Robust Estimation of Panels with Cross-Sectional Dependence in the Presence of Outliers

by

Teck Thye Chua

A Thesis Submitted to

The University of Birmingham

for the degree of

DOCTOR OF PHILOSOPHY

Birmingham Business School

Department of Economics

The University of Birmingham

September 2023

Abstract

Outliers are common in empirical datasets. There is a reluctance to use robust techniques due to lack of expertise, despite the large availability of robust literature. Cross-section dependence in panel data is also the norm and not the exception. The Common Correlated Effects estimator (CCE) is robust to cross-section dependence, but not robust to outliers as it is based on least squares. We improved on the robustness of the current existing robust CCE estimator, the Trimmed CCE (TCCE) estimators. We applied our robust TCCE estimators to economic growth data, and we find overall positive and significant results under aggregated data. We also find that the global economy is more varied in structure than the typical mixed-distribution outlier model in the literature. We explored outlier robust versions of the CCE estimators in general, and discovered potential new robust estimators for further research.

Acknowledgements

I would like the thank my supervisors, Dr Yiannis K. and Prof Anindya B. for their continued patience and support throughout the thesis. Special mentions to my PhD classmate, Maria Theresa G., Dr. Alessandra G., and the rest of the Department of Economics at the University of Birmingham for always being positive and encouraging. I would like to thank my lecturers and support staff from the University of Birmingham for their support over the years. Most of all I would like to thank my wife Gwyneth for ideas on women's studies, and for her loving and faithful support for which this thesis would not be possible. Finally, this thesis is a dedication to our newborn daughter, Rebecca.

Contents

List of Figures

List of Tables

Chapter 1

Introduction

Outliers are common in statistical data, being the norm rather than the exception (see Rouseeuw and Leroy, 1987 for example). There is a reluctance to use robust statistics due to lack of expertise, despite the large availability in the literature since the 1960s. Consideration of cross-section dependence in panel data is also the exception rather than the norm. The Common Correlated Effects (CCE) estimator addresses the issue of cross-section dependence in factor models. For a review of cross-section dependence and the CCE estimators, see Pesaran (2006); Chudik and Pesaran (2013), and others. However, the CCE estimators are extremely sensitive to outliers. A small perturbation is sufficient to distort estimates and render standard errors biased and inconsistent. On the other hand, robust literature is plentiful for least squares regressions, but is currently very rare in factor models. Our research aims to introduce robust estimation into the CCE estimators, in order to make the CCE estimators resistant to outliers.

1.1 Outline of the Thesis

The next chapter reviews the robust estimation literature. Chapter 3 reviews cross-section dependence and CCE estimation, culminating in simulation studies to illustrate the effects of outliers on the CCE estimators. In Chapter 4, we take a look at robustifying the Trimmed CCE (TCCE estimators; Lee and Sul, 2020). We then apply our outlier-robust TCCE estimators to economic growth data in chapters 5 and 6. Chapter 7 explores the development of a general one-step robust CCE estimator. Chapter 8 concludes with future research.

References

Chudik, A. and Pesaran, M., 2013. Large Panel Data Models with Cross-Sectional Dependence: A Survey. *SSRN Electronic Journal*

Pesaran, M.H., 2006. Estimation and inference in large heterogeneous panels with a multifactor error structure. *Econometrica*, 74(4), pp.967-1012.

Rousseeuw, P. and Leroy, A., 1987. Robust Regression And Outlier Detection. *Hoboken: John Wiley & Sons*.

Chapter 2

Introduction to Robust Statistics: A Review of Statistical Methods

2.1 Outliers in the Data

The presence of outliers can lead to erroneous estimates in regression models. The classical least-squares (LS) estimator is known to be very sensitive to outliers. In classical LS, it takes only one outlying data point to produce arbitrarily far estimates (see Rousseeuw, 1984). Thus, all estimators based on LS are extremely sensitive to outliers. Outliers can be traced as far back as the 18th century. Swiss mathematician Daniel Bernoulli, in his memoir (Bernoulli, 1777), talked of the "absurdity" of assuming that all observations are "of the same weight or moment" and that they could not be "of the same probability" everywhere.

Douglas M. Hawkins (1980) gave a working definition of an outlying data point:

> "An outlier is an observation which deviates so much from the other observations as to arouse suspicions that it was generated by a different mechanism."

Indeed, outliers can be modelled by the mixed distribution model (Bramati & Croux, 2007; Maronna et al., 2019):

$$F(x) = \epsilon H(x) + (1 - \epsilon)G(x) \tag{2.1}$$

The distribution F of the data consists of the main distribution G and a fraction ϵ of the data being contaminated by an unknown distribution H. For example, when G and H are distinct normal distributions, the resulting data is normally distributed with possibly different mean and variance from the bulk of the data (or the "true" distribution). This, of course produces biases in the estimates of scale and location.

To illustrate the effect of an outlier, consider the following sample of observations (adapted from Rousseeuw & Hubert, 2011):

8.25	8.51	8.36	8.81	7.29

The classical estimates of location is the sample mean

$$\bar{x} = \frac{1}{N}\sum_{i=1}^{N} x_i \tag{2.2}$$

In this case, the mean is simply 8.244. Now suppose that there is an incorrect keystroke in the second observation:

8.25	85.1	8.36	8.81	7.29

The resulting mean is $\bar{x} = 23.562$, which is very far from the true mean. On the other hand, the sample median is 8.36, which is much closer. The sample median is the robust analogue of the mean.

There are 2 main types of outliers: vertical and leverage points (Rousseeuw & Leroy 1987; Bramati & Croux 2007). Vertical outliers are outliers in the y-direction, while leverage points are outliers in the x-direction. Rousseeuw and Leroy illustrated the effect of a single outlier in their published book (see Rousseeuw & Leroy 1987, pp. 3-8). Suppose there are 5 data points (y_i, x_i), $i = 1, 2, \ldots, 5$. Figure 2.1(a) depicts these points and their corresponding LS regression line. In Figure 2.1(b), the fourth y-observation is increased to simulate a vertical outlier. We see that the resulting regression line is tilted, resulting in an upward bias in the regression parameter.

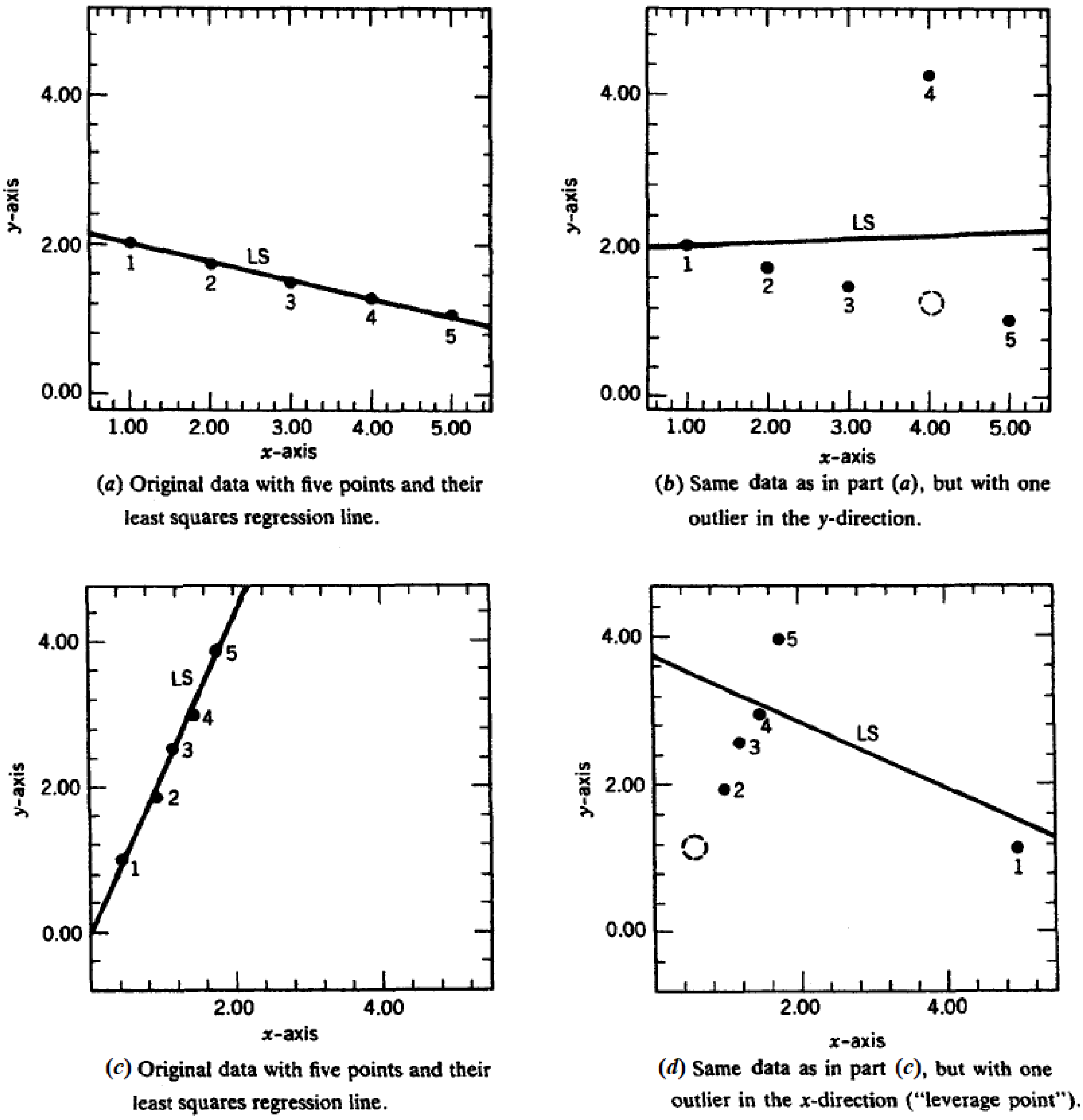
In Figure 2.1(c), we look at another set of five data points. The first x-observation is now increased to simulate the outlier. The resulting regression line is rotated, hence the name 'leverage point'.

