If $k_1 > g_2$, HT suggest estimating the variance-components as follows:

   $$\widehat{\sigma}_\nu^2 = (y_{it} - X_{it}'\widetilde{\beta}_W)'Q(y_{it} - X_{it}'\widetilde{\beta}_W)/N(T-1) \tag{4}$$

   and

   $$\widehat{\sigma}_1^2 = (y_{it} - X_{it}'\widetilde{\beta}_W - Z_i'\gamma_{2SLS})'P(y_{it} - X_{it}'\widetilde{\beta}_W - Z_i'\gamma_{2SLS})/N \tag{5}$$

   where $\sigma_1^2 = T\sigma_\mu^2 + \sigma_\nu^2$. $P = I_N \otimes \bar{J}_T$ and $\bar{J}_T = J_T/T$, with I_N being a matrix of dimension N, and J_T is a matrix of ones of dimension T. P is a matrix which averages the observation across time for each individual. $Q = I_{NT} - P$. Once the variance-components estimates are obtained, the model is transformed using $\widehat{\Omega}^{-1/2}$ where

   $$\Omega^{-1/2} = \frac{1}{\sigma_1}P + \frac{1}{\sigma_\nu}Q \tag{6}$$

   see Baltagi (2008). Note that $y^* = \widehat{\sigma}_\nu \widehat{\Omega}^{-1/2} y$ has a typical element $y_{it}^* = y_{it} - \widehat{\theta}\bar{y}_{i.}$ where $\widehat{\theta} = 1 - (\widehat{\sigma}_\nu/\widehat{\sigma}_1)$ and X_{it}^* and Z_i^* are defined similarly. In fact, the transformed regression becomes:

   $$\widehat{\sigma}_\nu\widehat{\Omega}^{-1/2} y_{it} = \widehat{\sigma}_\nu\widehat{\Omega}^{-1/2} X_{it}\beta + \widehat{\sigma}_\nu\widehat{\Omega}^{-1/2} Z_i\gamma + \widehat{\sigma}_\nu\widehat{\Omega}^{-1/2} u_{it} \tag{7}$$

where $u_{it} = \mu_i + \nu_{it}$. The asymptotically efficient HT estimator is obtained by running 2SLS on this transformed model using $A_{HT} = [\widetilde{X}, \bar{X}_1, Z_1]$ as the set of instruments. In this case, $\widetilde{X}$ denotes the within transformed X and $\bar{X}_1$ denotes the time average of X_1. More formally, the HT estimator under over-identification is given by:

$$\begin{pmatrix} \widehat{\beta} \\ \widehat{\gamma} \end{pmatrix}_{HT} = \left[\begin{pmatrix} X^{*\prime} \\ Z^{*\prime} \end{pmatrix} P_{A_{HT}} (X^*, Z^*) \right]^{-1} \begin{pmatrix} X^{*\prime} \\ Z^{*\prime} \end{pmatrix} P_{A_{HT}} y^* \qquad (8)$$

where $P_{A_{HT}}$ is the projection matrix on $A_{HT} = [\widetilde{X}, \bar{X}_1, Z_1]$, see also Breusch, Mizon and Schmidt (1989).

3 A brief review of M, MS and GM robust estimators

To robustify the HT estimator for the possible presence of outliers, we used two MS estimators: the one proposed by Bramati and Croux (2007) for the first step of the HT estimator and a two stage generalized MS estimator inspired from Wagenvoort and Waldmann (2002) for the second step of the HT estimator. In this section, we briefly review M and S and MS estimators from the robust statistics literature.

Huber (1964) generalized the median regression to a wider class of estimators, called M-estimators, by considering other functions besides the absolute value of the residuals. This increases Gaussian efficiency while keeping robustness with respect to vertical outliers. Consider the panel data fixed effects regression disturbances: $r_{it}(\alpha, \beta) \equiv r_{it} = y_{it} - X_{it}'\beta - \alpha_i$. The M-estimator is defined as

$$\widehat{\theta}_M \left(\equiv (\alpha, \beta')'\right) = \arg\min_{\alpha,\beta} \sum_{i=1}^{N} \sum_{t=1}^{T} \rho\left(\frac{r_{it}}{\sigma}\right) \qquad (9)$$

To guarantee scale equivariance (*i.e.* independence with respect to the measurement units of the dependent variable), residuals are standardized by a measure of dispersion σ. This can be implemented as an iterative weighted least-squares. The M-estimator of θ based on the function $\rho(.)$ is the vector $\widehat{\theta}_M$ of size $(K \times 1)$ which is the solution of the following system:

$$\min_{\theta} \sum_{i=1}^{N} \sum_{t=1}^{T} \psi\left(\frac{r_{it}}{\sigma}\right) \frac{d(r_{it})}{d\theta_j} = 0 \ , \ j = 1, ..., K \qquad (10)$$

$\psi(u) = \rho'(u)$ is called the influence function. If we define a weight function $W_r(u) = \psi(u)/u$, then $\widehat{\theta}_M$ is a weighted least-squares estimator:

$$\widehat{\theta}_M = \arg\min_{\theta} \sum_{i=1}^{N} \sum_{t=1}^{T} W_r(r_{it}) r_{it}^2 \tag{11}$$

and the first order condition which defines the class of M-estimators is given by:

$$\sum_{i=1}^{N} \sum_{t=1}^{T} X'_{it} r_{it} W_r(r_{it}) = 0 \tag{12}$$

The loss function $\rho(.)$ is a symmetric, positive-definite function with a unique minimum at zero. There are several constraints that a robust M-estimator should meet: a) the first is of course to have a bounded influence function $\psi(r_{it}/\sigma)$; b) The robust estimator should be unique. This requires that the function $\rho(.)$ is convex in θ. The literature on robust statistics proposed several specifications for the ρ-function. The choice of the loss function $\rho(.)$ is crucial to having good robustness properties and high Gaussian efficiency. The Tukey biweight function is a common choice[5]:

$$\rho(u) = \begin{cases} \frac{u^2}{2} - \frac{u^4}{2c^2} + \frac{u^6}{6c^4} & \text{if} \quad |u| \leq c \\ \frac{c^2}{6} & \text{if} \quad |u| > c \end{cases} \tag{13}$$

The associated influence function and weight function are defined as:

$$\psi(u) = \begin{cases} u\left[1 - \left(\frac{u}{c}\right)^2\right]^2 & \text{if} \quad |u| \leq c \\ 0 & \text{if} \quad |u| > c \end{cases} \text{ and } W_r(u) = \begin{cases} \left[1 - \left(\frac{u}{c}\right)^2\right]^2 & \text{if} \quad |u| \leq c \\ 0 & \text{if} \quad |u| > c \end{cases} \tag{14}$$

In this case, the high breakdown point M-estimator is defined as:

$$\widehat{\theta}_M = \left(X'W_rX\right)^{-1} X'W_r y \tag{15}$$

where y is the $(NT \times 1)$ vector denoting the dependent variable, and X is the $(NT \times K)$ matrix of the explanatory variables. W_r is an $(NT \times NT)$ matrix with diagonal elements given by:

$$W_r(r_{it}) = \begin{cases} \left[1 - \left(\frac{r_{it}}{c\sigma}\right)^2\right]^2 & \text{if} \quad \left|\frac{r_{it}}{\sigma}\right| \leq c \\ 0 & \text{if} \quad \left|\frac{r_{it}}{\sigma}\right| > c \end{cases} \tag{16}$$

[5] For the advantages and disadvantages of several influence functions considered in this literature, see Rousseeuw and Leroy (2003).

For the tuning constant $c = 2.937$ (or $c = 1.547$), the corresponding M-estimator resists contamination up to 25% (or up to 50%) of outliers. In other words, it is said to have a breakdown point of 25% (or 50%). Unfortunately, this M-estimator suffers from some deficiencies. If it is able to identify isolated outliers, it is inappropriate in case of the existence of clusters of outliers, i.e., where one outlier can mask the presence of another. Hence, it is not guaranteed to identify all leverage points. Furthermore, the initial values for the iterative reweighted least squares algorithm are monotone M-estimators that are not robust to bad leverage points and may cause the algorithm to converge to a local instead of a global minimum (see Croux and Verardi (2008)).[6]

Rousseeuw and Yohai (1987) proposed minimizing a measure of dispersion of the residuals that is less sensitive to extreme values. They call this class of estimators the S-estimators. In order to increase robustness, they suggest finding the smallest robust scale of the residuals. This robust dispersion, that will be called $\widehat{\sigma}^S$, satisfies

$$\frac{1}{NT}\sum_{i=1}^{N}\sum_{t=1}^{T}\rho\left(\frac{r_{it}\left(\alpha,\beta\right)}{\widehat{\sigma}_S}\right) = b \tag{17}$$

where $b = E\left[\rho\left(Q\right)\right]$ with $Q \sim N(0,1)$. The value of θ that minimizes $\widehat{\sigma}_S$ is then called an S-estimator defined as:

$$\widehat{\theta}_M^S = \arg\min_{\theta}\ \widehat{\sigma}_S\left(r_{11}\left(\theta\right),...,r_{NT}\left(\theta\right)\right) \tag{18}$$

with the corresponding $\widehat{\sigma}_S$ being the robust estimator of scale.

Rousseeuw and Yohai (1987) computed the asymptotic efficiency of the S-estimator of a Gaussian model for different values of the breakdown point (see appendix 2). Unfortunately, this S-estimator has a Gaussian efficiency of only 28.7%. If the tuning constant (c) of the Tukey biweight loss function $\rho\left(.\right)$ is high, for instance $c = 5.182$, the Gaussian efficiency climbes to 96.6% but the breakdown point drops to 10%.[7]

[6]M-estimators are called monotone if the loss function ρ is convex over the entire domain and are called redescending if the influence function ψ is bounded. Redescending M-estimators have high breakdown points (close to 0.5) and their influence function can be chosen to redescend smoothly to 0 as for the Tukey biweight function.

[7]Monotone M-estimators are robust to outliers in the response variable, but are not resistant to outliers in the explanatory variables (leverage points). In contrast, redescend-

To cope with this, Yohai (1987) introduced M-estimators that combine high-breakdown point and high efficiency. These estimators are redescending M-estimators, but where the scale is fixed at $\widehat{\sigma}_S$. The preliminary S-estimator guarantees a high breakdown point and the final M-estimate allows a high Gaussian efficiency. Following the proposition of Rousseeuw and Yohai (1987), the tuning constant can be set to $c = 1.547$ for the S-estimator to guarantee a 50% breakdown point, and it can be set to $c = 5.182$ for the second step M-estimator to guarantee 96% efficiency of the final estimator.
Generally, the S and M-estimator use the algorithm of Salibian-Barrera and Yohai (2006) (see also Maronna and Yohai (2006)). The algorithm starts by randomly picking p subsets of K observations where K is the number of regression parameters to estimate. For each of the p-subsets, residuals are computed and a scale estimate $\widehat{\sigma}_S$ is obtained. An approximation for the final scale estimate $\widehat{\sigma}_S$ is then given by the value that leads to the smallest scale over all p-subsets[8]. Maronna and Yohai (2006) introduce the MS-estimator

ing M-estimators are resistant to bad leverage points but are difficult to implement from a computational point of view. S-estimation, which finds an hyperplane that minimizes a robust estimate of the scale of the residuals, is highly resistant to leverage points, and is robust to outliers in the response. However, this method can be inefficient. MM-estimation (not used here) tries to capture both the robustness and resistance of S-estimation, while at the same time gaining the efficiency of M-estimation. The method proceeds in 3 steps: a) with a first loss function, we get an initial M-estimator, b) we obtain an M-estimate of the scale of the residuals, c) the estimated scale is then held constant whilst an M-estimate of the parameters is located with a new loss function. MM-estimators are robust and efficient.

[8] From equation (18), the algorithm calculates the hyperplane of K observations that fits all points perfectly if all K points are regular observations and do not contain outliers. For each subset, the residuals are defined as the vertical distance separating each observation from the hyperplane. Using these residuals, a scale estimate $\widehat{\sigma}_S$ is obtained as in (17) for each p-subset. Salibian-Barrera and Yohai (2006) proposed the following number of generated sub-samples $N_{\mathbf{sub}}$:

$$N_{sub} = \left\lceil \frac{\log(1-P)}{\log\left(1-(1-\nu)^K\right)} \right\rceil$$

where ν is the maximal expected proportion of outliers. P is the desired probability of having at least one p-subset without outliers among the N_{sub} subsamples and $\lceil x \rceil$ is the ceiling operator of x, i.e., the smallest integer not less than x. The number of sub-samples is chosen to guarantee that at least one p-subset without outliers is selected with high probability (see Salibian-Barrera and Yohai (2006), Maronna and Yohai (2006), Croux and Verardi (2008)). In our Monte-Carlo study, we use $N_{sub} = 500$. As Croux and Verardi that alternates an S-estimator and an M-estimator, until convergence. This estimator has been adapted for the fixed effects panel data case by Bramati and Croux (2007). They call this estimator the WMS (Within MS) estimator. This will be our estimator in place of the classical within estimator in the first step of our robust Hausman-Taylor estimator.