2.1.1 The Need for Robust Statistics

Formally, a method or technique is robust if it is not affected by the deletion or addition of a few (possibly spurious) points in the data.

In statistical methods, the two main approaches to dealing with outliers are Robust Regression and outlier diagnostics (see Rousseeuw & Leroy, 1987; Aggarwal, 2017; Maronna et al., 2019). Outlier diagnostics is the employment of statistical methods to identify outliers. Once the suspect data points are identified, they are either discarded or downweighted at the discretion of the experimenter.

Figure 2.1: Vertical and leverage points in LS regression. Source: Rousseeuw & Leroy 1987, pp. 4-5



(*a*) Original data with five points and their least squares regression line.

(*b*) Same data as in part (*a*), but with one outlier in the *y*-direction.

(*c*) Original data with five points and their least squares regression line.

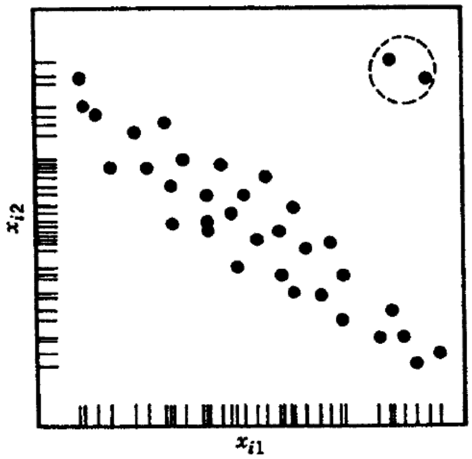
(*d*) Same data as in part (*c*), but with one outlier in the *x*-direction ("leverage point").

The altered dataset is then usually estimated by LS regression. Robust regression, on the other hand, proceeds to fit the data by some outlier-robust regression technique. The outliers are typically identified from the residuals of this robust regression. The corresponding estimators are called robust estimators.

Outliers are more pervasive in the data than we would like to admit in the literature. At the time of this paper, the vast majority of literature either does not employ robust techniques, or they continue to apply classical methods of outlier detection. Zahman, Rousseeuw and Orhan (Zaman et al., 2001) listed reasons the reluctance of using robust methods:

- Belief that large sample sizes make robust methods uneccessary

Figure 2.2: Plot of the explanatory variables (x_{i1}, x_{i2}) of a regression data set. There are two leverage points (dashed circle), neither which are outlying in their respective coordinates. Source: Rousseeuw & Leroy 1987, p. 7



- Belief that outliers can be detected simply by eye or classical methods like unusual residuals or sensitivity analysis
- Little guidance available as to which existing robust techniques are appropriate
- Unfamiliarity with the interpretation of results
- Unawareness of the gains produced by robust statistics
- Many estimation techniques labelled 'robust' are actually not robust techniques at all.

In the same paper (Zaman et al., 2001), Zahman, Rousseeuw and Orhan applied the (robust) Least Trimmed Squares (LTS) estimator to three LS regressions in the literature (De Long & Summer, 1991; Nonneman & Vanhoudt, 1996 and Benderly & Zwick, 1985). In each of these regressions, Zaman , Rousseeuw and Orhan showed that deleting outliers identifed by robust methods tremendously impacts the results, produces much better fits and changes significances of regressors with impacts to economic theory. Misleading results are also eliminated.

This paper attempts to address the above issues in the likeness of Zaman, Rousseeuw and Orhan, in order to promote awareness to robust statistics. We start by reviewing robust statistics in the next section, followed by reviewing extensions of robust statistics in the literature for Panel Data. We follow up with a simple application to demonstrate the use of robust statistics in LS and in Panel regressions. The final section concludes this chapter.

2.2 Robust Statistics

2.2.1 A Brief History of Outliers

The statistical term "robust" was first coined by Box in 1953 (see Stigler, 1973). Outlier detection can be traced as far back as the 18th century, with Bernoulli (1777). Bernoulli's ideas about outliers were as advanced as the foundations of today's outlier analysis. In his memoir (Bernoulli, 1777), he noted that astronomers in his time were already rejecting aberrant observations (taking into account adjustments according to theory) in order to obtain better fits. However, in the same memoir, Bernoulli himself found no systematic way to distinguish the outliers from the main bulk of the data. He also did not discount the possibility that a seemingly deviant observation could actually provide an optimal fit to the bulk of the data. Finally, he illustrated the concept of maximum likelihood (in his example of the archer), noting that outlying observations are of extremely low probability (likelihood). The legacy of this treatment is evident in the simplest of outlier diagnostics, the Z-score. Many statistical tests for outliers have been proposed since (Grubbs, 1950). For a comprehensive history of the development of robust estimation from the time of Bernoulli till the advent of M-estimators, see Stephen M. Stigler's historical review (Stigler, 1973).

2.2.2 Modern Outlier Detection Methods

There is much literature in outlier detection today. However, most of the methods in the literature are based on classical methods. In the past two centuries or so since Bernoulli, the normal distribution has been the foundation of all classical statistical methods. Other non-statistical outlier detection algorithms exist (see Aggarwal, 2017), but the focus of this paper will be statistical methods. Most statistical outlier detection methods (also called outlier diagnostics) are based on classical methods. However, there are two key main drawbacks to outlier diagnostics. Firstly, they are less reliable than robust fits of the data. Secondly, flagged outliers are at the discretion of the experimenter (Maronna et al., 2019; Rousseeuw & Leroy, 1987). This may result in lost information (as also noted by Bernoulli above), in contrast to an objective method of flagging outliers (Maronna et al., 2019). Outlier detection methods are also applied to numerous other fields, such as intrusion detection systems, credit card fraud, voting irregularities, law enforcement, data-cleansing, interesting sensor events, medical diagnosis, clinical trials, geographical information systems and severe weather prediction (Aggrawal, 2017; Ben-Gal, 2009).

Despite a slow uptake, there is much literature for robust statistics. Influential textbooks covering the foundations of robust statistics have been written, beginning with Huber (1981), Rousseeuw & Leroy

(1987) and (Barnett & Lewis (1994). Many of the robust estimators have been historically difficult to compute due to the lack of closed form expressions and complex algorithms, however with advances in computing, this is no longer a problewm for most techniques. More recent and influential textbooks include Maronna et al, (2019), Wilcox (2022) and Huber & Ronchetti (2011). Modern robust statistics began with Peter Huber, John W. Tukey and Frank Hampel(Huber, 1964; 1967; Tukey, 1960,; 1962; Hampel, 1971; 1974, cited by Maronna et. al., 2019). Huber developed the Huber-M estimator, while Tukey developed the Tukey Bisquare function. In the 1980s, Peter Rouseeuw developed the Least Median of Squares (LMS, Rouseeuw, 1984) and the Least Trimmed Squares (LTS) as robust analogues to LS. Although incredibly robust to outliers (resisting up to 50% contamination of data), the relative efficiencies of the LMS and LTS estimators were dismal (0% and 7% respectively).
Contrary to popular opinion (Rousseeuw & Leroy, 1987), it is insufficient to identify outliers from LS residuals alone. Extremely large LS residuals may hint at an outlier, but LS residuals are computed from mean-based functions, as illustrated above. Hence, it is possible for outliers to pull the estimates away from their true means, resulting in misleading results and consequently, incorrect residuals (see Rousseeuw & Hubert, 2001). This is also true of many outlier diagnostic techinques, which are often based on some function of or computation of the mean and standard deviation. Based on the example above, just like the mean, the standard deviation is very sensitive to outliers. Robust alternatives for outlier diagnostics (see Rousseeuw & Hubert 2001, Hodges & Austin 2004 among many others) have also been proposed in the literature.

2.2.3 The Concept of Breakdown Point

Classical methods of outlier detection assume a Gaussian distribution of the data. However, as illustrated in section 2.1, the sample mean is non-robust, where only a single point is required to significantly distort the estimate. Similarly, the median is robust to data contamination up to half of the observations. In the robust statistics literature, the breakdown point is a key measure of a method to data contamination.