Hinloopen and Wagenvoort (1997) proposed further protection against observations with a high leverage. They suggest using location weights indirectly proportional to the values of covariates:

$$W_x\left(X_{it}\right)=\min\left(1,\frac{\sqrt{\chi^2_{K,0.975}}}{RD_{it}}\right) \tag{19}$$

where $\chi^2_{K,0.975}$ is the upper 97.5% quantile of a chi-squared distribution with K degrees of freedom,

$$RD_{it}=\sqrt{\left(X_{it}-\widehat{\mu}_x\right)' V_x^{-1}\left(X_{it}-\widehat{\mu}_x\right)} \tag{20}$$

is a robust version of the Malahanobis distance (or Rao's distance) and $\widehat{\mu}_x$ and V_x are the robust estimates of the location and variance matrix[9] of X_{it}. Wagenvoort and Waldmann (2002) proposed the use of this class of generalized M-estimators (GM hereafter)[10]. The first order condition which

(2008) warn, subsampling algorithms can easily lead to collinear sub-samples if various dummies are present. A rough solution is to use subsets of size a little bit larger than K. An exact solution is given by Maronna and Yohai (2006) who introduce the MS-estimator that alternates an S-estimator (for continuous variables) and an M-estimator (for dummy ones), till convergence.

[9] The robust estimates $\widehat{\mu}_x$ and V_x can be obtained using the Minimum Covariance Determinant (MCD) estimator (see Rousseeuw (1984)). The MCD method looks for the $h(> NT/2)$ observations whose classical covariance matrix has the lowest possible determinant. The raw MCD estimate of location $\widehat{\mu}_x$ is then the average of these h points, whereas the raw MCD estimate of the scatter V_x is their covariance matrix, multiplied by a consistency factor. The MCD estimates can resist $(NT-h)$ outliers and a value of $h=0.75NT$ is recommended to obtain a high finite-sample efficiency. The computation of the MCD estimator is non-trivial. Rousseeuw and Van Driessen (1999) suggest a fast resampling algorithm (FAST-MCD) that we use. Several other algorithms have been proposed (see Olive (2008), chap.10 for a discussion).

[10] Bramati and Croux (2007) also use this GM estimator for fixed effects panel data model and call it the Within GM (or WGM) estimator. They show that this estimator gives similar results when compared to the WMS estimator.

defines this class of GM estimators is:

$$\sum_{i=1}^{N}\sum_{t=1}^{T} X'_{it} W_x\left(X_{it}\right) r_{it} W_r\left(r_{it}\right) = 0 \tag{21}$$

In this case, the high breakdown point generalized M-estimator is defined as:

$$\widehat{\theta}_{GM} = \left(X' W_x W_r X\right)^{-1} X' W_x W_r y \tag{22}$$

where W_x is the $(NT \times NT)$ matrix with diagonal elements given by $W_x\left(X_{it}\right)$.

4 The robust Hausman-Taylor estimator

To robustify the HT estimator for the possible presence of outliers, two MS estimators are successively used: the one proposed by Bramati and Croux (2007) for the first step of the HT estimator and a two stage generalized MS estimator inspired from Wagenvoort and Waldmann (2002) for the second step of the HT estimator.

4.1 The WMS estimator

The Within MS (WMS) $\widetilde{\beta}_{WMS}$ estimator proposed by Bramati and Croux (2007) is then defined as:

$$\widetilde{\beta}_{WMS} = \arg\min_{\beta} \widehat{\sigma}_S\left(r_{11}\left(\beta\right), ..., r_{NT}\left(\beta\right)\right) \tag{23}$$

with

$$r_{it}\left(\beta\right) = \left(y_{it} - X'_{it}\beta\right) - \text{median}_{t=1}^{T}\left(y_{it} - X'_{it}\beta\right) \tag{24}$$

Given an initial estimate β_0, they use an iterative algorithm to get closer to the minimum of eq.(23). This algorithm is based upon the generation of random subsamples suggested by Maronna and Yohai (2006) to compute the robust scale estimate of the residuals $\widehat{\sigma}_S$. They suggest iterating a fixed number of times ($\max m = 20$), and to choose the $\widetilde{\beta}_{WMS}^{(m)}$ which produces the minimum value of the objective function in (23).
Unfortunately, for the HT model, the WMS estimator, like the within estimator, gives us only an estimate of β and not γ. Once again, the Z_i variables

drop out as they are time invariant. The variance-covariance matrix of the WMS estimate $\widetilde{\beta}_{WMS}$ is given by:

$$Var\left(\widetilde{\beta}_{WMS}\right)=\widehat{\sigma}_S^2\left(\widehat{X}'D_1\widehat{X}\right)^{-1}\widehat{X}'D_2\widehat{X}\left(\widehat{X}'D_1\widehat{X}\right)^{-1} \tag{25}$$

where D_1 and D_2 are diagonal matrices with diagonal elements given by:

$$D_{1,it}=\frac{d}{du_{it}}\left[u_{it}W_r\left(u_{it}\right)\right] \text{ and } D_{2,it}=\left[u_{it}W_r\left(u_{it}\right)\right]^2 \text{ with } u_{it}=\frac{r_{it}}{c\widehat{\sigma}_S} \tag{26}$$

4.2 The two stage generalized MS estimator

Instead of averaging the within residuals over time as HT suggest (see eq.(2)), we take the *median* of the resulting residuals over time:

$$\widehat{r}_i=\text{median}_{t=1}^{T}(y_{it}-X_{it}'\widetilde{\beta}_{WMS}) \tag{27}$$

and instead of the 2SLS procedure suggested by HT, we propose a two stage generalized MS-estimate (2SGMS) following Wagenvoort and Waldmann (2001).

More specifically:

1. (a) *Stage 1*: suppose that there are m_1 instrumental variables A_{it} which are correlated with the explanatory factors Z_i but independent of the error term $\varepsilon_{it}\left(=\mu_i+\nu_{it}\right)$. The explanatory variable Z_k (the k^{th} column of Z) is regressed on the instrumental variables $A=[X_1,Z_{1,}]$: $Z_{it,k}=A_{it}\eta_k+\xi_{it,k}$. The high breakpoint generalized M-estimate (GM) and the prediction of the k^{th} column of Z is computed according to[11]:

$$\widehat{Z}_k=A\left(A'W_A\left(A\right)W_r\left(r_{1,k}\right)A\right)^{-1}A'W_A\left(A\right)W_r\left(r_{1,k}\right)Z_k \tag{28}$$

where $W_A\left(A\right)$ and $W_r\left(r_{1,k}\right)$ are the diagonal matrices comprising the weight functions $W_A\left(A_{it}\right)$ and $W_r\left(r_{1it,k}\right)$. $r_{1,k}$ are the first stage GM residuals associated with Z_k ($r_{1,k}=Z_k-A\widehat{\eta}_k$) and

[11] In our specific case $Z=[Z_1,Z_2]$ and $A=[Z_1,X_{11},X_{12}]$, so $\widehat{Z}_1\equiv Z_1$ and

$$\widehat{Z}_2=A\left(A'W_A\left(A\right)W_r\left(r_{1,2}\right)A\right)^{-1}A'W_A\left(A\right)W_r\left(r_{1,2}\right)Z_2$$

$W_r\left(r_{1,k}\right)$ differs for every distinct column of Z. Thus (g_1+g_2) separate GM regressions are performed if $dim(Z)=g_1+g_2$. Contrary to Wagenvoort and Waldmann (2001), we suggest using the residuals $\left(r_{1,k}\right)$ to estimate a new robust scale estimator of the residuals $\widehat{\sigma}_S$, which is then used to re-estimate a new weight function $W_r\left(r_{1it,k}\right)$, and so on. Following the suggestion of Maronna and Yohai (2006), we compute this iterated MS procedure using a maximum of 20 iterations.

(b) *Stage 2*: replacing the explanatory variables of the original equation by their robust projection on A. This returns the high breakpoint generalized MS-estimator, called the 2SGMS estimator:

$$\widetilde{\gamma}_{2SGMS}=\left(\widehat{Z}'W_Z\left(\widehat{Z}\right)W_r\left(r_2\right)\widehat{Z}\right)^{-1}\widehat{Z}'W_Z\left(\widehat{Z}\right)W_r\left(r_2\right)\widehat{r}_2 \quad (29)$$

where $W_Z\left(\widehat{Z}\right)$ and $W_r\left(r_2\right)$ are diagonal matrices containing the second step GMS weights and r_2 are the second stage GMS residuals $\left(r_2=y-\widehat{Z}\widetilde{\gamma}_{2SGMS}\right)$.

4.2.1 The second step: a two stage generalized MS estimator

The variance-components estimates are obtained as follows:

$$\widetilde{\sigma}^2_\nu=(y_{it}-X'_{it}\widetilde{\beta}_{WMS})'Q(y_{it}-X'_{it}\widetilde{\beta}_{WMS})/N(T-1) \quad (30)$$

and

$$\widetilde{\sigma}^2_1=(y_{it}-X'_{it}\widetilde{\beta}_{WMS}-Z'_i\widetilde{\gamma}_{2SGMS})'P(y_{it}-X'_{it}\widetilde{\beta}_{WMS}-Z'_i\widetilde{\gamma}_{2SGMS})/N \quad (31)$$

where $\sigma^2_1=T\sigma^2_\mu+\sigma^2_\nu$. Once the variance-components estimates are obtained, we compute

$$y^*_{it}=y_{it}-\widetilde{\theta}\bar{y}_{i.} \quad (32)$$

where $\widetilde{\theta}=1-\widetilde{\sigma}_\nu/\widetilde{\sigma}_1$ and X^*_{it} and Z^*_i are defined similarly. The 2SGMS procedure applied to this transformed model can be described as follows:

1. (a) *Stage 1*: each explanatory variable of $V=[X^*_{it},Z^*_i]$ is regressed on the m_2 instrumental variables $A_{HT}=[\widetilde{X},\bar{X}_1,Z_1]$. The k^{th} explanatory variable V_k is regressed on the instrumental variables:

$V_{it,k} = A_{HT_{it}}\delta_k + \xi_{it,k}$. This returns the GM estimate, and the prediction of the k^{th} column of $V = [X^*_{it}, Z^*_i]$ is computed according to:

$$\begin{aligned} \widehat{V}_k &= A_{HT}\left(A'_{HT}W_{A_{HT}}\left(A_{HT}\right)W_r\left(r_{1,k}\right)A_{HT}\right)^{-1} \times \\ &\quad A'_{HT}W_{A_{HT}}\left(A_{HT}\right)W_r\left(r_{1,k}\right)V_k. \end{aligned} \tag{33}$$

$W_{A_{HT}}(A)$ and $W_r(r_{1,k})$ are the diagonal matrices comprising the weight functions $W_{A_{HT}}(A_{HT,it})$ and $W_r(r_{1it,k})$. $r_{1,k}$ are the first stage GM residuals associated with V_k $\left(r_{1,k} = V_k - A_{HT}\widehat{\delta}_k\right)$, and $W_r(r_{1,k})$ differs for every distinct column of V.
Thus $(K = k_1 + k_2 + g_1 + g_2)$ separate GM regressions are performed if $dim(V) = K$. With these residuals $(r_{1,k})$, we estimate a new robust scale estimator of the residuals $\widehat{\sigma}_S$ which is used to re-estimate a new weight function $W_r(r_{1it,k})$, and so on. Following the suggestion of Maronna and Yohai (2006), we compute this iterated MS procedure up to a maximum of 20 iterations.

(b) *Stage 2*: replacing the explanatory variables of the original equation by their robust projection on A_{HT} and applying the GM technique one more time provides the 2SGMS estimates:

$$\widetilde{\lambda}_{2SGM} = \begin{pmatrix} \widetilde{\beta} \\ \widetilde{\gamma} \end{pmatrix}_{2SGMS} = \begin{array}{c} \left(\widehat{V}'W_V\left(\widehat{V}\right)W_r\left(r_2\right)\widehat{V}\right)^{-1} \times \\ \widehat{V}'W_V\left(\widehat{V}\right)W_r\left(r_2\right)y^* \end{array} . \tag{34}$$

$W_V\left(\widehat{V}\right)$ and $W_r(r_2)$ are diagonal matrices containing the second step GMS weights and r_2 are the second stage GMS residuals $\left(r_2 = y^* - \widehat{V}\widetilde{\lambda}_{2SGMS}\right)$.