The Breakdown Point

A widely used quantitative measure of the robustness of a technique is the breakdown point (BP). The BP was first introduced by Hodges in 1967 in a univariate context and generalised by Hampel in 1971 (Hodges 1967, Hampel 1971, cited in Rousseeuw & Leroy 1987, p. 9). A common working definition of the breakdown point is given by Donoho and Huber (1980, cited in Rousseeuw & Leroy 1987, p. 9):

Let $Z = \{(\mathbf{x_i}, y_i) \mid i = 1, 2, \dots, N\}$ be an uncontaminated dataset. Let $Z' = \{(\tilde{\mathbf{x_i}}, \tilde{y_i}) \mid (\tilde{\mathbf{x_i}}, \tilde{y_i}) = (\mathbf{x_i}, y_i)$ for some $i = 1, 2, \dots, N$ and $\#\{(\tilde{\mathbf{x_i}}, \tilde{y_i}) \neq (\mathbf{x_i}, y_i)\} \leq M\}$, where $\#A$ is the cardinality, which is the number of elements of of the set A. Thus, Z' represents a contaminated dataset of Z containing at most M contaminated points. Let R be a regression estimator of Z. Thus, $R(Z)$ is the vector of regression estimates obtained from the dataset Z, and $R(Z')$ is the estimate obtained from applying the regression estimator R to contaminated Z'. The breakdown point of R, $\epsilon_N^*(R, Z)$ is given by

$$\epsilon_N^*(R, Z) = \min_{1 \leq M \leq N} \left\{ \frac{M}{N} \;\middle|\; \sup_{Z'} \left\| R(Z') - R(Z) \right\| = \infty \right\} \tag{2.3}$$

where sup is the supremum over all possible contaminated datasets Z' and $\|\cdot\|$ is the Euclidean norm. Intuitively, the BP is the smallest fraction of outliers that an estimator can withstand before the estimates deviate arbitrarily far from their true values.

As illustrated in the *Pilot-Plant* data in figure 2.2, a single aberrant point is enough to take an LS estimate far from its true mean. Thus, the finite sample BP of the LS estimator is $1/N$. For large N, the asymptotic breakdown point is 0%. Similarly, the classical estimates of location and scale are also 0%.

Rousseeuw (1984) showed that the finite sample BP of the LMS is $([N/2] + k - 2)/N$, where k is the number of regressors in the model (in the case of the *Pilot-Plant* data, $k = 1$). For large N,

$$\begin{aligned} \frac{[\frac{N}{2}] + k - 2}{N} &\cong \frac{\frac{N}{2} + k - 2}{N} \\ &= \frac{1}{2} + \frac{k}{N} - \frac{2}{N} \\ &\longrightarrow \frac{1}{2} \text{ as } N \longrightarrow \infty \end{aligned}$$

Thus, the asymptotic BP of the LMS is 50%.

Regression, Scale and Affine Equivariance

Three desirable properties of estimators are regression, scale and affine equivariance (see Rousseeuw & Leroy 1987). Given an estimator R and a dataset $(\mathbf{x}_i, y_i)$:

Regression equivariance is defined as

$$R(\mathbf{x}_i, y_i + \mathbf{x}_i\mathbf{v}) = R(\mathbf{x}_i, y_i) + \mathbf{v} \tag{2.4}$$

for any $k \times 1$ vector $\mathbf{v}$. Monte Carlo studies often select a fixed value of the regression parameter $\boldsymbol{\beta}$ in their simulations. Regression equivariance ensures that the results of their simulations apply to all other translated values of $\boldsymbol{\beta}$.

Scale equivariance is defined as

$$R(\mathbf{x}_i, cy_i) = cR(\mathbf{x}_i, y_i) \tag{2.5}$$

for any constant c. Scale equivariance ensures that the fit holds irregardless of the choice of units used (Rousseeuw & Leroy 1987).

Affine equivariance is definied as

$$R(\mathbf{x}_i\mathbf{A}, y_i) = \mathbf{A}^{-1}R(\mathbf{x}_i, y_i) \tag{2.6}$$

for any $k \times k$ nonsingular matrix $\mathbf{A}$. This implies that a given linear transformation of $\mathbf{x}_i$ should transform the estimator R accordingly. This allows different coordinate systems to be used for the observations, without affecting the estimate $\hat{y}_i$ (Rousseeuw & Leroy 1987).

The LS and LMS estimators are known to be regression, scale and affine equivariant. We now take a look at classical outlier detection methods.

2.2.4 Classical Outlier Diagnostics

Classical outlier detection methods assume the data follows a Gaussian distribution.

Z-score

The simplest outlier diagnostic is the Z-score, available in many published books, such as Aggarwal (2017) and Maronna et al. (2019). The Z-score measures the standardised distance between an observation and its mean:

$$Z = \frac{|X_i - \bar{X}|}{\sigma} \tag{2.7}$$

For iid univariate data $X_i,\ i = 1, 2, \ldots, n$ and σ the standard deviation of X. The "three-signma edit rule" is to reject observations with Z-scores larger than 3 (Aggarwal, 2017; Maronna et al., 2019) or 2.5 (Rousseeuw & Leroy, 1987). Although this rule is based on the standard normal distribution, it can be easily extended to any unimodal symmetric distribution with positive density function, including the multivariate case (Ben-Gal, 2009). Similar diagnostics to the Z-score include Chauvenet's criterion

(Chauvenet, 1863) or the more general Peirce's criterion (Peirce, 1852; Ross, 2003) and the Grubb's test (Grubbs, 1950). The fundamental issue with the Z-score and its related methods is the sensitivty of the sample mean and standard deviation to outliers. It only takes one sufficiently large observation to distort the statistics, i.e. the BP of these diagnostic techniques is 0%.

Residuals and Leverage

Adrien-Marie Legendre (Legendre, 1805; Stigler 1973) spoke of trimming suspicious residuals:

> If among these errors are some which appear too large to be admissible, then those equations which produced these errors will be rejected, as coming from too faulty experiments, and the unknowns will be determined by means of the other equations, which will then give much smaller errors.

Vertical outliers are often dealt with by examining the residuals of a least squares regression (see Rousseeuw & Leroy, 1987; Hansen, 2022). However, the residuals themselves are suceptible to masking by multiple outliers. Residual analysis also does not always detect leverage points. A classical statistic to detect leverage points is the leverage statistic, given by the diagonal entries of the hat matrix

$$P = \mathbf{X}(\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}' \tag{2.8}$$

The i-th leverage statistic, h_{ii}, is the i-th diagonal entry of P. The i-th leverage statistic is also related to the i-th observation by

$$h_{ii} = \mathbf{x}_i(\mathbf{X}'\mathbf{X})^{-1}\mathbf{x}_i' \tag{2.9}$$

Which can be interpreted as the standardised distance between each data point $\mathbf{x}_i$ and the mean of the data cloud $\mathbf{X}$. The leverage has the following properties: for p-variate $\mathbf{x}_i$,

$$\sum_{i=1}^{N} h_{ii} = p \tag{2.10}$$

$$0 \leq h_{ii} \leq 1 \tag{2.11}$$

An observation is deemed influential if it has an unusually large leverage value compared to the rest. Thus, by (2.10), leverages much larger than $\frac{p}{N}$ are flagged as suspicious. Although the leverage does not include the dependent variable in its compuation, it is actually related to the residuals:

$$\hat{\mathbf{y}} = P\mathbf{y} \tag{2.12}$$

Thus, anomalies in the leverage affects the residuals, however influential residuals could either be due to leverage points or vertical outliers. The leverage statistic is effective at singling out a single outlier sufficiently far away from the data cloud. However, if multiple outliers exist, they can mask the presence of each other, reducing their leverage size and possibly avoiding detection altogether.

Mahalanobis Distance

The multivariate case of the Z-score is the Mahalanobis distance (Mahalanobis, 1936) The Mahalanobis distance is relatively ubiquitous in statistical multivariate outlier detection. A search on Semantic Scholar for the original paper (Mahalanobis, 1936) returned 410 influential citations, 1183 background citations and 2358 method citations. It measures the standardised Euclidean distances between each observation point and the centre of the data cloud. The Mahalanobis distance, MD_i, is given by

$$MD_i^2 = (\mathbf{X}_i - \boldsymbol{\mu})'\boldsymbol{\Sigma}^{-1}(\mathbf{X}_i - \boldsymbol{\mu}) \tag{2.13}$$

The Mahalanobis distance is also related to leverage by the following (Rousseeuw & Leroy, 1987, p.224):

$$MD_i^2 = (N-1)\left[h_{ii} - \frac{1}{N}\right] \tag{2.14}$$

Since the Mahalanobis distance uses the sample mean and sample covariance matrix in its population, the Mahalanobis distance suffers from distorted estimates and the masking effect. Robust analogues to the Mahalanobis distance are the Minimum Covariance Determinant (MCD) and the Minimum Volume Ellipsoid (MVE) (Rouseeuw, 1985; also in Rousseeuw & Leroy, 1987; Maronna et. al. 2019). Each of these estimators estimate robust mean vectors and covariance matrices that minimise the determinant of the covariance matrix (effectively minimising the sample variances).

The Masking Effect

A key reason why techniques based on classical methods fail is due to the masking effect (Rousseeuw & Leroy, 1987). The masking effect can be defined as the effect of one or more other influential points that make a given influential point appear non-outlying. Those other influential points are said to ”mask” the given outlier. Thus, the detrimental effects of masking give misleading results with many classical methods.

Figure 2.3: *Pilot-Plant* data (Daniel & Wood 1971, cited in Rousseeuw & Leroy 1987, p. 22)

Observation (i)	Extraction (x_i)	Titration (y_i)
1	123	76
2	109	70
3	62	55
4	104	71
5	57	55
6	37	48
7	44	50
8	100	66
9	16	41
10	28	43
11	138	82
12	105	68
13	159	88
14	75	58
15	88	64
16	164	88
17	169	89
18	167	88
19	149	84
20	167	88

2.2.5 Robust Regression

Robust regression is the development of estimators that are resistant to outliers. Outliers arise from many sources: incorrect keystrokes, including misplaced decimal points, errors in recording or transmission, shocks such as natural disasters or strikes, or possibly members of a different population included in the sample. With much data now processed by computers with little or no screening, outliers often go unnoticed.