Following Wagenvoort and Waldmann (2001), the variance-covariance matrix of the 2SGMS estimate $\widetilde{\lambda}_{2SGMS}$ is given by:

$$Var\left(\widetilde{\lambda}_{2SGMS}\right) = \widehat{D}^{-1}M\left(\widehat{D}^{-1}\right)' \tag{35}$$

with

$$\widehat{D} = \widehat{V}'W_V\left(\widehat{V}\right)D_1\widehat{V} \text{ and } M = GG' \tag{36}$$

where

$$G = \left(\widehat{V}'W_V\left(\widehat{V}\right)W_r\left(r_2\right)\widehat{V}\right)\left(R_2+R_3\right) \quad (37)$$
$$-\widehat{\Lambda}'\left(A'_{HT}W_V\left(\widehat{V}\right)W_r\left(r_2\right)A_{HT}\right)BAC$$

and where R_2, R_3, A, $\widehat{\Lambda}$, B and C are $(NT \times NT)$, $(NT \times NT)$, $(Km_2 \times Km_2)$, $(m_2 \times K)$, $(m_2 \times Km_2)$, $(Km_2 \times NT)$ matrices defined as follows:

$$R_2 = diag\left(y_{it} - V_{it}\widetilde{\lambda}_{2SGMS}\right) \quad (38)$$
$$R_3 = diag\left(\left(V_{it} - \widehat{V}_{it}\right)\widetilde{\lambda}_{2SGMS}\right)$$

$$A = diag\left[\left\{A'_{HT}W_{A_{HT}}\left(A_{HT}\right)W_r\left(r_{1,it,k}\right)A_{HT}\right\}^{-1}\right] \;,\; k = 1, ..., K \quad (39)$$

$$\widehat{\Lambda} = [\delta_1, \delta_2, ..., \delta_K] \quad (40)$$

$$B = \widetilde{\lambda}'_{2SGMS} \otimes I_{m_2} \quad (41)$$

$$C = [C_{11}, ..., C_{NT}] \quad (42)$$

where C_{ij} is a $(Km_2 \times 1)$ vector given by

$$C_{ij} = \begin{pmatrix} A'_{HT_{ij}}W_{A_{HT}}\left(A_{HT_{ij}}\right)\left(V_{ij,1} - \widehat{V}_{ij,1}\right)W_r\left(r_{1,it,1}\right) \\ \vdots \\ \vdots \\ A'_{HT_{ij}}W_{A_{HT}}\left(A_{HT_{ij}}\right)\left(V_{ij,K} - \widehat{V}_{ij,K}\right)W_r\left(r_{1,it,K}\right) \end{pmatrix} \quad (43)$$

For ease of comparison, the next table gives the steps for the Hausman-Taylor and the corresponding robust Hausman-Taylor estimator.

Hausman-Taylor

$$
\text{Step 1}\begin{cases}
\text{FE } \widetilde{\beta}_W = \left(\widetilde{X}'\widetilde{X}\right)^{-1}\widetilde{X}'\widetilde{y} \\
\widehat{d}_i = \bar{y}_{i.} - \bar{X}_{i.}'\widetilde{\beta}_W \\
\text{2SLS}\begin{cases}
Z = [Z_1, Z_2],\ A = [X_1, Z_1] \\
\text{stage 1: } \widehat{Z} = P_A Z \\
\text{stage 2: } \widehat{\gamma}_{2SLS} = (\widehat{Z}'\widehat{Z})^{-1}\widehat{Z}'\widehat{d}
\end{cases}
\end{cases}
$$

$$
\text{Step 2}\begin{cases}
\widehat{\sigma}_\nu^2\ ,\ \widehat{\sigma}_1^2\ ,\ \widehat{\theta} \\
y_{it}^* = y_{it} - \widehat{\theta}\bar{y}_{i.} \\
\text{2SLS}\begin{cases}
Z^* = [Z_1^*, Z_2^*],\ A_{HT} = [\widetilde{X}, \bar{X}_1, Z_1] \\
\text{stage 1: } \widehat{Z}^* = P_{A_{HT}} Z^* \\
\text{stage 2: } \begin{pmatrix} \widehat{\beta} \\ \widehat{\gamma} \end{pmatrix}_{HT} = \left(\widehat{Z}^{*\prime}\widehat{Z}^*\right)^{-1}\widehat{Z}^{*\prime} y^*
\end{cases}
\end{cases}
$$

Robust Hausman -Taylor

$$
\text{Step 1}\begin{cases}
\text{WMS } \widetilde{\beta}_{WMS} \\
\widehat{r}_i = \text{median}_{t=1}^T (y_{it} - X_{it}'\widetilde{\beta}_{WMS}) \\
\text{2SGMS}\begin{cases}
Z = [Z_1, Z_2],\ A = [X_1, Z_1] \\
\text{stage 1: } \widehat{Z}_k = P_{A_{ROB}} Z_k \\
\text{with } P_{A_{ROB}} = \begin{array}{l} A\left(A'W_A\left(A\right)W_r\left(r_{1,k}\right)A\right)^{-1}\times \\ A'W_A\left(A\right)W_r\left(r_{1,k}\right) \end{array} \\
\text{stage 2: } \widetilde{\gamma}_{2SGM} = \begin{array}{l} \left(\widehat{Z}'W_Z\left(\widehat{Z}\right)W_r\left(r_2\right)\widehat{Z}\right)^{-1}\times \\ \widehat{Z}'W_Z\left(\widehat{Z}\right)W_r\left(r_2\right)\widehat{r}_2 \end{array} \\
\text{with } r_2 = y - \widehat{Z}\widetilde{\gamma}_{2SGMS}
\end{cases}
\end{cases}
$$

$$
\text{Step 2}\begin{cases}
\widetilde{\sigma}_\nu^2,\ \widetilde{\sigma}_1^2,\ \widetilde{\theta} \\
y_{it}^* = y_{it} - \widetilde{\theta}\bar{y}_{i.} \\
\text{2SGMS}\begin{cases}
V = [X_{it}^*, Z_i^*],\ A_{HT} = [\widetilde{X}, \bar{X}_1, Z_1] \\
\text{stage 1: } \widehat{V}_k = P_{A_{HTROB}} V_k \\
\text{with } P_{A_{HTROB}} = \begin{array}{l} A_{HT}\left(A_{HT}'W_{A_{HT}}\left(A_{HT}\right)W_r\left(r_{1,k}\right)A_{HT}\right)^{-1}\times \\ A_{HT}'W_{A_{HT}}\left(A_{HT}\right)W_r\left(r_{1,k}\right) \end{array} \\
\text{stage 2: } \begin{pmatrix} \widetilde{\beta} \\ \widetilde{\gamma} \end{pmatrix}_{2SGMS} = \begin{array}{l} \left(\widehat{V}'W_V\left(\widehat{V}\right)W_r\left(r_2\right)\widehat{V}\right)^{-1}\times \\ \widehat{V}'W_V\left(\widehat{V}\right)W_r\left(r_2\right)y^* \end{array}
\end{cases}
\end{cases}
$$

5 The simulation study

We first simulate a Hausman-Taylor world (see below) and in the next step, contamination is carried out on the y's only (vertical outliers) and then on both the y and the X variables (leverage) and last on y, X and the Z variables (leverage).

5.1 The DGP of a Hausman-Taylor world

We consider the following panel data regression model with one-way error component disturbances:

$$y_{i,t} = \beta_1 X_{1,i,t} + \beta_2 X_{2,i,t} + \eta_{11} Z_{11,i} + \eta_{12} Z_{12,i} + \eta_2 Z_{2,i} + \mu_i + \nu_{it} \tag{44}$$

with $X_{1,i,t} = [X_{11,i,t}\ X_{12,i,t}]$, $\beta_1 = [\beta_{11}\ \beta_{12}]'$, $X_{1,it}$ is $1 \times k_1$ (here $k_1 = 2$), $X_{2,i,t}$ is $1 \times k_2$ (here $k_2 = 1$), $Z_{1,i}$ is $1 \times g_1$ (here $g_1 = 2$, $Z_{1,i} = [Z_{11,i}, Z_{12,i}]$) and $Z_{2,i}$ is $1 \times g_2$ (here $g_2 = 1$). The $Z_{11,i}$ is a constant ($Z_{11,i} \equiv 5$), $Z_{12,i}$ is a cross-sectional time-invariant $(0,1)$ dummy variable and $Z_{2,i}$ is a cross-sectional time-invariant variable. Our experiments are designed as follows:

- $\beta_{11} = \beta_{12} = \beta_2 = \eta_{11} = \eta_{12} = \eta_2 = 1$.
- μ_i is IIN(0,σ_μ^2), ν_{it} is IIN(0,σ_ν^2), the total variance is fixed $\sigma_u^2 = \sigma_\mu^2 + \sigma_\nu^2 = 3$. The proportion of the total variance due to the individual effects is set at $\rho = \sigma_\mu^2/\sigma_u^2 = 0.5$.
- $N = 100, 200$, and $T = 5, 10$.
- The number of replications is 1000.

The $X_{j,i,t}$ variables are generated by:

$$\begin{aligned} X_{11,i,t} &= \delta_i + \zeta_{i,t} \\ X_{12,i,t} &= \theta_i + w_{i,t} \\ X_{2,i,t} &= \mu_i + \tau_{i,t} \end{aligned} \tag{45}$$

where δ_i, $\zeta_{i,t}$, θ_i, $w_{i,t}$, $\tau_{i,t}$ are uniform on $[-2, 2]$. It is clear that X_2 is correlated with μ_i by construction. The cross-sectional time-invariant $(0,1)$ dummy variable $Z_{12,i}$ has been generated randomly such that its mean is 0.2.

The Hausman-Taylor world is defined with Z_2 correlated with μ_i, X_{11}, X_{12} and X_2.

$$Z_{2,i} = \mu_i + \delta_i + \theta_i + \xi_i \tag{46}$$

where ξ_i is uniform on $[-2, 2]$. So, the $Z_{2,i}$ variable is correlated with $X_{11,i,t}$ (by the term δ_i) with $X_{12,i,t}$ (by the term θ_i) and with $X_{2,i,t}$ (by the term μ_i).

5.2 Contamination

Once the observations are generated for our model in (44), contamination is carried out as follows:

- the y's only (vertical outliers).
- both y and the time-varying explanatory variables (X) by introducing bad leverage points.
- y, X and Z_{12} by introducing bad leverage points.
- y, X, Z_{12} and Z_2 by introducing bad leverage points.

Contamination is generated in two different ways:

- either completely randomly over all observations (*random contamination*),
- or concentrating the contamination in a number of blocks such that half of the observations in the affected time-series are contaminated (*concentrated contamination*). In other words, few individuals in the sample have 50% of their data corrupted while the other individuals have clean observations[12].

Outliers generated by random contamination are either vertical outliers or leverage points, whereas in the case of concentrated contamination, they are either block-concentrated vertical outliers or block-concentrated leverage points.

Vertical outliers are obtained by adding a term $\sim N\left(5\overline{y}, \sigma_y^2/40\right)$ to the y's

[12] For instance, if $N = 100$, $T = 5$ and 10% of the observations are corrupted, it means that 10 individuals have all 5 time period observations corrupted.

originally generated.
Bad leverage points are obtained by replacing X-values (and the Z_{12} and Z_2) corresponding to the observations already contaminated in the y-direction, by points coming from a K-variate normal distribution $N(e_K, 0.5I_K)$, where e_K is a $K \times 1$ vector of ones and I_K is a $K \times K$ identity matrix. We use the Tukey biweight functions $W_r(.)$ for the WMS estimator (eq. 23), for the first stage (eq. 28) and the second stage (eq. 29) of step 1 of the 2SGMS, and for the first stage (eq. 33) and the second stage (eq. 34) of step 2 of the 2SGMS.[13] For all these functions, we need to define the breakdown points and the associated tuning constants. We used the same breakdown point of 25% with a tuning constant $c = 2.937$ yielding an asymptotic efficiency of 76%.[14]
The percentages of contamination considered are 5% and 10%. We report results for the case of no outliers as well as 8 different cases of contamination:

case 1	vertical outliers (y)
case 2	leverage points (y, X_1, X_2)
case 3	concentrated vertical outliers (y)
case 4	concentrated leverage points (y, X_1, X_2)
case 5	leverage points with Z_{12} (y, X_1, X_2, Z_{12})
case 6	concentrated leverage points with Z_{12} (y, X_1, X_2, Z_{12})
case 7	leverage points with Z_{12} and Z_2 (y, X_1, X_2, Z_{12}, Z_2)
case 8	concentrated leverage points with Z_{12} and Z_2 (y, X_1, X_2, Z_{12}, Z_2)

5.3 The results

Table 1 reports the MSE of the coefficients for the Hausman-Taylor (HT) estimator and its robust counterpart (robust HT) based on 1000 replications[15].

[13] There are several weighting functions defined in the robust statistics literature. Since the Tukey biweight function is widely used, we only test the properties of the robust Hausman-Taylor estimator with this weight function.

[14] Bramati and Croux (2007) used a unique breakdown point of 20% with a tuning constant $c = 3.420$ for their Monte-Carlo study on samples of $N = 100$, $T = 4, 20$ and 5% or 10% contamination.

[15] Following Kapoor, Kelejian and Prucha (2007), our measure of dispersion is closely related to the standard measure of the MSE, but it is based on quantiles rather than moments because, unlike moments, quantiles are assured to exist. For ease of presentation,

The results in Table 1 pertain to $N = 100$, $T = 5$, with no outliers as well as 8 different cases of contamination, where the level of contamination is 5% or 10%.
When no outliers are present, the robust HT shows loss in MSE relative to the standard HT estimator. The absolute magnitudes are small (except for η_{12}, the coefficient of Z_{12}), but the relative MSE of robust HT with respect to classical HT could be as small as 1 and as big as 2, depending on the coefficient. Contrasting that to the various types of 5% and 10% contaminations considered, it is clear that the gain in absolute as well as relative MSE is huge for the robust HT estimator compared to the classical HT estimator. Note also that the largest absolute magnitude of this MSE is for η_{12} (the coefficient of Z_{12}, which is the exogenous time-invariant dummy variable). This is 0.12 for the HT estimator compared to 0.20 for our robust HT estimator in case of no outliers. However, when we introduce 5% contamination and vertical outliers, the MSE of HT rises to 0.62 compared to 0.20 for the robust HT estimator. In case of bad leverage points, the MSE of HT rises to 0.48 compared to 0.20 for the robust HT estimator. But when you add bad leverage points in Z_{12}, the MSE of HT becomes really bad 31.7 compared to 0.39 for the robust HT estimator. This is true for contamination cases 5,6,7 and 8 with bad leverage points and concentrated leverage points. The gains in absolute and relative MSE of robust HT over HT can be huge. For example, in the presence of vertical outliers, the robust HT estimator with 5% contamination, yields large gains in MSE with respect to the classical HT procedure. The HT MSE is 8 to 9 times higher than its robust counterpart for the coefficient estimates of X_1, X_2; 23 times higher for the intercept (Z_{11}) and 3 to 5 times higher for the coefficient estimates of Z_{12} and Z_2. Similarly, for bad leverage points, these MSE are respectively 74 to 107 times higher for the coefficient estimates of X_1, X_2; 18 times higher for the intercept (Z_{11}) and 2 to 24 times higher for the coefficient estimates of Z_{12} and Z_2.
When the outliers are block-concentrated vertical outliers, we get similar re-

we also refer to our measure as MSE. It is defined by:

$$MSE = bias^2 + \left[\frac{IQ}{1.35}\right]^2$$

where $bias$ is the difference between the median and the true value and IQ is the interquantile range $Q_3 - Q_1$ where Q_3 is the 0.75 quantile and Q_1 is the 0.25 quantile. If the distribution is normal, the median is the mean and, aside from a slight rounding error, $IQ/1.35$ is the standard deviation.

sults but when the outliers are block-concentrated leverage points, the gain in MSE of the robust HT estimate becomes more pronounced. Whatever the sampling scheme, the MSE of η_{12} — the parameter of the dummy variable Z_{12} — is always more affected than that of the other parameters. Of course, the robust version yields better results than HT no matter what type of contamination.

When we increase the level of contamination from 5% to 10%, the classical HT estimator yields much larger MSE and the gains from relative MSE using the robust HT procedure are much larger than the 5% contamination case no matter what sampling scheme is used.

When we increase the size of N and T, we get similar conclusions. Table 2 keeps N fixed at 100, but double T from 5 to 10, while Table 3 keeps T fixed at 5 and doubles N from 100 to 200. Table 4 doubles both N and T from $(100, 5)$ to $(200, 10)$. While the magnitudes of the MSE are different the gains in MSE as we go from HT to robust HT are maintained throughout the experiments and for various types of contamination. Once again the largest values for the MSE are reported for η_{12} (the coefficient of Z_{12}).

These results may be conditional on the fact that we only have 10% contamination. What happens if we increase the percentage of corrupted data? In order to investigate this, we used the largest allowable values of the breakdown points (i.e, 50% and $c = 1.547$) for each estimator (WMS and 2SGMS). We ran simulations for $N = 100$, $T = 5$, for case 2 (leverage points (y, X_1, X_2)) and for $5\%, 10\%, 15\%, 20\%, 25\%, 30\%, 35\%$ and 40% contamination.[16] Results in Table 5 show that the robust HT estimator resists quite well the increase in the percentage of contamination up to 35%. When the level of contamination is even higher, the gain in relative MSE decreases quickly even if the robust HT estimator is a little bit better than the classical HT estimator. However, when 40% of the data are corrupted, the MSE for the time invariant variable $\boldsymbol{Z_2}$ converges and even exceeds that of the MSE of HT. Figures 1 to 4 show average HT and robust HT estimates with their 95% confidence intervals for the coefficients of X_{11}, X_2, Z_{12} and Z_2 respectively. For the time varying variables (X_{11} and X_2), Figures 1 and 2 show that the robust HT estimator is stable with narrow confidence intervals showing a small bias and a good precision of the estimators leading to a relatively small MSE. For time invariant variables (Z_{12} and Z_2), Figures 3 and 4 show good stability of the robust HT estimator associated with narrow confidence

[16] We thank a referee for this suggestion.

intervals up to 25% of contamination. When the percentage of data pollution is higher, the confidence interval for the dummy variable Z_{12} widens appreciably. For the time invariant variable Z_2, both bias and confidence interval increase remaining always lower than the corresponding magnitudes for the standard HT estimator.

In order to evaluate the potential impact of the breakdown point values on the final 2SGMS estimator, we run simulations for $N = 100$, $T = 5$ for case 2 (leverage points $(y,\ X_1,\ X_2)$) for 10% and 25% contamination but with several breakdown point (bdp) values. We define $bdp_{WMS,S}$ ($bdp_{WMS,M}$) as the breakdown point for the S-estimator (M-estimator) of the WMS. We also define $bdp_{2SGMS\_1\_j,S}$ ($bdp_{2SGMS\_1\_j,M}$) as the breakdown point for the first stage of step j ($j = 1, 2$) for the S-estimator (M-estimator) of 2SGMS; and $bdp_{2SGMS\_2\_j,S}$ ($bdp_{2SGMS\_2\_j,M}$) as the breakdown point for the second stage of step j ($j = 1, 2$) for the S-estimator (M-estimator) of 2SGMS. We studied the following 5 cases to check the sensitivity of our results to different breakdown point values:[17]

	case A	case B	case C	case D	case E
$bdp_{WMS,S}$	0.50	0.25	0.50	0.50	0.50
$bdp_{WMS,M}$	0.25	0.25	0.25	0.25	0.25
$bdp_{2SGMS\_1\_j,S}$	0.25	0.50	0.50	0.25	0.50
$bdp_{2SGMS\_1\_j,M}$	0.25	0.25	0.25	0.25	0.25
$bdp_{2SGMS\_2\_j,S}$	0.25	0.50	0.25	0.50	0.50
$bdp_{2SGMS\_2\_j,M}$	0.25	0.25	0.25	0.25	0.25

In case A, we used different values of the breakdown points only for WMS. As suggested by Rousseeuw and Yohai (1987), we set the tuning constant to $c = 1.547$ for the S-estimator to guarantee a 50% breakdown point and we set the tuning constant to $c = 2.937$ for the second step M-estimator to guarantee a higher efficiency of 76% for the final estimator. Results in Table 6 for 10% contamination are similar to those of Table 1 for which all the breakdown points are 25%. When only 10% of the data are corrupted, there is no significant differences between cases A to E. But, when we increase the percentage of data pollution up to 25%, the results deteriorate for cases A and C. In these two cases, the breakdown points values of the second stage of step j ($j = 1, 2$) for the S-estimator of 2SGMS $bdp_{2SGMS\_2\_j,S}$ are small (0.25 as compared to 0.5 in the three other cases). Cases B, D and E give

[17] We thank a referee for this suggestion.

similar results showing that the crucial values are those of $bdp_{2SGMS\_2\_j,S}$ and not necessarily those of bdp_{WMS} or $bdp_{2SGMS\_1\_j}$.

What about the interesting case where outliers only exist in the time invariant variables (for instance in Z_2)?[18] To check this potential negative influence, we run simulations for $N = 100$, $T = 5$, and for 20% contamination for leverage points. First, we suppose that only Y_2 and Z_2 are randomly contaminated or block-contaminated and, second, we suppose that only Z_2 is randomly contaminated or block-contaminated. In Table 7, results for leverage points for both Y_2 and Z_2 show that the robust HT estimator yields better results than HT no matter what type of contamination. In contrast, when we simulate contamination only on Z_2, the impact of this contamination appears to be marginal.

Last, we tried an hybrid setup where we have a quasi-robust HT estimator where only one of the two robust estimators is deployed.[19] Two cases are possible: either the robust within estimator is followed by the generic IV regression, or either the classic within estimator is followed by the two stage generalized MS estimator. We only run these two quasi-robust HT estimators for $N = 100$, $T = 5$, for case 2 (leverage points (y, X_1, X_2)) and for 10% contamination. The results in Table 8 show that the second quasi-robust estimator (labeled quasi-robust HT (Within, 2SGMS)) gives similar results as compared to the robust HT estimator whatever the type of contamination (vertical outliers, leverage points, random or block-contamination). In contrast, the first quasi-robust estimator (labelled quasi-robust HT (WMS, HT)) does not seem to clean effectively the negative effects of contamination as there seems to be no significant gain in absolute or relative MSE as compared to the standard HT estimate. This gives further evidence that a robust version of the second step of the Hausman-Taylor estimator is necessary and highly recommended.

6 An empirical example: the Cornwell-Rupert (1988) Mincer wage equation

Cornwell and Rupert (1988) applied the Hausman-Taylor estimator to a returns to schooling example based on a panel of 595 individuals observed

[18] We thank a referee for this suggestion.

[19] We thank a referee for this suggestion.

over the period 1976-1982 drawn from the Panel Study of Income Dynamics (PSID). In particular, log wage is regressed on years of education (ED), weeks worked (WEEKS), years of full-time work experience (EXP) and its square (EXP2), occupation (OCC=1, if the individual is in a blue-collar occupation), residence (SOUTH = 1, SMSA = 1, if the individual resides in the South, or in a standard metropolitan statistical area), industry (IND=1, if the individual works in a manufacturing industry), marital status (MAR=1, if the individual is married), sex and race (FEM=1, BLK=1, if the individual is female or black), union coverage (UNION = 1, if the individual's wage is set by a union contract). Table 9 reports the Within, WMS, HT and robust HT estimators for this Mincer wage equation in which the X_1 variables are (OCC, SOUTH, SMSA, IND), the X_2 variables are (EXP, EXP2, WKS, MAR, UNION), the Z_1 variables are (FEM, BLK) and the Z_2 variable is (ED). In this specification, there are 8 dummies for 12 explanatory variables. As Croux and Verardi (2008) note, the need of random picking p subsets of K observations become Achille's heel of the MS algorithm when several dummy variables are present. Subsampling algorithms can easily lead to collinear sub-samples if various dummies are present. A rough solution is to use subsets of size a little bit larger than K. For this empirical example, we generated 500 random subsets of $(T.K)$ observations.[20] If we compare the Within and WMS estimators, we see that all the coefficients of WMS are statistically significant contrary to those of the Within estimator. MAR switches sign and becomes significant. Weeks worked has a larger effect that is now statistically significant. The HT estimator indicates that an additional year of schooling yields a 13.8% wage gain, and that gender discrimination (FEM) is statistically insignificant. An additional year of schooling yields almost the same effect as for HT, 13.5%, but with lower standard errors. The FEM coefficient estimate yields a lower but statistically significant effect on wages. Interestingly, SOUTH is insignificant for both HT and robust HT, while MAR is significant only for the robust HT estimator. In fact, with the robust HT estimator, industry and weeks worked are also statistically significant. Only a small proportion of the observations (2.5% and 8.4%) change values for SOUTH and MAR over the period observed. This indicates that these variables would have been eliminated using the within transformation

[20] One can also use the exact solution proposed by Maronna and Yohai (2006) who introduce the MS-estimator that alternates an S-estimator (for continuous variables) and an M-estimator (for dummy ones), till convergence.

if this proportion of observations changing values was zero. To check the sensitivity of our results, we removed SOUTH and MAR, and the results are reported at the bottom of Table 9. For the robust HT estimator deleting these two dummy variables, the returns to education is about the same and the FEM coefficient is smaller but statistically significant.

7 Conclusion

This paper applies the useful robust panel data methods suggested by Bramati and Croux (2007) and Wagenvoort and Waldmann (2001) to the Hausman and Taylor (1981) estimator. We demonstrate using Monte Carlo experiments the substantial gains in efficiency as measured by MSE of this robust HT estimator over its classical counterpart. The magnitude of the gains in MSE depend upon the *type* and *degree* of contamination of the observations. We illustrate this robust HT method by applying it to the classical Mincer wage equation using the empirical study of Cornwell and Rupert (1988). For this empirical study, the returns to education seem to be robust to outliers, while the magnitude and significance of the female coefficient is sensitive to robustification of the HT estimator. We performed several sensitivity analysis but there remains a lot of questions for future research. For example, we did not prove that the proposed robust Hausman-Taylor estimator is scale, regression and affine equivariant. There is also a need to derive formal tests or metrics to use in applied panel data setting to determine the presence of outliers. This analysis can also be extended to dynamic HT type models, where one can check the sensitivity of using the difference transformation, rather than a within transformation that subtracts a mean or a median to get rid of the individual effects, on the performance of the contaminated classical dynamic panel data estimators.