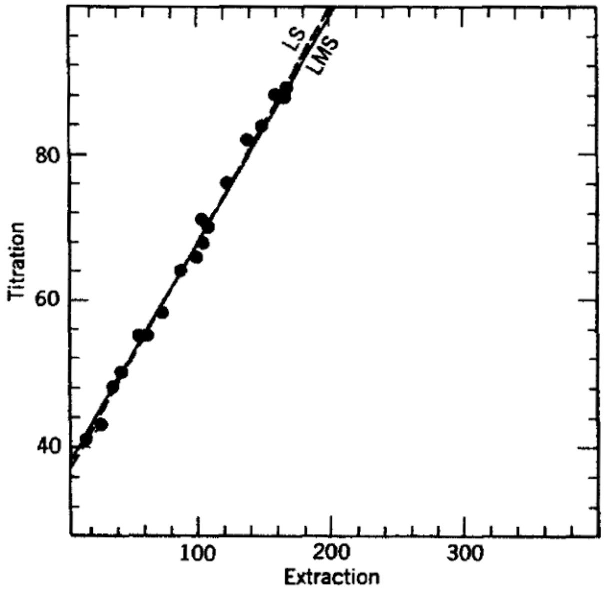
A robust alternative to LS estimation is the Least Median of Squares (LMS) estimator (Rousseeuw 1984), given by

$$\hat{\beta}_{LMS} = \underset{\beta}{\operatorname{argmin}}(\underset{i}{\operatorname{med}}\ r_i^2) \tag{2.15}$$

Where med is the median over all the squared residuals $r_i^2 = (y_i - \mathbf{x}_i\hat{\boldsymbol{\beta}})^2$ of the classical regression model $y_i = \mathbf{x}_i\boldsymbol{\beta} + e_i$. Thus, the LMS minimises the median of the squared residuals instead of their sum.

We now demonstrate the robustness of the LMS compared to LS. The following example is from Rousseeuw & Leroy (1987, ch 2), which features the *Pilot-Plant* data (figure 2.3) from Daniel & Wood (1971). The scatterplot of Figure 2.2 indicates a strong linear relationship between regressor and regressand.

Figure 2.4: *Pilot-Plant* data with LS fit (dashed line) and LMS fit (dotted line)(Rousseeuw & Leroy 1987, p. 23)



The LS fit of the data (Figure 2.4, dashed line) is

$$\hat{y} = 0.322x + 35.458 \tag{2.16}$$

The LMS fit (Figure 2.4, dotted line) is

$$\hat{y} = 0.311x + 36.519 \tag{2.17}$$

The plot in Figure 2.4 contains no outliers: the LMS coincides very strongly with the LS. The authors, Rousseeuw and Leroy then replace observation 6 with 370 instead of 37 (illustrated on the scatterplot in Figure 2.5). The resulting LS becomes

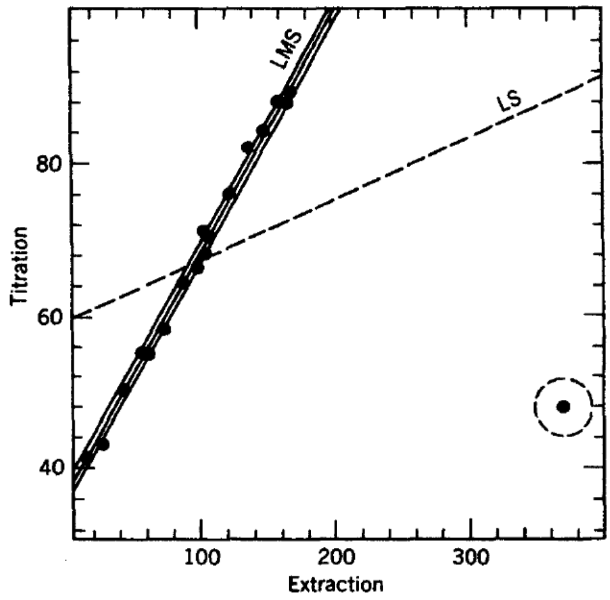
$$\hat{y} = 0.081x + 58.939 \tag{2.18}$$

and the resulting LMS is

$$\hat{y} = 0.314x + 36.343 \tag{2.19}$$

The LMS has successfully resisted the outlier. Furthermore, it is also very close to the uncontaminated estimate. Rousseeuw and Leroy then proceeded to demonstrate the same effectiveness of LMS in real data (see Rousseeuw & Leroy 1987, pp. 25-28).

Figure 2.5: *Pilot-Plant* data but with one outlier. The LMS line is surrounded by the narrowest strip containing half of the points.(Rousseeuw & Leroy 1987, p. 24)



Least Trimmed Squares

Although it has a high breakdown point, the LMS estimator has a low efficiency as compared to the LS. To improve on its efficiency, Rousseeuw (1984, cited in Rousseeuw & Leroy 1987) devised the least trimmed squares (LTS) estimator:

$$\hat{\beta}_{LTS} = \underset{\beta}{\text{argmin}} \sum_{i=1}^{N} (r^2)_{i:N} \tag{2.20}$$

Where

$$(r^2)_{1:N} \leq (r^2)_{2:N} \leq \ldots \leq (r^2)_{N:N} \tag{2.21}$$

are the N squared residuals $r_i^2 = (y_i - \mathbf{x}_i\hat{\boldsymbol{\beta}})^2$, arranged in ascending order. The residuals are first ordered, and the sum of the smallest h residuals are computed. The LTS is the value of β that gives rise to the smallest sum of h residuals. By selecting the h smallest residuals, the largest $N-h$ residuals are not used, effectively allowing the estimator to stay away from the outliers.

This estimator is also regression, scale and affine equivariant, with a breakdown point of no more than $([(N-k)/2]+1)/N$, where β is a $k \times 1$ vector. Here, $[\cdot]$ is the *floor* function, that is, $[x]$ is the integer component of x.

Rousseeuw (Rousseeuw & Leroy, 1987) also showed that the maximum breakdown point a regression equivariant estimator can attain is $([(N-k)/2]+1)/N$. The LTS attains this maximum if h is set to $h = [n/2] + [(k+1)/2]$.

For large N:

$$\frac{[\frac{N-k}{2}]+1}{N} \cong \frac{\frac{N-k}{2}+1}{N}$$
$$= \frac{1}{2}\Big(\frac{N-k}{N}\Big) + \frac{1}{N}$$
$$\longrightarrow \frac{1}{2} \text{ as } N \longrightarrow \infty$$

Thus, like the LMS, the asymptotic BP of the LTS is also 50%. Because of their high breakdown point and simplicity, the LMS and LTS are often used in preliminary estimations before a more efficient estimator is applied. We will see an example of the LTS used as a preliminary estimator in section 2.3.

Although the LMS and LTS are very robust against outliers, they are known to be very inefficient due to their extremely slow asymptotic convergence rates. Many more efficient and still relatively robust alternatives have been put forward. In the following section, we take a look at some notably more efficient robust estimators in the literature. These examples will again be further applied in the panel data context in chapter 2.

2.2.6 Robust Estimators

M-Estimators

In traditional LS, the motivation is to minimise the sum of squared residuals $r_i = (y_i - \mathbf{x}_i'\boldsymbol{\beta})$

$$\min_{\beta} \sum_{i=1}^{N} r_i^2 \tag{2.22}$$

with first-order conditions (FOC)

$$\sum_{i=1}^{N} r_i \mathbf{x}_i = \mathbf{0} \tag{2.23}$$

To generalise the concept in (2.22), the basic idea was to minimise a loss function of the standardised residuals instead:

$$\hat{\boldsymbol{\beta}} = \underset{\beta}{\text{argmin}} \sum_{i=1}^{N} \rho\big(\frac{r_i}{\hat{\sigma}}\big) \tag{2.24}$$

ρ is a robust loss function of the residuals r_i.

The scale estimate $\hat{\sigma}$ is any positive statistic that is scale equivariant (see Huber & Ronchetti, 2009 p.105), that is $\hat{\sigma}(\alpha\mathbf{x}_1, \alpha\mathbf{x}_2, \ldots, \alpha\mathbf{x}_N) = \alpha\hat{\sigma}(\mathbf{x}_1, \mathbf{x}_2, \ldots, \mathbf{x}_N)$ for any positive real number α. Scale often occurs as a nuisance parameter in robust regression and M-estimates are not scale equivariant without a scale estimate, hence the need to standardise the residuals in (2.24). Again, by being scale equivariant, one does not need to worry about the choice of units used in the measurements.

Two robust estimates of scale are (see Rousseeuw & Leroy 1987, Rousseeuw & Hubert 2001):

$$\hat{\sigma} = C_1 \sqrt{\underset{i}{\text{med}}\, r_i} \tag{2.25}$$

and

$$\hat{\sigma} = C_2 \sqrt{\sum_{i=1}^{h} r_{i:N}} \tag{2.26}$$

(2.25) and (2.26) are, respectively, LMS and LTS estimates of the standard deviation of the errors. Being highly robust to outliers, they are often used as preliminary estimates of scale.

C_1 and C_2 are constants selected for consistency under normal errors.