Appendix 1

Definition of the breakdown point

As Bramati and Croux (2007, pp.523) noted, the breakdown point of an estimator is defined as the smallest fraction of outlying observations that can cause a 'breakdown' of the estimator. Let our panel data sample be composed of NT observations $\Omega = \{y_{it}, X_{it}\}$ and let $\theta(\Omega)$ be our estimator. Let $\widetilde{\Omega} = \left\{\widetilde{y}_{it}, \widetilde{X}_{it}\right\}$ be a contaminated set of NT observations, any m of the original points of Ω are replaced by arbitrary values, and $\widetilde{\theta}\left(\widetilde{\Omega}\right)$ is our estimator of the corrupted sample $\widetilde{\Omega}$. If $\omega(m, \theta, \Omega)$ is the supremum of $\| \widetilde{\theta}\left(\widetilde{\Omega}\right) - \theta(\Omega) \|$, then the breakpoint of θ at Ω is defined as:

$$\begin{aligned}
\varepsilon^*_{NT}(m, \theta, \Omega) &= \min\left\{\frac{m}{NT}; \omega(m, \theta, \Omega) \text{ is infinite}\right\} \\
&= \min\left\{\frac{m}{NT}; \sup_{\widetilde{\Omega}} \| \widetilde{\theta}\left(\widetilde{\Omega}\right) - \theta(\Omega) \| = \infty\right\}
\end{aligned}$$

The breakdown point of the estimator θ at Ω is the smallest proportion of observations replaced by outliers which can cause the estimator θ to take on values arbitrarily far from $\theta(\Omega)$. (see Bramati and Croux (2007) and Croux and Verardi (2008)). In case of block-contaminated data, we suppose that for some individuals half of the time, the data is contaminated. In particular, $\widetilde{\Omega} = \left\{\widetilde{y}_{it}, \widetilde{X}_{it}\right\}$ where for some individuals, $\widetilde{y}_{it}, \widetilde{X}_{it}$ are contaminated for $t = 1, 2, ..., T/2$ and not for $t = (T/2)+1, ..., T)$.[21] For instance, if $N = 100$, $T = 5$ and 10% of the observations are corrupted, it means that 10 individuals each have 5 time observations which are corrupted.

Appendix 2

Breakdown point and asymptotic efficiency.

Rousseeuw and Yohai (1987) computed the asymptotic efficiency of the S-estimator of a Gaussian model for different values of the breakdown point for the Tukey biweight function:

[21] The $T/2$ corrupted time data can be randomly chosen between $t = 1$ and $t = T$.

breakdown point	asymp. efficiency	tuning constant	
ε^*	as.eff	c	b
50%	28.7%	1.547	0.1995
45%	37.0%	1.756	0.2312
40%	46.2%	1.988	0.2634
35%	56.0%	2.251	0.2957
30%	66.1%	2.560	0.3278
25%	75.9%	2.937	0.3593
20%	84.7%	3.420	0.3899
15%	91.7%	4.096	0.4194
10%	96.6%	5.182	0.4475

source: Rousseeuw and Yohai (1987) table 3 pp. 268.

where $as.eff = \left(\int \psi' d\Phi\right)^2 / \left(\int \psi^2 d\Phi\right)$ where Φ is the c.d.f of $N(0,1)$ and $b = E\left[\rho\left(Q\right)\right]$ with $Q \sim N(0,1)$.

References

Aquaro, M. and P. Čížek, 2010, One-step robust estimation of fixed-effects panel data models, Working paper CentER n° 2010-110, Dept. of Econometrics, *Tilburg University*, The Netherlands.

Baltagi, B.H., 2008, *Econometric Analysis of Panel Data*, 4th edition, John Wiley & sons, Chichester.

Bramati, M. C. and C. Croux, 2007, Robust estimators for the fixed effects panel data model, *Econometrics Journal* 10, 521-540.

Breusch, T.S., G.E. Mizon and P. Schmidt, 1989, Efficient estimation using panel data, *Econometrica,* 57, 695-700.

Čížek, P., 2008, Robust and efficient adaptive estimation of binary-choice regression models, *Journal of the American Statistical Association* 103, 687-696.

Cornwell, C. and P. Rupert, 1988, Efficient estimation with panel data: an empirical comparison of instrumental variables estimators, *Journal of Applied Econometrics* 3, 149-155.

Croux, C. and V. Verardi, 2008, Robust regression in Stata, *The Stata Journal* 9, 439-453.

Dehon, C., Gassner, M. and V. Verardi, 2009, Beware of 'good' outliers and overoptimistic conclusions, *Oxford Bulletin of Economics and Statistics* 71, 437-452.

Dhaene, G. and Y. Zhu, 2009, Median-based estimation of dynamic panel models with fixed effects, Working paper, Faculty of Business and Economics, *Catholic University of Leuven*, Belgium.

Donoho D. L. and P. J. Huber, 1983, The notion of breakdown point, in *A Festschrift for Erich L. Lehmann*, (Bickel, P.J., Doksum, K.A. and J. L. Hodges, Eds.), Belmont, Wadsworth, 157–184.

Egger, P. and M. Pfaffermayr, 2004, Distance, trade and FDI: a SUR Hausman-Taylor approach, *Journal of Applied Econometrics,* 19, 227-246.

Hausman, J.A. and W.E. Taylor, 1981, Panel data and unobservable individual effects, *Econometrica* 49, 1377–1398.

Hinloopen, J. and J.L.M. Wagenvoort, 1997, On the computation and efficiency of a HBP-GM estimator: some simulation results, *Computational Statistics and Data Analysis* 25, 1-15.

Huber, P., 1964, Robust estimation of a location parameter, *Annals of Mathematical Statistics* 35, 73-101.

Huber, P.J., 1981, *Robust Statistics*, Series in Probability and Mathematical Statistics, 1st Edition, John Wiley, New York

Huber, P.J. and E.M. Ronchetti, 2009, *Robust Statistics*, Series in Probability and Mathematical Statistics, 2nd Edition, John Wiley, New York.

Kapoor, M., H.H. Kelejian and I.R. Prucha, 2007, Panel data models with spatially correlated error components, *Journal of Econometrics* 140, 97-130.

Lucas, A., van Dijk, R. and T. Kloek, 2007, Outlier robust GMM estimation of leverage determinants in linear dynamic panel data models, Unpublished manuscript, *Vrije Universiteit*, Amsterdam, the Netherlands.

Maronna, R. A., R. D. Martin, and V. J. Yohai, 2006, *Robust Statistics*, John Wiley & Sons, New York.

Maronna, R., and V. Yohai, 2006, Robust regression with both continuous and categorical predictors, *Journal of Statistical Planning and Inference,* 89, 197-214.

Olive, D.J., 2008, *Applied Robust Statistics*, Department of Mathematics, Southern Illinois University, Unpublished manuscript.

Ronchetti, E. and F. Trojani, 2001, Robust inference with GMM estimators. *Journal of Econometrics* 101, 37-69.

Rousseeuw, P. J., 1984, Least median of squares regression, *Journal of the American Statistical Association* 79, 87-780.

Rousseeuw, P. J. and A. M. Leroy, 2003, *Robust Regression and Outlier Detection*, John Wiley, New York.

Rousseeuw, P.J. and K. Van Driessen, 1999, A fast algorithm for the minimum covariance determinant estimator, *Technometrics*, 41, 212-223.

Rousseeuw, P. J. and V. Yohai, 1987, Robust Regression by means of S-estimators, in *Robust and Nonlinear Time Series Analysis*, (J. Franke, W. Härdle and D. Martin, Eds.), Springer Verlag, Berlin, 256-272.

Salibian-Barrera, M. and V. Yohai, 2006, A fast algorithm for S-regression estimates, *Journal of Computational and Graphical Statistics* 15, 414-427.

Serlenga, L. and Y. Shin, 2007, Gravity models of intra-EU trade: application of the CCEP-HT estimation in heterogeneous panels with unobserved common time-specific factors, *Journal of Applied Econometrics,* 22, 361–381.

Wagenvoort R. and R. Waldmann, 2002, On B-robust instrumental variable estimation of the linear model with panel data, *Journal of Econometrics* 106, 297-324.

Yohai, V. 1987, High breakdown-point and high efficiency estimates for regression, *The Annals of Statistics* 15, 642-665.

Table 1 - MSE of coefficients N =100, T = 5, 5% contamination (1000 replications)

		β_{11}	β_{12}	β_{2}	η_{11}	η_{12}	η_{2}
no outlier	HT	0.0023	0.0026	0.0028	0.0010	0.1235	0.0102
	robust HT	0.0042	0.0049	0.0046	0.0014	0.2010	0.0164
				Contamination (5%)			
Vertical outliers	HT	0.0453	0.0446	0.0554	0.0663	0.6205	0.0882
	robust HT	0.0054	0.0058	0.0056	0.0028	0.1981	0.0181
Leverage points	HT	0.6299	0.6357	0.7643	0.0420	0.4827	0.5856
	robust HT	0.0084	0.0064	0.0071	0.0023	0.2027	0.0239
Concentrated vertical outliers	HT	0.0446	0.0435	0.0542	0.0581	0.6984	0.0942
	robust HT	0.0067	0.0055	0.0058	0.0044	0.2198	0.0217
Concentrated leverage points	HT	0.5142	0.4460	0.6145	0.0383	0.5567	0.3740
	robust HT	0.0087	0.0098	0.0094	0.0041	0.2411	0.0338
Leverage points with Z_{12}	HT	0.3464	0.3683	0.3964	0.0227	31.7112	0.3983
	robust HT	0.0080	0.0087	0.0078	0.0026	0.3851	0.0239
Concentrated leverage points with Z_{12}	HT	0.2693	0.2994	0.3147	0.0232	33.5696	0.2638
	robust HT	0.0089	0.0078	0.0090	0.0040	0.4215	0.0324
Leverage points with Z_{12} and Z_{2}	HT	0.3356	0.3710	0.3387	0.0205	29.8074	0.3845
	robust HT	0.0089	0.0085	0.0077	0.0025	0.4621	0.0236
Concentrated leverage points with Z_{12} and Z_{2}	HT	0.2782	0.2821	0.3120	0.0211	31.3235	0.2694
	robust HT	0.0084	0.0076	0.0087	0.0029	0.4787	0.0296
				Contamination (10%)			
Vertical outliers	HT	0.0718	0.0770	0.1111	0.2481	0.9964	0.1379
	robust HT	0.0067	0.0076	0.0056	0.0103	0.2523	0.0197
Leverage points	HT	1.6804	1.6431	1.9832	0.1348	0.7264	1.5328
	robust HT	0.0120	0.0112	0.0117	0.0087	0.2242	0.0481
Concentrated vertical outliers	HT	0.0707	0.0682	0.0914	0.2116	1.5824	0.1503
	robust HT	0.0073	0.0064	0.0051	0.0225	0.4390	0.0327
Concentrated leverage points	HT	1.2417	1.2015	1.3094	0.1189	1.0429	0.7424
	robust HT	0.0204	0.0211	0.0209	0.0194	0.4306	0.1102
Leverage points with Z_{12}	HT	0.7343	0.7314	0.7407	0.0503	64.7983	0.8470
	robust HT	0.0231	0.0225	0.0197	0.0095	1.1644	0.0609
Concentrated leverage points with Z_{12}	HT	0.5671	0.6129	0.6127	0.0620	70.5724	0.5393
	robust HT	0.0244	0.0261	0.0276	0.0198	0.8332	0.1154
Leverage points with Z_{12} and Z_{2}	HT	0.7279	0.8791	0.8224	0.0535	63.8211	0.8714
	robust HT	0.0208	0.0192	0.0168	0.0087	0.9404	0.0486
Concentrated leverage points with Z_{12} and Z_{2}	HT	0.6235	0.5389	0.5822	0.0567	67.5897	0.4981
	robust HT	0.0195	0.0218	0.0221	0.0189	0.8958	0.1032

Table 2 - MSE of coefficients N =100, T = 10, 5% and 10% contamination (1000 replications)

		β_{11}	β_{12}	β_{2}	η_{11}	η_{12}	η_{2}
no outlier	HT	0.0012	0.0011	0.0013	0.0009	0.1065	0.0081
	robust HT	0.0019	0.0018	0.0023	0.0012	0.1921	0.0107
			Contamination (5%)				
Vertical outliers	HT	0.0189	0.0219	0.0278	0.0662	0.3180	0.0382
	robust HT	0.0030	0.0033	0.0028	0.0063	0.1967	0.0140
Leverage points	HT	0.5893	0.5989	0.7037	0.0401	0.3378	0.5309
	robust HT	0.0043	0.0040	0.0037	0.0064	0.1717	0.0144
Concentrated vertical outliers	HT	0.0183	0.0184	0.0216	0.0563	0.6558	0.0566
	robust HT	0.0028	0.0027	0.0023	0.0141	0.3836	0.0270
Concentrated leverage points	HT	0.3703	0.3607	0.4397	0.0368	0.4525	0.2016
	robust HT	0.0053	0.0052	0.0058	0.0123	0.3808	0.0468
Leverage points with Z_{12}	HT	0.3273	0.3369	0.3962	0.0393	43.6259	0.3550
	robust HT	0.0040	0.0041	0.0032	0.0056	0.2905	0.0159
Concentrated leverage points with Z_{12}	HT	0.2457	0.2362	0.2591	0.0608	61.7257	0.2023
	robust HT	0.0057	0.0054	0.0058	0.0126	0.5750	0.0481
Leverage points with Z_{12} and Z_2	HT	0.3564	0.3147	0.3327	0.0387	43.4782	0.3588
	robust HT	0.0035	0.0045	0.0033	0.0050	0.4028	0.0146
Concentrated leverage points with Z_{12} and Z_2	HT	0.2185	0.2282	0.2452	0.0529	54.5255	0.1640
	robust HT	0.0047	0.0049	0.0049	0.0112	0.5467	0.0389
			Contamination (10%)				
Vertical outliers	HT	0.0375	0.0365	0.0486	0.2450	0.4764	0.0641
	robust HT	0.0033	0.0036	0.0030	0.0380	0.2535	0.0158
Leverage points	HT	1.6572	1.7292	1.8854	0.1332	0.5636	1.4878
	robust HT	0.0064	0.0067	0.0064	0.0304	0.2079	0.0272
Concentrated vertical outliers	HT	0.0255	0.0256	0.0255	0.1703	1.8336	0.1151
	robust HT	0.0025	0.0028	0.0019	0.0244	0.7454	0.0480
Concentrated leverage points	HT	0.6623	0.7548	0.7656	0.1007	1.0182	0.1882
	robust HT	0.0092	0.0103	0.0112	0.0282	0.7329	0.1434
Leverage points with Z_{12}	HT	0.7021	0.7079	0.8027	0.0958	86.4403	0.7952
	robust HT	0.0091	0.0080	0.0079	0.0326	0.6247	0.0364
Concentrated leverage points with Z_{12}	HT	0.3648	0.3919	0.3797	0.1376	118.2726	0.2568
	robust HT	0.0106	0.0124	0.0124	0.0278	1.2446	0.1568
Leverage points with Z_{12} and Z_2	HT	0.7350	0.7159	0.7651	0.1029	88.9891	0.7858
	robust HT	0.0077	0.0080	0.0065	0.0292	0.7635	0.0322
Concentrated leverage points with Z_{12} and Z_2	HT	0.3989	0.3785	0.3601	0.1410	112.4906	0.2629
	robust HT	0.0111	0.0107	0.0109	0.0275	1.1932	0.1431

Table 3 - MSE of coefficients N =200, T = 5, 5% and 10% contamination (1000 replications)

		β_{11}	β_{12}	β_{2}	η_{11}	η_{12}	η_{2}
no outlier	HT	0.0013	0.0013	0.0014	0.0004	0.0514	0.0048
	robust HT	0.0025	0.0020	0.0026	0.0007	0.0833	0.0073
			Contamination (5%)				
Vertical outliers	HT	0.0250	0.0221	0.0298	0.0649	0.2665	0.0422
	robust HT	0.0032	0.0033	0.0030	0.0023	0.0959	0.0091
Leverage points	HT	0.5862	0.5724	0.7283	0.0404	0.2227	0.5624
	robust HT	0.0043	0.0045	0.0046	0.0019	0.1042	0.0145
Concentrated vertical outliers	HT	0.0204	0.0188	0.0272	0.0595	0.3678	0.0468
	robust HT	0.0032	0.0028	0.0029	0.0043	0.1171	0.0092
Concentrated leverage points	HT	0.4848	0.5399	0.5920	0.0388	0.2185	0.3857
	robust HT	0.0060	0.0061	0.0062	0.0039	0.1107	0.0239
Leverage points with Z_{12}	HT	0.3718	0.3591	0.3730	0.0194	30.7602	0.3762
	robust HT	0.0044	0.0040	0.0040	0.0017	0.1660	0.0125
Concentrated leverage points with Z_{12}	HT	0.2909	0.2907	0.3192	0.0273	36.0306	0.2600
	robust HT	0.0060	0.0061	0.0059	0.0032	0.2196	0.0239
Leverage points with Z_{12} and Z_{2}	HT	0.3208	0.3391	0.3504	0.0190	31.4212	0.3545
	robust HT	0.0044	0.0046	0.0038	0.0013	0.2090	0.0124
Concentrated leverage points with Z_{12} and Z_{2}	HT	0.2935	0.2786	0.3013	0.0225	32.2625	0.2621
	robust HT	0.0048	0.0043	0.0042	0.0027	0.2308	0.0169
			Contamination (10%)				
Vertical outliers	HT	0.0363	0.0371	0.0483	0.2444	0.4273	0.0634
	robust HT	0.0035	0.0028	0.0029	0.0095	0.1089	0.0088
Leverage points	HT	1.6679	1.6053	1.8884	0.1283	0.3845	1.5108
	robust HT	0.0088	0.0086	0.0084	0.0077	0.1066	0.0297
Concentrated vertical outliers	HT	0.0389	0.0377	0.0456	0.2128	0.7118	0.0750
	robust HT	0.0032	0.0035	0.0030	0.0237	0.2230	0.0161
Concentrated leverage points	HT	1.2614	1.2761	1.3243	0.1183	0.5092	0.7000
	robust HT	0.0182	0.0179	0.0189	0.0188	0.2231	0.1007
Leverage points with Z_{12}	HT	0.7173	0.8023	0.8121	0.0490	63.7971	0.8610
	robust HT	0.0137	0.0136	0.0131	0.0082	0.4565	0.0410
Concentrated leverage points with Z_{12}	HT	0.5688	0.6160	0.6019	0.0583	69.9450	0.4638
	robust HT	0.0196	0.0198	0.0207	0.0189	0.4322	0.1037
Leverage points with Z_{12} and Z_{2}	HT	0.7576	0.7869	0.7428	0.0442	61.7973	0.8441
	robust HT	0.0101	0.0102	0.0099	0.0071	0.5515	0.0312
Concentrated leverage points with Z_{12} and Z_{2}	HT	0.5793	0.5650	0.6015	0.0587	67.0209	0.4536
	robust HT	0.0151	0.0165	0.0166	0.0162	0.3982	0.0962

Table 4 - MSE of coefficients N =200, T = 10, 5% contamination (1000 replications)

		β_{11}	β_{12}	β_2	η_{11}	η_{12}	η_2
no outlier	HT	0.0006	0.0006	0.0007	0.0004	0.0589	0.0042
	robust HT	0.0009	0.0008	0.0010	0.0006	0.0820	0.0056
Contamination (5%)							
Vertical outliers	HT	0.0103	0.0102	0.0145	0.0662	0.1621	0.0179
	robust HT	0.0016	0.0014	0.0014	0.0060	0.0864	0.0060
Leverage points	HT	0.6265	0.6294	0.6866	0.0387	0.1558	0.5890
	robust HT	0.0024	0.0023	0.0019	0.0056	0.0827	0.0089
Concentrated vertical outliers	HT	0.0088	0.0081	0.0094	0.0536	0.3951	0.0311
	robust HT	0.0014	0.0013	0.0009	0.0138	0.2244	0.0134
Concentrated leverage points	HT	0.3720	0.3862	0.4201	0.0349	0.2455	0.1966
	robust HT	0.0039	0.0050	0.0047	0.0117	0.2508	0.0363
Leverage points with Z_{12}	HT	0.3106	0.3373	0.3943	0.0341	43.9376	0.3435
	robust HT	0.0025	0.0023	0.0019	0.0051	0.1343	0.0089
Concentrated leverage points with Z_{12}	HT	0.2284	0.2518	0.2546	0.0563	60.5669	0.1669
	robust HT	0.0043	0.0047	0.0045	0.0117	0.3075	0.0348
Leverage points with Z_{12} and Z_2	HT	0.3290	0.3300	0.3460	0.0371	42.8320	0.3779
	robust HT	0.0024	0.0024	0.0016	0.0042	0.1872	0.0074
Concentrated leverage points with Z_{12} and Z_2	HT	0.2292	0.2334	0.2245	0.0622	59.7633	0.1554
	robust HT	0.0036	0.0032	0.0036	0.0096	0.3200	0.0348
Contamination (10%)							
Vertical outliers	HT	0.0198	0.0217	0.0234	0.2441	0.2536	0.0293
	robust HT	0.0019	0.0019	0.0013	0.0383	0.1061	0.0084
Leverage points	HT	1.7081	1.5612	1.9221	0.1305	0.2831	1.5793
	robust HT	0.0050	0.0046	0.0044	0.0301	0.1096	0.0223
Concentrated vertical outliers	HT	0.0125	0.0120	0.0163	0.1670	0.9008	0.0571
	robust HT	0.0012	0.0012	0.0010	0.0220	0.4063	0.0233
Concentrated leverage points	HT	0.7113	0.6753	0.7371	0.1019	0.5339	0.1489
	robust HT	0.0083	0.0077	0.0087	0.0264	0.3498	0.1061
Leverage points with Z_{12}	HT	0.7000	0.7468	0.7239	0.1062	93.9170	0.8229
	robust HT	0.0057	0.0059	0.0057	0.0309	0.2828	0.0245
Concentrated leverage points with Z_{12}	HT	0.3691	0.3613	0.3495	0.1540	120.1300	0.2022
	robust HT	0.0091	0.0094	0.0104	0.0250	0.6855	0.1190
Leverage points with Z_{12} and Z_2	HT	0.6805	0.6756	0.7437	0.0955	88.5690	0.7853
	robust HT	0.0046	0.0043	0.0038	0.0279	0.3611	0.0206
Concentrated leverage points with Z_{12} and Z_2	HT	0.3536	0.3545	0.3616	0.1317	114.3700	0.2204
	robust HT	0.0086	0.0084	0.0091	0.0250	0.6182	0.1260

Table 5 - MSE of coefficients for leverage points, from 5% to 40% contamination, N =100, T = 5 (1000 replications)

Contamination		β_{11}	β_{12}	β_2	η_{11}	η_{12}	η_2
5%	HT	0.6299	0.6357	0.7643	0.0420	0.4827	0.5856
	robust HT	0.0084	0.0064	0.0071	0.0023	0.2027	0.0239
10%	HT	1.6804	1.6431	1.9832	0.1348	0.7264	1.5328
	robust HT	0.0120	0.0112	0.0117	0.0087	0.2242	0.0481
15%	HT	2.6447	2.8059	3.1403	0.2324	1.1057	2.4817
	robust HT	0.0202	0.0236	0.0222	0.0284	0.3318	0.0865
20%	HT	3.6674	3.7195	4.0780	0.3833	1.2714	3.4245
	robust HT	0.0364	0.0356	0.0306	0.0768	0.4462	0.1649
25%	HT	4.6089	4.4924	4.9189	0.5097	1.4811	3.9350
	robust HT	0.0757	0.0812	0.0880	0.1507	0.6847	0.4081
30%	HT	5.0355	5.0298	5.5646	0.6478	1.7890	4.7209
	robust HT	0.0980	0.0885	0.0917	0.2959	1.8805	0.5418
35%	HT	5.5962	5.5722	6.6318	0.8208	1.7701	4.7974
	robust HT	0.1924	0.2178	0.1897	0.5065	2.8786	0.9733
40%	HT	6.3150	6.1837	6.9899	0.9517	1.9593	5.5632
	robust HT	3.3007	3.2071	3.2774	0.9650	5.8274	3.2028

Figure 1. - Average HT and Robust HT estimates of β_{11}
with corresponding 95% confidence intervals

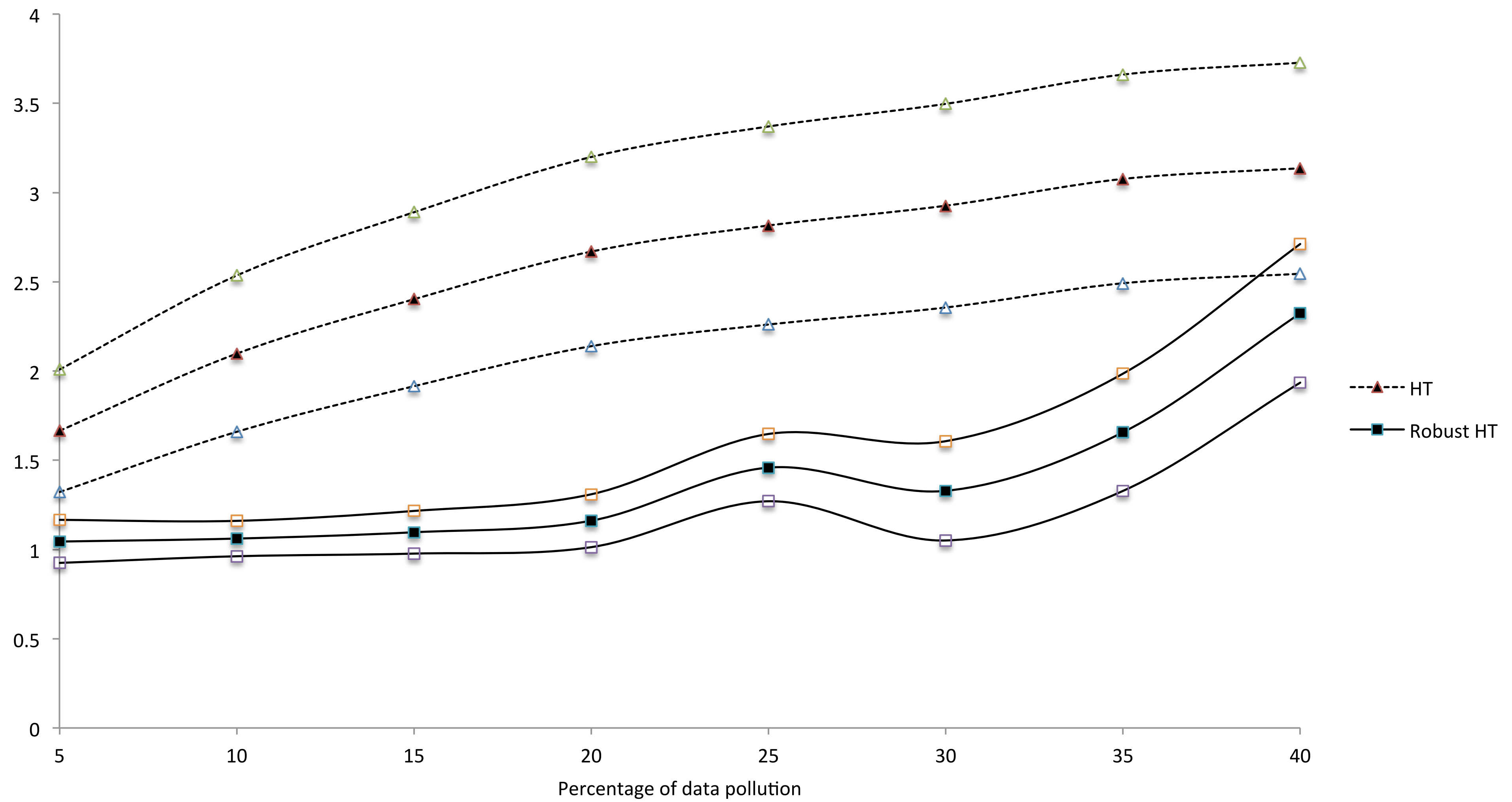

Figure 2. - Average HT and Robust HT estimates of β_2
with corresponding 95% confidence intervals