The solution to (2.24), $\hat{\boldsymbol{\beta}}$, is known as an M-estimator (of location). The FOC to (2.24) is given by

$$\sum_{i=1}^{N} \psi\left(\frac{r_i}{\hat{\sigma}}\right)\mathbf{x}_i' = \mathbf{0} \tag{2.27}$$

The derivative of ρ, $\rho'(x) = \psi(x)$, is known as the influence function. It measures the effect of an additional data point on the estimator. A popular influence function is Tukey's Bisquare (Tukey & Beaton 1974):

$$\psi(x) = \begin{cases} x\left(1 - \left(\dfrac{x}{c}\right)^2\right)^2 & |\, x \,| \leq c \\ 0 & |\, x \,| > c \end{cases} \tag{2.28}$$

with a corresponding (robust) loss function

$$\rho(x) = \begin{cases} 1 - \left(1 - \left(\dfrac{x}{c}\right)^2\right)^3 & |\, x \,| \leq c \\ \dfrac{c^2}{6} & |\, x \,| > c \end{cases} \tag{2.29}$$

Figure 4 shows the graphs of Tukey's bisquare loss and influence functions. The loss function ρ acts

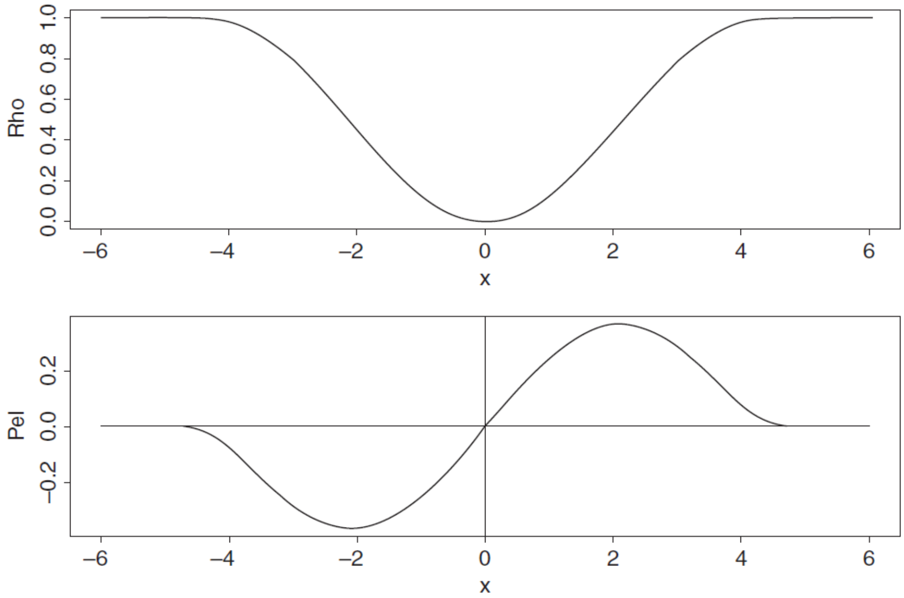


Figure 4. ρ- (labelled Rho) and ψ- (labelled Pel) bisquare functions. (Maronna et. al. 2019)

on the standardised residuals and is capped at a maximum value of 1. This prevents residuals from getting too large in magnitude, thereby reducing the effect of vertical outliers.

The influence function ψ vanishes at the extreme ends. This negates the additional effect a vertical outlier would have on the estimate $\hat{\boldsymbol{\beta}}$. Therefore, a robust loss function and its corresponding influence function bounds the effect of a vertical outlier in the M-estimator. For this reason, M-estimators are also called *bounded-influence estimators.*

The value c is selected as a Terms of Trade-off between efficiency and robustness. Smaller values of c down-weigh more observations, with lower weights applied to vertical outliers (see Wagenvoort a& Waldmann 2001). On the other hand, higher values of c lead to more efficient estimates but lower robustness. Under normal errors, a common choice of c is 4.685 (Beaton & Tukey 1974, see also Wagenvoort and Waldmann 2001), found by simulation. This leads to a relative efficiency of 95% compared to standard maximum likelihood estimates.

To give an example, an M-estimator using the LTS and Tukey's Bisquare as the preliminary scale estimate (2.26) and loss function (2.29) respectively is given by (assuming normal errors):

$$\hat{\boldsymbol{\beta}} = \underset{\beta}{\operatorname{argmin}} \begin{cases} \displaystyle\sum_{i=1}^{N} \left(\frac{y_i - \mathbf{x}_i \boldsymbol{\beta}}{C_2 \sqrt{\sum_{i=1}^{h} r_{i:N}}} \right) \left(1 - \left(\frac{y_i - \mathbf{x}_i \boldsymbol{\beta}}{4.685 (C_2 \sqrt{\sum_{i=1}^{h} r_{i:N}})} \right)^2 \right)^3 & \mid x \mid \leq 4.685 \\ \dfrac{4.685^2}{6} & \mid x \mid > 4.685 \end{cases}$$

If instead, we set $\rho(x) = \dfrac{1}{2} x^2$ and $\hat{\sigma} = 1$, we obtain the familiar OLS.

The M-estimator is very robust against vertical outliers, but the $\mathbf{x}_i$ terms in (2.24) make the M-estimator vulnerable to leverage points, similar to the illustrations in the previous sections. To address the issue of leverage points, Mallows (1975, described in Rousseeuw & Leroy 1987, Maronna et. al. 2019) devised the first generalised M-estimator (GM-estimator).

GM-estimators

The GM-Estimator additionally downweighs the regressors in the FOC:

$$\sum_{i=1}^{N} w(\mathbf{x}_i)\psi\left(\frac{r_i}{\hat{\sigma}}\right)\mathbf{x}_i = \mathbf{0} \tag{2.30}$$

S-Estimators

Another class of high-breakdown estimators with better relative efficiency than the LTS are the S-estimators (Rousseeuw 1984, cited in Rousseeuw & Leroy 1987, Yu & Yao 2017). S-estimators are given by

$$\hat{\beta} = \underset{\beta}{\operatorname{argmin}}\ s(r_1(\beta), r_2(\beta), \ldots, r_N(\beta)) \tag{2.31}$$

with scale estimate

$$\hat{\sigma} = s(r_1(\hat{\beta}), r_2(\hat{\beta}), \ldots, r_N(\hat{\beta})) \tag{2.32}$$

$\sigma \equiv s(r_1(\beta), r_2(\beta), \ldots, r_N(\beta))$ is called the dispersion of the residuals, which is defined as the solution of the scale estimate

$$\hat{\sigma} = \sup_{s}\left\{ s \in \mathbb{R}^{+} \,\middle|\, \frac{1}{N}\sum_{i=1}^{N}\rho\big(\frac{r_i}{s}\big) = K \right\} \tag{2.33}$$

for some real number K, which is often set to $E_{\Phi}[\rho]$, with Φ as the standard normal. ρ must be symmetric, continuously differentiable, strictly increasing on some interval $[0, c]$ with $\rho(0) = 0$, but is often the loss function of an M-estimator of β.

MS-Estimator

The MS-Estimator (Maronna & Yohai 2000) combines both the M- and the S-estimator in a single-step estimation. The MS-estimator was also constructed to accomodate categorical and continuous explanatory variables at the same time. The base model is

$$y_i = \mathbf{x}_{1i}\boldsymbol{\beta}_1 + \mathbf{x}_{2i}\boldsymbol{\beta}_2 + e_i \ , \ i = 1, 2, \ldots, N \tag{2.34}$$

Where $\mathbf{x}_{1i}$ is a $p_1 \times 1$ vector of 0-1 categorical variables (including intercepts) and $\mathbf{x}_{2i}$ is a $p_2 \times 1$ vector of continuous explanatory variables. $\boldsymbol{\beta}_1$ and $\boldsymbol{\beta}_2$ are the location parameters of interest. Let $\mathbf{X} = (\mathbf{X}_1, \mathbf{X}_2)'$ be the $N \times (p_1 + p_2)$ matrix of regressors with $\mathbf{X}_1 = (\mathbf{x}_{11}, \mathbf{x}_{12}, \ldots, \mathbf{x}_{1N})$ and $\mathbf{X}_2 = (\mathbf{x}_{21}, \mathbf{x}_{22}, \ldots, \mathbf{x}_{2N})$. let $\mathbf{y} = (y_1, y_2, \ldots, y_N)'$ the $N \times 1$ vector of dependent variables, let

$$M(\mathbf{X}, \mathbf{y}) = \underset{\boldsymbol{\beta}}{\text{argmin}} \sum_{i=1}^{N} \rho(y_i - \mathbf{x}_i\boldsymbol{\beta}) \tag{2.35}$$

be an M-estimate with loss function ρ.