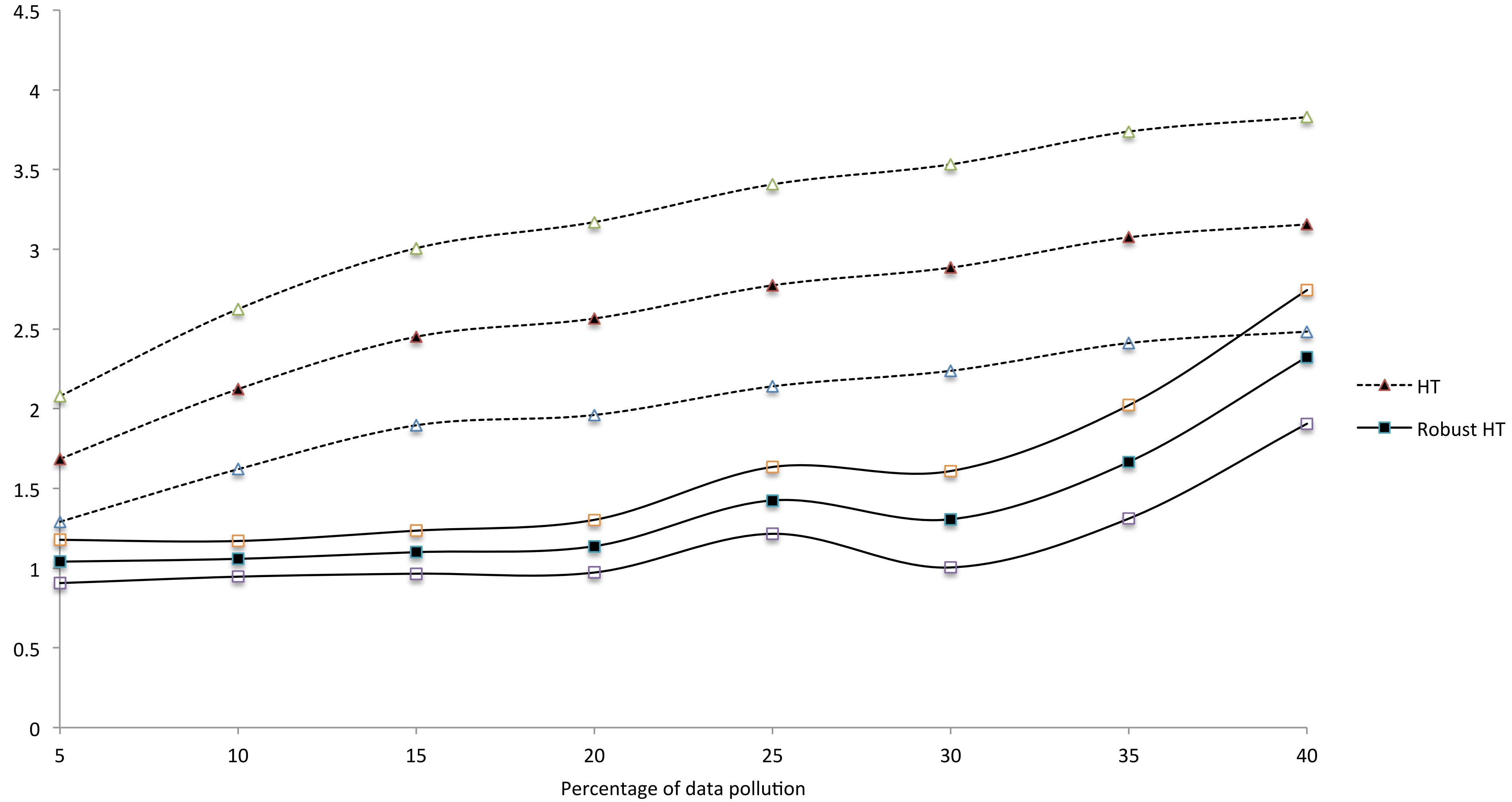

Figure 3. - Average HT and Robust HT estimates of η_{12} with corresponding 95% confidence intervals

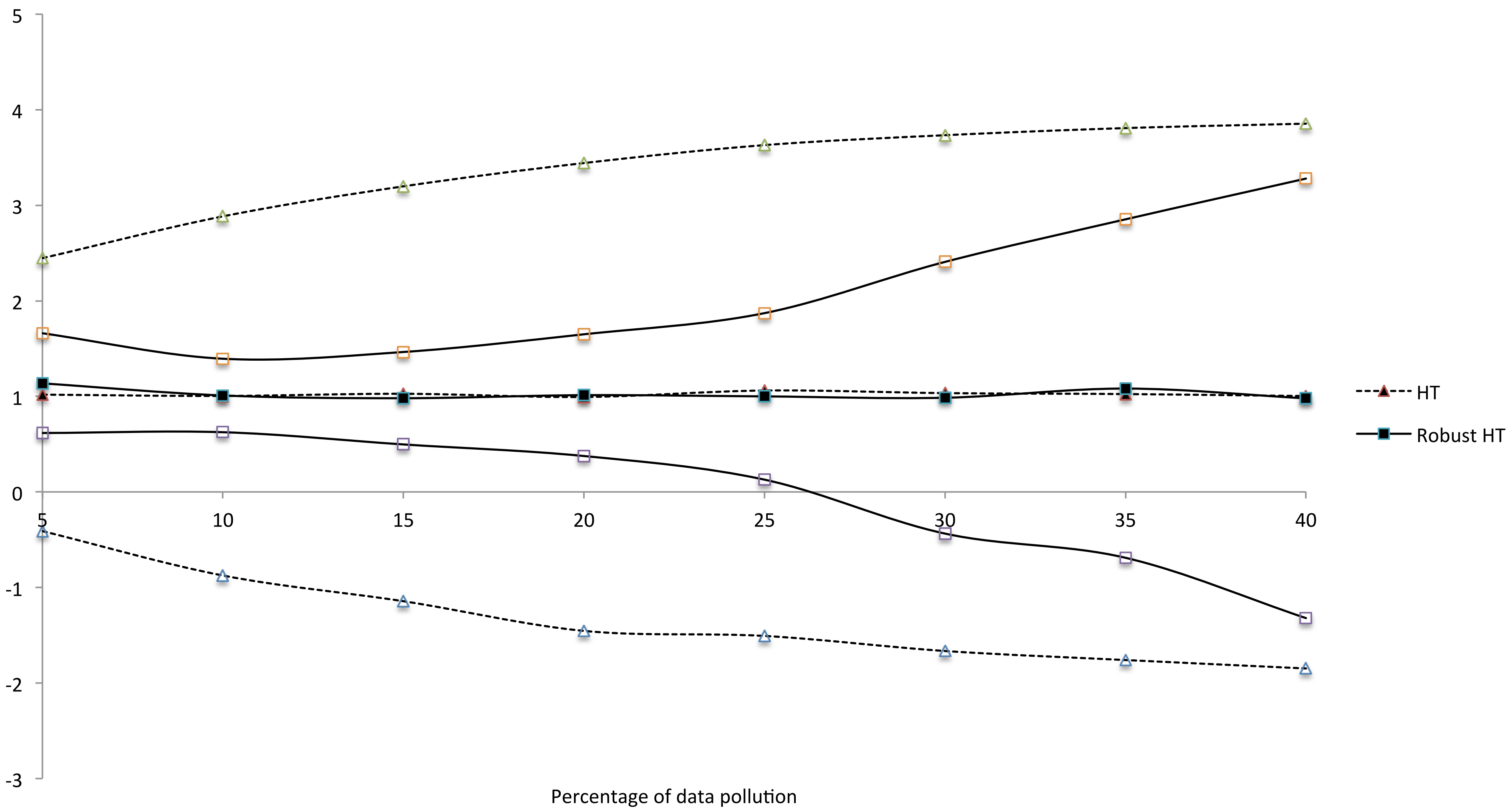

Figure 4. - Average HT and Robust HT estimates of η_2
with corresponding 95% confidence intervals

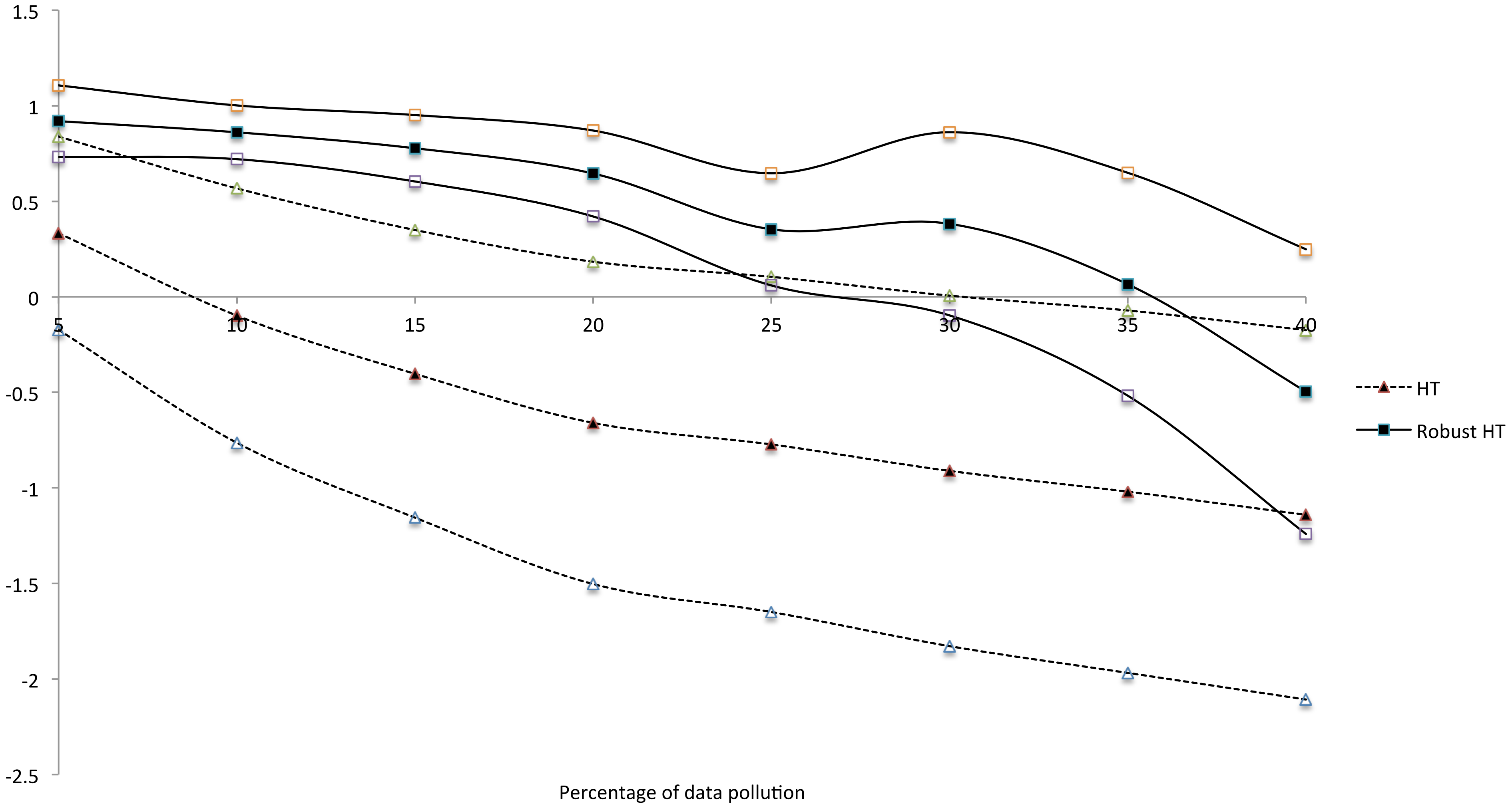

Table 6 - MSE of coefficients for leverage points N =100, T = 5 (1000 replications)