If $\boldsymbol{\beta}_2$ were known, one can estimate $\boldsymbol{\beta}_1$ with an M-estimator. Rewrite (2.34) as

$$y_i - \mathbf{x}_{2i}\boldsymbol{\beta}_2 = \mathbf{x}_{1i}\boldsymbol{\beta}_1 + e_i \ , \ i = 1, 2, \ldots, N \tag{2.36}$$

Then, for each $\boldsymbol{\beta}_2 \in \mathbb{R}^{p_2}$, define $\boldsymbol{\beta}_1$ as a function (in particular, an M-estimate) of $\boldsymbol{\beta}_2$:

$$\boldsymbol{\beta}_1^*(\boldsymbol{\beta}_2) = \mathbf{M}(\mathbf{X}, \mathbf{y} - \mathbf{X}_2\boldsymbol{\beta}_2) \tag{2.37}$$

Let $\mathbf{r}(\boldsymbol{\beta}_1, \boldsymbol{\beta}_2) = y_i - \mathbf{x}_{1i}\boldsymbol{\beta}_1 - \mathbf{x}_{2i}\boldsymbol{\beta}_2$ represent the residuals obtained by estimating (2.36). Let $S(\mathbf{r})$ be a robust scale estimate of the residuals $\mathbf{r}(\boldsymbol{\beta}_1, \boldsymbol{\beta}_2)$. Substituting (2.37) into the expression for estimating $S(\mathbf{r})$, the MS-estimator is defined as

$$\hat{\boldsymbol{\beta}}_2 = \underset{\boldsymbol{\beta}_2}{\text{argmin}} \ S(\mathbf{r}(\boldsymbol{\beta}_1, \boldsymbol{\beta}_2), \boldsymbol{\beta}_2) \tag{2.38}$$

That is, the MS-estimator combines the formulae for the M-estimate of $\boldsymbol{\beta}_1$ and the robust scale estimate $S(\mathbf{r})$ into a single expression for the estimation of $\boldsymbol{\beta}_2$. The estimate of $\boldsymbol{\beta}_1$ can be retrieved by substituting the estimates obtained in (2.38) into (2.37).

We now look at the application of robust estimation in the context of panel data. We will be looking at the classical within-groups (also known as fixed effects) panel model. This model is based on LS (for a comprehensive description, see Wooldridge, 2010 chapter 10). The following section will showcase a robust estimation technique found in the literature of robust panel data estimation, similar to the example of section 2.2.6.

2.3 Robust Regression in Panel Data

2.3.1 The Within Groups Estimator

The standard fixed effects model is

$$y_i = \alpha_i + \mathbf{x}_{it}\boldsymbol{\beta} + u_{it} \qquad i = 1, 2, \ldots, N \qquad t = 1, 2, \ldots, T \tag{2.39}$$

where $\mathbf{x}_{it} = (x_{it}^{(1)}, x_{it}^{(2)}, \ldots, x_{it}^{(k)})$ is a $1 \times p$ vector of regressors of cross section i at time period t, $\boldsymbol{\beta} = (\beta_1, \beta_2, \ldots, \beta_k)'$ is the $k \times 1$ regression vector and α_i are the individual time-invariant effects. To obtain the standard Within-Groups (WG) estimator, we first center the data for each time series. Define

$$\tilde{y}_{it} = y_{it} - \frac{1}{T}\sum_{i=1}^{T} y_{it} \tag{2.40}$$

$$\tilde{\mathbf{x}}_{it} = \mathbf{x}_{it} - \frac{1}{T}\sum_{i=1}^{T} \mathbf{x}_{it} \tag{2.41}$$

$$\tilde{u}_{it} = u_{it} - \frac{1}{T}\sum_{i=1}^{T} u_{it} \tag{2.42}$$

The WG estimator, $\hat{\boldsymbol{\beta}}_{WG}$, is the OLS regression of

$$\tilde{y}_{it} = \tilde{\mathbf{x}}_{it}\boldsymbol{\beta} + \tilde{u}_{it} \tag{2.43}$$

From the previous sections, we have seen that the mean is very sensitive to outliers. The WG estimator applies the mean in each stage of its estimation (during the centering and the resgression). In the presence of outliers, the mean is biased. Each stage of the WG estimation would compound the bias. We would therefore expect large biases in the WG estimator in the presence of outliers. This would be illustrated in the following examples.

The Breakdown Point and Equivariance in Panel Estimation

2.3.2 Breakdown Point in Panels

Bramati and Croux (2007) expanded the concept of the breakdown point to two panel data analogues: The cell breakdown point (cell BP) and the block breakdown point (block BP).

Given an uncontaminated panel dataset $Z = \{(\mathbf{x_{it}}, y_{it}) \mid i = 1, 2, \ldots, N \; t = 1, 2, \ldots, T\}$ and the contaminated counterpart $Z' = \{(\tilde{\mathbf{x_{it}}}, \tilde{y_{it}}) \mid \#\{(\tilde{\mathbf{x_{it}}}, \tilde{y_{it}}) \neq (\mathbf{x_{it}}, y_{it})\} \leq M, \; i = 1, 2, \ldots, N \; t = 1, 2, \ldots, T\}$, the cell breakdown point is, analogously,

$$\epsilon^*_{NT}(R, Z) = \min_{1 \le M \le NT} \left\{ \frac{M}{NT} \;\middle|\; \sup_{Z'} \left\| R(Z') - R(Z) \right\| = \infty \right\} \tag{2.44}$$

which is the smallest number of observations M that need to be corrupted in order for the estimand of R to deviate arbitrarily far from its true mean.

Sometimes, outliers in panel data can be concentrated in a few time series (for example, outliers confined to a cluster of countries). The block outlier is defined as a time series in a panel dataset that contains at least one outlying observation. The block breakdown point is the minimum fraction of block outliers that is needed to carry the estimates arbitrarily far away.

Let $Z = \{(\mathbf{X}_i, \mathbf{y}_i) \mid i = 1, 2, \ldots, N\}$ be the same uncontaminated panel data set, written in stacked form. Let $Z'_B = \{(\tilde{\mathbf{X}}_i, \tilde{\mathbf{y}}_i) \mid \#\{(\tilde{\mathbf{X}}_i, \tilde{\mathbf{y}}_i) \ne (\mathbf{X}_i, \mathbf{y}_i)\} \le B,\ i = 1, 2, \ldots, N\}$ represent the contaminated dataset with at most B block outliers. The block breakdown point of an estimator R is given by

$$\epsilon^B_N(R, Z) = \min_{1 \le B \le N} \left\{ \frac{B}{N} \;\middle|\; \sup_{Z'_B} \left\| R(Z') - R(Z) \right\| = \infty \right\} \tag{2.45}$$

2.3.3 Regression, Scale and Affine Equivariance

Bramati & Croux (2007) adapted the definitions of regression, scale and affine equivariance to panel data. Given an estimator R and a dataset $(\mathbf{x}_{it}, y_{it})$, for all $i = 1, 2, \ldots, N$ and $t = 1, 2, \ldots, T$:

Regression equivariance is defined in panels as

$$R(\mathbf{x}_{it}, y_{it} + \mathbf{x}_{it}\mathbf{v}) = R(\mathbf{x}_{it}, y_{it}) + \mathbf{v} \tag{2.46}$$

for any $k \times 1$ vector $\mathbf{v}$.

Scale equivariance is given by

$$R(\mathbf{x}_{it}, cy_{it}) = cR(\mathbf{x}_{it}, y_{it}) \tag{2.47}$$

for any constant c.

Affine equivariance is given by

$$R(\mathbf{x}_{it}\mathbf{A}, y_{it}) = \mathbf{A}^{-1}R(\mathbf{x}_{it}, y_{it}) \tag{2.48}$$

for any $k \times k$ nonsingular matrix $\mathbf{A}$.

The classical WG estimator is regression, affine and scale equivariant. However, $\hat{\boldsymbol{\beta}}_{LTS}$ is only scale equivariant. This is due to the non-linearity of the centering transformation of the median, unlike that of the mean (see Bramati & Croux 2007).

2.3.4 The Within Groups Generalised M-Estimator

Bramati & Croux (2007) developed robust versions of the WG estimator. Since the median is highly robust to outliers, Bramati and Croux first centered the data around the median instead of the mean:

$$\tilde{y}_{it} = y_{it} - \underset{t}{\text{med}}\ y_{it} \tag{2.49}$$

$$\tilde{\mathbf{x}}_{it} = \mathbf{x}_{it} - (\underset{t}{\text{med}}\ x_{it}^{(1)}, \underset{t}{\text{med}}\ x_{it}^{(2)}, \ldots, \underset{t}{\text{med}}\ x_{it}^{(k)}) \tag{2.50}$$

After which $\tilde{y}_{it}$ is regressed on $\tilde{\mathbf{x}}_{it}$ using a robust estimator. Bramati and Croux initially used the LTS, adapted to the centered data:

$$\hat{\boldsymbol{\beta}}_{LTS} = \text{arg}\underset{\boldsymbol{\beta}}{min} \sum_{j=1}^{h} (\tilde{y}_j - \tilde{\mathbf{x}}_j\boldsymbol{\beta})^2_{j:NT} \tag{2.51}$$

where $(\tilde{y}_j - \tilde{\mathbf{x}}_j\boldsymbol{\beta})^2_{1:NT} \leq (\tilde{y}_j - \tilde{\mathbf{x}}_j\boldsymbol{\beta})^2_{2:NT} \leq \ldots \leq (\tilde{y}_j - \tilde{\mathbf{x}}_j\boldsymbol{\beta})^2_{NT:NT}$ are the NT ordered squared residuals. The LTS minimises the sum of the h smallest ordered squared residuals. When $h = NT$, we get the usual pooled OLS regression. larger values of h increases robustness, up to a maximum of but decreases efficiency. A common choice of h is $3/4NT$, which allows the estimator to attain a breakdown point of 25%.

To further improve on the statistical efficiency, while still maintaining robustness, in the same paper Bramati and Croux developed Within-Groups versions of the generalised M-estimator and the MS-estimator, called the Within-Groups Generalised M- (WGM) and Within-Groups MS- (WMS) estimators respectively.