	β_{11}	β_{12}	β_2	η_{11}	η_{12}	η_2
Contamination (10%)						
HT	1.6804	1.6431	1.9832	0.1348	0.7264	1.5328
Robust HT - case A	0.0119	0.0130	0.0119	0.0094	0.2493	0.0496
Robust HT - case B	0.0132	0.0139	0.0140	0.0090	0.2885	0.0414
Robust HT - case C	0.0130	0.0138	0.0135	0.0096	0.2649	0.0395
Robust HT - case D	0.0144	0.0118	0.0134	0.0096	0.2969	0.0497
Robust HT - case E	0.0122	0.0137	0.0127	0.0102	0.2796	0.0493
Contamination (25%)						
HT	4.6089	4.4924	4.9189	0.5097	1.4811	3.9350
Robust HT - case A	2.0403	1.9726	1.9137	0.0776	1.2611	2.0880
Robust HT - case B	0.0841	0.0841	0.0879	0.1469	0.7055	0.4223
Robust HT - case C	2.1863	2.1296	2.1639	0.0862	1.4104	2.2172
Robust HT - case D	0.0837	0.0863	0.0883	0.1480	0.7856	0.4094
Robust HT - case E	0.0757	0.0812	0.0880	0.1507	0.6847	0.4081

case A: $bdp_{WMS,S} = 0.50$, $bdp_{WMS,M} = 0.25$, $bdp_{2SGMS\ 1\ j,S} = 0.25$, $bdp_{2SGMS\ 1\ j,M} = 0.25$, $bdp_{2SGMS\ 2\ j,S} = 0.25$, $bdp_{2SGMS\ 2\ j,M} = 0.25$, $j=1,2$

case B: $bdp_{WMS,S} = 0.25$, $bdp_{WMS,M} = 0.25$, $bdp_{2SGMS\ 1\ j,S} = 0.50$, $bdp_{2SGMS\ 1\ j,M} = 0.25$, $bdp_{2SGMS\ 2\ j,S} = 0.50$, $bdp_{2SGMS\ 2\ j,M} = 0.25$, $j=1,2$

case C: $bdp_{WMS,S} = 0.50$, $bdp_{WMS,M} = 0.25$, $bdp_{2SGMS\ 1\ j,S} = 0.50$, $bdp_{2SGMS\ 1\ j,M} = 0.25$, $bdp_{2SGMS\ 2\ j,S} = 0.25$, $bdp_{2SGMS\ 2\ j,M} = 0.25$, $j=1,2$

case D: $bdp_{WMS,S} = 0.50$, $bdp_{WMS,M} = 0.25$, $bdp_{2SGMS\ 1\ j,S} = 0.25$, $bdp_{2SGMS\ 1\ j,M} = 0.25$, $bdp_{2SGMS\ 2\ j,S} = 0.50$, $bdp_{2SGMS\ 2\ j,M} = 0.25$, $j=1,2$

case E: $bdp_{WMS,S} = 0.50$, $bdp_{WMS,M} = 0.25$, $bdp_{2SGMS\ 1\ j,S} = 0.50$, $bdp_{2SGMS\ 1\ j,M} = 0.25$, $bdp_{2SGMS\ 2\ j,S} = 0.50$, $bdp_{2SGMS\ 2\ j,M} = 0.25$, $j=1,2$

Table 7 - MSE of coefficients N =100, T = 5, 20% contamination (1000 replications)

		β_{11}	β_{12}	β_2	η_{11}	η_{12}	η_2
Leverage points (Y and Z_2 only)	HT	0.1273	0.1322	0.1729	0.8234	1.3979	0.2124
	robust HT	0.0088	0.0101	0.0078	0.1072	0.5063	0.0357
Concentrated leverage points (Y and Z_2 only)	HT	0.0963	0.1019	0.1051	0.6144	2.6819	0.2728
	robust HT	0.0097	0.0096	0.0069	0.1165	1.0016	0.0831
Leverage points (only Z_2)	HT	0.0028	0.0030	0.0026	0.0027	0.1316	0.0105
	robust HT	0.0048	0.0039	0.0049	0.0029	0.2000	0.0165
Concentrated leverage points (only Z_2)	HT	0.0026	0.0025	0.0028	0.0023	0.1384	0.0089
	robust HT	0.0043	0.0042	0.0046	0.0025	0.2121	0.0185

Table 8 - MSE of coefficients N =100, T = 5, 10% contamination, robust and quasi-robust Hausman-Taylor estimators (1000 replications)

		β_{11}	β_{12}	β_2	η_{11}	η_{12}	η_2
no outlier	HT	0.0023	0.0026	0.0028	0.0010	0.1235	0.0102
	robust HT	0.0042	0.0049	0.0046	0.0014	0.2010	0.0164
	quasi-robust HT (WMS, HT)	0.0027	0.0026	0.0025	0.0008	0.1183	0.0105
	quasi-robust HT (Within, 2SGMS)	0.0043	0.0042	0.0045	0.0015	0.1950	0.0152
	Contamination (10%)						
Vertical outliers	HT	0.0718	0.0770	0.1111	0.2481	0.9964	0.1379
	robust HT	0.0067	0.0076	0.0056	0.0103	0.2523	0.0197
	quasi-robust HT (WMS, HT)	0.0843	0.0848	0.1173	0.2512	0.9760	0.1348
	quasi-robust HT (Within, 2SGMS)	0.0070	0.0072	0.0069	0.0042	0.2315	0.0180
Leverage points	HT	1.6804	1.6431	1.9832	0.1348	0.7264	1.5328
	robust HT	0.0120	0.0112	0.0117	0.0087	0.2242	0.0481
	quasi-robust HT (WMS, HT)	1.7202	1.6030	1.9746	0.1312	0.8288	1.5557
	quasi-robust HT (Within, 2SGMS)	0.0174	0.0163	0.0118	0.0152	0.2835	0.0650
Concentrated vertical outliers	HT	0.0707	0.0682	0.0914	0.2116	1.5824	0.1503
	robust HT	0.0073	0.0064	0.0051	0.0225	0.4390	0.0327
	quasi-robust HT (WMS, HT)	0.0745	0.0809	0.1009	0.2187	1.3779	0.1685
	quasi-robust HT (Within, 2SGMS)	0.0064	0.0072	0.0054	0.0161	0.3855	0.0335
Concentrated leverage points	HT	1.2417	1.2015	1.3094	0.1189	1.0429	0.7424
	robust HT	0.0204	0.0211	0.0209	0.0194	0.4306	0.1102
	quasi-robust HT (WMS, HT)	1.2562	1.2113	1.3700	0.1187	0.9489	0.7312
	quasi-robust HT (Within, 2SGMS)	0.0193	0.0215	0.0177	0.0157	0.3987	0.1083
Leverage points with Z_{12}	HT	0.7343	0.7314	0.7407	0.0503	64.7983	0.8470
	robust HT	0.0231	0.0225	0.0197	0.0095	1.1644	0.0609
	quasi-robust HT (WMS, HT)	0.7181	0.7615	0.7533	0.0479	61.4710	0.7458
	quasi-robust HT (Within, 2SGMS)	0.0253	0.0263	0.0203	0.0209	0.9807	0.0917
Concentrated leverage points with Z_{12}	HT	0.5671	0.6129	0.6127	0.0620	70.5724	0.5393
	robust HT	0.0244	0.0261	0.0276	0.0198	0.8332	0.1154
	quasi-robust HT (WMS, HT)	0.5920	0.5941	0.6386	0.0854	81.9930	0.5391
	quasi-robust HT (Within, 2SGMS)	0.0231	0.0251	0.0216	0.0159	0.8725	0.1028
Leverage points with Z_{12} and Z_2	HT	0.7279	0.8791	0.8224	0.0535	63.8211	0.8714
	robust HT	0.0208	0.0192	0.0168	0.0087	0.9404	0.0486
	quasi-robust HT (WMS, HT)	0.7456	0.8357	0.8438	0.0513	57.6370	0.8320
	quasi-robust HT (Within, 2SGMS)	0.0242	0.0212	0.0182	0.0233	0.9265	0.0906
Concentrated leverage points with Z_{12} and Z_2	HT	0.6235	0.5389	0.5822	0.0567	67.5897	0.4981
	robust HT	0.0195	0.0218	0.0221	0.0189	0.8958	0.1032
	quasi-robust HT (WMS, HT)	0.6550	0.5567	0.6400	0.0731	76.7050	0.4874
	quasi-robust HT (Within, 2SGMS)	0.0227	0.0213	0.0200	0.0173	0.8710	0.0899

Table 9 - Cornwell and Rupert (1988) Mincer wage equation

	WITHIN			WMS		
	Coeff.	s.e	T-stat	Coeff.	s.e	T-stat
OCC	-0.0215	0.0138	-1.5581	-0.0224	0.0027	-8.2684
SOUTH	-0.0019	0.0343	-0.0543	-0.0975	0.0025	-38.8978
SMSA	-0.0425	0.0194	-2.1859	-0.0209	0.0024	-8.6947
IND	0.0192	0.0154	1.2437	-0.0177	0.0024	-7.3852
EXP	0.1132	0.0025	45.8141	0.1105	0.0005	220.0491
EXP2	-0.0004	0.0001	-7.6629	-0.0005	0.0000	-42.4252
WKS	0.0008	0.0006	1.3940	0.0020	0.0001	16.9810
MAR	-0.0297	0.0190	-1.5659	0.0345	0.0032	10.9441
UNION	0.0328	0.0149	2.1970	0.0363	0.0025	14.2990

	HAUSMAN-TAYLOR			ROBUST HAUSMAN-TAYLOR		
	Coeff.	s.e	T-stat	Coeff.	s.e	T-stat
OCC	-0.0207	0.0138	-1.5024	-0.0184	0.0038	-4.8763
SOUTH	0.0074	0.0320	0.2328	-0.0162	0.0130	-1.2451
SMSA	-0.0418	0.0190	-2.2066	-0.0327	0.0069	-4.7163
IND	0.0136	0.0152	0.8928	-0.0205	0.0045	-4.5335
EXP	0.1131	0.0025	45.7850	0.1077	0.0007	164.1500
EXP2	-0.0004	0.0001	-7.6718	-0.0004	0.0000	-26.5820
WKS	0.0008	0.0006	1.3963	0.0009	0.0002	5.4072
MAR	-0.0299	0.0190	-1.5728	-0.0121	0.0052	-2.3474
UNION	0.0328	0.0149	2.1982	0.0116	0.0050	2.3116
INTERCEPT	2.9127	0.2837	10.2690	3.1716	0.0690	45.9800
FEM	-0.1309	0.1267	-1.0337	-0.0650	0.0254	-2.5564
BLK	-0.2858	0.1557	-1.8352	-0.0832	0.0347	-2.3970
ED	0.1379	0.0212	6.4919	0.1320	0.0051	25.8780

	WITHIN			WMS		
	Coeff.	s.e	T-stat	Coeff.	s.e	T-stat
OCC	-0.0216	0.0137	-1.5712	-0.0286	0.0027	-10.4797
SMSA	-0.0445	0.0193	-2.3059	-0.0265	0.0024	-11.0088
IND	0.0189	0.0154	1.2222	0.0007	0.0024	0.2778
EXP	0.1133	0.0025	45.8456	0.1134	0.0005	224.5598
EXP2	-0.0004	0.0001	-7.6682	-0.0004	0.0000	-38.5920
WKS	0.0008	0.0006	1.3819	0.0026	0.0001	23.9203
UNION	0.0327	0.0149	2.1944	-0.0148	0.0025	-5.9975

	HAUSMAN-TAYLOR			ROBUST HAUSMAN-TAYLOR		
	Coeff.	s.e	T-stat	Coeff.	s.e	T-stat
OCC	-0.0210	0.0137	-1.5279	-0.0155	0.0039	-3.9153
SMSA	-0.0434	0.0189	-2.3013	-0.0287	0.0069	-4.1785
IND	0.0133	0.0152	0.8746	-0.0195	0.0047	-4.1591
EXP	0.1132	0.0025	45.8140	0.1077	0.0007	157.6600
EXP2	-0.0004	0.0001	-7.6728	-0.0004	0.0000	-24.5100
WKS	0.0008	0.0006	1.3789	0.0009	0.0002	5.9036
UNION	0.0331	0.0149	2.2196	0.0095	0.0051	1.8813
INTERCEPT	2.8666	0.2839	10.0960	3.0986	0.0717	43.2180
FEM	-0.1046	0.1257	-0.8322	-0.0535	0.0271	-1.9760
BLK	-0.2789	0.1559	-1.7896	-0.0729	0.0375	-1.9419
ED	0.1396	0.0214	6.5265	0.1365	0.0054	25.3530