The WGM estimator is based on the Mallows GM-estimator (2.30) reviewed in section 2.2.6. Adapted to the panel data context, the WGM is given by

$$\sum_{i=1}^{N}\sum_{t=1}^{T} w(\tilde{\mathbf{x}}_{it})\psi\left(\frac{r_{it}}{\hat{\sigma}_{LTS}}\right)\tilde{\mathbf{x}}_{it} = \mathbf{0} \tag{2.52}$$

where

$$\hat{\sigma}_{LTS} = c_{LTS}\frac{1}{h}\sum_{j=1}^{h} \left(\tilde{y}_j - \tilde{\mathbf{x}}_j\hat{\boldsymbol{\beta}}_{LTS}\right)^2_{j:NT} \tag{2.53}$$

is the preliminary LTS scale estimate of the data with c_{LTS} a constant correction factor for consistency under normal errors. $r_{it} \equiv \tilde{y}_{it} - \tilde{\mathbf{x}}_{it}\boldsymbol{\beta}_{LTS}$ are the residuals obtained from the preliminary LTS estimation of the data.

Tukey's Bisquare (2.28) was used as the influence function ψ of the standardised residuals, with $c = 4.685$ selected for normal errors, as in Wagenvoort & Waldmann (2001, see also section 2.2.6). For computation, a $NT \times NT$ diagonal matrix $W_r = \text{diag}((w_r)_{11}, (w_r)_{12}, \dots, (w_r)_{NT})$ is set up to downweigh residuals that are too large. The resulting diagonal entries of the weighting matrix are

$$(w_r)_{it} = \begin{cases} \left(1 - \left(\frac{r_{it}}{c\hat{\sigma}_{LTS}}\right)^2\right)^2 & \left|\frac{r_{it}}{\hat{\sigma}_{LTS}}\right| \leq c \\ 0 & \left|\frac{r_{it}}{\hat{\sigma}_{LTS}}\right| > c \end{cases} \tag{2.54}$$

The weights $w(\tilde{\mathbf{x}}_{it})$ for the regressors are compiled into a $NT \times NT$ diagonal matrix $W_x = \text{diag}(w(\tilde{\mathbf{x}}_{11}),$ $w(\tilde{\mathbf{x}}_{12}, \dots, w(\tilde{\mathbf{x}}_{NT}))$. The weight for each $\mathbf{x}_{it}$ is given by

$$w(\tilde{\mathbf{x}}_{it}) = \min\left(1, \frac{\sqrt{\chi^2_{p(0.975)}}}{\text{RD}_{it}}\right) \tag{2.55}$$

where RD_{it} is the robust version of the Mahalanobis distance (see Rousseeuw & Hubert 2011):

$$\text{RD}_{it} = \sqrt{(\tilde{\mathbf{x}}_{it} - \hat{\boldsymbol{\mu}})'\mathbf{V}^{-1}(\tilde{\mathbf{x}}_{it} - \hat{\boldsymbol{\mu}})} \tag{2.56}$$

where $\hat{\boldsymbol{\mu}}$ =mean of the h points of the dataset X such that the resulting covariance matrix $\mathbf{V}$ has the smallest determinant. Rouseeuw and Leroy (1987) showed that $h = [n + p + 1]/2$ yields the highest asymptotic breakdown point of 50%. $\hat{\boldsymbol{\mu}}$ and $\mathbf{V}$ are the LTS analogues of the empirical mean $\bar{\mathbf{x}} = \frac{1}{NT}\sum_{i=1}^{N}\sum_{t=1}^{T}\mathbf{x}_{it}$ and empirical covariance matrix $\mathbf{S}_x = \frac{1}{NT-p}\sum_{i=1}^{N}\sum_{t=1}^{T}(\mathbf{x}_{it} - \bar{\mathbf{x}})'(\mathbf{x}_{it} - \bar{\mathbf{x}})$, which are known to have a BP of $1/n$ and thus non-robust (see Rousseeuw & Hubert 2011).

The Robust Mahalanobis distance measures the euclidean distance of a datapoint $\mathbf{x}_{it}$ from the (geometric) centre of a tolerance ellipse(see Rousseeuw & Leroy 1987, p261). In this case, the tolerance is taken to be 97.5%, that is, all points with a distance greater than $\chi^2_{p(0.975)}$, the 97.5 percentile chi-squared statistic with p degrees of freedom, is considered an outlier to be downweighted according to (2.54).

The resulting WGM estimator is computed as

$$\hat{\boldsymbol{\beta}}_{WGM} = (\tilde{\mathbf{X}}' W_x W_r \tilde{\mathbf{X}})^{-1} \tilde{\mathbf{X}}' W_x W_r \tilde{\mathbf{y}} \tag{2.57}$$

2.3.5 The Within Groups MS Estimator

The median-centred data (2.49) and (2.50) can also be estimated by Maronna & Yohai's (2000) MS-estimator. In the same paper, Bramati & Croux (2007) set (based on the notation of the MS estimator reviewed in section 2.2.6)

$$\boldsymbol{\beta}_1 = \boldsymbol{\alpha}_i$$
$$\boldsymbol{\beta}_2 = \boldsymbol{\beta}$$

Assuming $\boldsymbol{\beta}$ is known, then (as in section 2.2.6) $\hat{\boldsymbol{\alpha}} = (\hat{\alpha}_1, \hat{\alpha}_2, \ldots, \hat{\alpha}_N)$ is an M-estimator

$$\hat{\boldsymbol{\alpha}}(\boldsymbol{\beta}) = \underset{\boldsymbol{\alpha}}{\operatorname{argmin}} \sum_{i=1}^{N} \sum_{t=1}^{T} \rho_M (y_{it} - \mathbf{x}'_{it}\boldsymbol{\beta} - \boldsymbol{\alpha}_i) \tag{2.58}$$

For quick and robust estimation, ρ_M was chosen such that $\hat{\alpha}_i(\boldsymbol{\beta}) = \underset{t}{\operatorname{median}}(y_{it} - \mathbf{x}'_{it}\boldsymbol{\beta})$ for each $i = 1, 2, \ldots, N$. The Within-Groups MS-estimator is therefore given by

$$\hat{\boldsymbol{\beta}}_{WMS} = \underset{\boldsymbol{\beta}}{\operatorname{argmin}}\ S\big(r_{11}(\hat{\boldsymbol{\alpha}}(\boldsymbol{\beta}), \boldsymbol{\beta}), r_{12}(\hat{\boldsymbol{\alpha}}(\boldsymbol{\beta}), \boldsymbol{\beta}), \ldots, r_{NT}(\hat{\boldsymbol{\alpha}}(\boldsymbol{\beta}), \boldsymbol{\beta})\big) \tag{2.59}$$

with

$$r_{it}(\hat{\boldsymbol{\alpha}}(\boldsymbol{\beta}), \boldsymbol{\beta}) = y_i t - \mathbf{x}'_{it}\boldsymbol{\beta} - \underset{t}{\operatorname{median}}(y_{it} - \mathbf{x}'_{it}\boldsymbol{\beta}) \tag{2.60}$$

Unlike the WGM estimator, the WMS is regression and affine equivariant, that is it is unchanged under linear transformations. Therefore, it is easier to modify the WMS than the WGM for other purposes, such as incorporating it into another estimation technique if neccessary.

Bramati & Croux (2007), in the same paper, showed that the cell breakdown points of the WGM and the WMS are equal to

$$\epsilon^*_{NT}(R, Z) = \frac{k[(T+1)/2] + \min(m - kT, [(T+1)/2])}{NT} \tag{2.61}$$

where $k = \left\lceil \frac{m}{[(T+1]/2} \right\rceil$ and $\frac{m}{NT}$ is the BP of the preliminary robust estimator. In this case, the preliminary estimator is the LTS), with $\frac{m}{NT} = \frac{1}{2}$. The asymptotic BP of the WGM and the WMS is 1/4. Indeed, this is a consequence of the "double-layered" structure of panel data. The preliminary estimator had the highest possible BP of 50%. This allows it to withstand up to half the centered data being outliers. Also, if more than half the data in one time series are outliers, the median centering breaks down. Therefore, to corrupt more than half of the time series, one only needs to corrupt over a quarter of the data $(\mathbf{x}_{it}, y_{it})$.

In the same paper, the block breakdown points of the WGM and the WMS estimator are also given by

$$\epsilon_N^B(R, Z) = \frac{\lceil m/T \rceil}{N} \tag{2.62}$$

where $\lceil \cdot \rceil$ is the ceiling operator, that is $\lceil x \rceil =$ the smallest integer that is larger than x. The authors also noted that the block BP of the WMS and WGM are identical to that of their preliminary LTS estimator. This follows naturally from the discussion of the cell breakdown points above: at least a quarter of the data contaminated is needed to contaminate half of the times series.

2.3.6 Simulation Studies

Bramati & Croux (2007) also conducted a simulation study showcasing the results of their estimators. For their simulation, they set up, according to (2.51)

$$\begin{aligned}
&\epsilon_{it} \sim N(0, 1) \\
&\alpha_i \sim U(0, 20) \\
&\boldsymbol{\beta} = \mathbf{0} \\
&\mathbf{x}_{it} \sim N(\mathbf{0}, \mathbf{I}) \quad \text{(multivariate standard normal)} \\
&\text{Number of regressors } p = 1
\end{aligned}$$

Contamination in the outliers is propagated randomly over the cells and a second time concentrated over a few time series, such that each of those time series have half of its data contaminated (known as block-concentrated outliers). Vertical outliers and leverage points are simulated separately, although over the same data points chosen for contamination.

The vertical outliers are generated by adding $N(50, 1)$ to each of the y_{it}'s to be contaminated. The

(a) MSE of the WG, WGM and WMS estimator under several sampling schemes for $N = 100, T = 4$ and levels of contamination 5% and 10%.

Sampling scheme	WG	WGM	WMS	WG	WGM	WMS
No outliers	0.004	0.003	0.004	0.004	0.003	0.004
		5% contamination			*10% contamination*	
Vertical outliers	0.395	0.004	0.004	0.936	0.004	0.004
Leverage points	16.933	0.004	0.003	20.260	0.006	0.005
Concentrated vertical outliers	0.273	0.004	0.010	0.640	0.004	0.009
Concentrated leverage points	14.818	0.003	0.004	18.885	0.003	0.004

(b) MSE of the WG, WGM and WMS estimator under several sampling schemes for $N = 100, T = 20$ and levels of contamination 5% and 10%.

Sampling scheme	WG	WGM	WMS	WG	WGM	WMS
No outliers	0.0005	0.0005	0.0007	0.0005	0.0005	0.0007
		5% contamination			*10% contamination*	
Vertical outliers	0.0700	0.0004	0.0007	0.1094	0.0006	0.0008
Leverage points	17.076	0.0007	0.0007	20.2340	0.0016	0.0009
Concentrated vertical outliers	0.0366	0.0006	0.0011	0.0690	0.0005	0.0021
Concentrated leverage points	13.175	0.0004	0.0008	17.6250	0.0005	0.0008

Table 1-3. Estimation results of the WG, WGM and WMS in the presence of outliers. Source: Bramati & Croux 2007

leverage points are generated by replacing the $\mathbf{x}_{it}$'s of the corresponding contaminated y_{it}'s with $N((10, 10, \ldots, 10)', \mathbf{I})$.

The percentages of contamination studied are 5% and 10%. The cross section size is set to $N = 100$, and the time dimension is set to $T = 4$ and $T = 20$ to simulate small and large time dimensions respectively. For each estimator (the classical within-groups (WG), the WGM and the WMS), 5 levels of contamination (no contamination, vertical outliers only, leverage points only, concentrated vertical outliers only, concentrated leverage points only), $M = 1000$ simulations are run for each level of contamination. The simulation results are presented in Table 1-3.

The performance of each estimator is measures by the mean squared error (MSE), defined as

$$\frac{1}{M}\sum_{j=1}^{M} \|\hat{\boldsymbol{\beta}}^(j) - \boldsymbol{\beta}\|^2 \qquad (2.63)$$

where $\hat{\boldsymbol{\beta}}^(j)$ is the estimated value of $\boldsymbol{\beta}$ for each simulation j of M of each category. Geometrically, the MSE is the average squared distance of the estimate from its true value. Thus, the higher the MSE, the more biased the estimate is.

From Table 1-3, we see that the WGM and WMS perform very similarly to the classical WG estimator when there is no contamination. As expected, the classical WG becomes very bad under contamination, particularly under leverage points, as evidenced by the large MSE values. In the WG estimator, we see the mean applied in two main stages (the centering and the regression). Each stage adds a further bias to the WG estimator. Thus, we expect the effects of outliers to be more severe in panel data compared to pure cross-section or pure time-series regressions.

On the other hand, the WGM and WMS maintained their MSE under both levels of contamination. We see that, like the *Pilot-Plant* data, the WGM and WMS have successfully resisted the outliers, demonstrating the robustness of the two estimators.

Other notable robust static panel data estimators can be found in the works of Wagenvoort & Waldmann (2001, of which much of the work of Bramati & Croux (2007) is based upon) Dhane & Zhu (2009) and Aquaro & Cizek (2013).

We now turn our attention to cross-sectional dependence in panel data, and related estimators. In the following section, we will review the topic of cross-sectional dependence in panel models and an appropriate estimator that follows. We will also discuss the effects of outliers on such estimators, and also discuss possible robust versions of these estimators.

2.4 Conclusion

We have reviewed the existing robust literature. The next chapter reviews the cross-section dependence literature and explores the effects of outliers on the CCE estimators.

References

Aggarwal, C.C. (2017) *Outlier analysis.* 2nd edn. Cham: Springer.

Barnett, V. and Lewis, T. (1994) *Outliers in statistical data.* Chichester: Wiley.

Beckman, R.J. and Cook, R.D. (1983) "Outlier..........s," *Technometrics*, 25(2), p. 119. Available at: https://doi.org/10.2307/1268541.

Ben-Gal, I. (2009) "Outlier detection," *Data Mining and Knowledge Discovery Handbook*, pp. 117–130. Available at: https://doi.org/10.1007/978-0-387-09823-47.

Bramati, M.C. and Croux, C. (2007) "Robust estimators for the fixed effects panel data model," *The Econometrics Journal*, 10(3), pp. 521–540. Available at: https://doi.org/10.1111/j.1368-423x.2007.00220.x.

Chauvenet, W. (1863) *A manual of spherical and practical astronomy: Embracing the general problems of spherical astronomy, the special applications to nautical astronomy, and the theory and use of fixed and portable astronomical instruments: With an appendix on the method of least sqares.* Philadelphia, London: (Ph.).

Grubbs, F.E. (1950) "Sample criteria for testing outlying observations," *The Annals of Mathematical Statistics*, 21(1), pp. 27–58. Available at: https://doi.org/10.1214/aoms/1177729885.

Hansen, B. (2022) *Econometrics.* Princeton, NJ: Princeton University Press.

Hawkins, D.M. (1980) "Identification of outliers." Available at: https://doi.org/10.1007/978-94-015-3994-4.

Huber, P.J. (1981) *Robust statistics.* New York: Wiley.

Huber, P.J. and Ronchetti, E.M. (2011) *Robust statistics.* Hoboken, NJ: Wiley.

Kendall, M.G. *et al.* (1961) "Studies in the history of probability and statistics: XI. Daniel Bernoulli on maximum likelihood," *Biometrika*, 48(1/2), p. 1. Available at: https://doi.org/10.2307/2333125.

Leys, C. *et al.* (2018) "Detecting multivariate outliers: Use a robust variant of the Mahalanobis distance," *Journal of Experimental Social Psychology*, 74, pp. 150–156. Available at: https://doi.org/10.1016/j.jesp.20

Mahalanobis, P. (1936) *On the generalized distance in statistics*, *Semantic Scholar.* Available at: https://www.semanticscholar.org/paper/On-the-generalized-distance-in-statistics-Mahalanobis/fa97c2238a16e922

(Accessed: December 21, 2022).

Mahalanobis, P.C. (1936) "On the Generalized Distance in Statistics," Proceedings of the National Institute of Science of India, 2, 49-55.

Maronna, R.A., Martin, R.D., Yohai, V.J. and Salibián-Barrera, M., 2019 *Robust statistics theory and methods (with R)*. 2nd edn. Hoboken, NJ: Wiley.

Peirce, B. (1852) "Criterion for the rejection of Doubtful Observations," *The Astronomical Journal*, 2, p. 161. Available at: https://doi.org/10.1086/100259.

Ross, S. (2003) "Peirce's criterion for the elimination of suspect experimental data," J. Eng. Technol.. 20.

Rousseeuw, P.J. (1984) "Least median of squares regression," *Journal of the American Statistical Association*, 79(388), pp. 871–880. Available at: https://doi.org/10.1080/01621459.1984.10477105.

Rousseeuw, P.J. and Leroy, A.M. (1987) *Robust regression and outlier detection.* Hoboken, NJ: Wiley-Interscience.

Stigler, S.M. (1973) "Simon Newcomb, Percy Daniell, and the history of robust estimation 1885-1920," *Journal of the American Statistical Association*, 68(344), p. 872. Available at: https://doi.org/10.2307/2284515.

Wilcox, R. (2022) *Introduction to robust estimation and hypothesis testing.* London: Academic Press, an imprint of Elsevier.

Zaman, A., Rousseeuw, P.J. and Orhan, M. (2001) "Econometric applications of high-breakdown robust regression techniques," *Economics Letters*, 71(1), pp. 1–8. Available at: https://doi.org/10.1016/s0165-1765(00)00404-3.

Chapter 3

Cross-sectional Dependence and the Common Correlated Effects - A Review

3.1 Cross-Sectional Dependence and Estimation

Cross-Sectional Dependence in Panels

Traditional panel data models assume that the cross-section units are independent. Much of the literature in panel data assumes no correlation between cross-sectional units. However, the notion of dependence between groups of individuals and entities in the social sciences has been familiar as early as the 1930s (see Sarafidis & Wansbeek, 2012). As quoted by Stephan:

> "... in dealing with social data, we know that by virtue of their very social character, persons, groups and their characteristics are interrelated and not independent"
> (Stephan 1934, quoted by Sarafidis & Wansbeek 2012)

Thus, it is an inescapable fact that cross-section dependence is very real and present in data. Given the increasing availability of data across regions, countries or even industries, we have not only the traditional micro-panles with small cross-section dimensions N and small time dimensions T, but macropanels with large N and T with much more opportunity for cross-section dependence to appear. Thus, cross-sectional correlation appears to be the norm rather than the exception (Chudik & Pesaran 2013). The consequences of ignoring cross-section dependence ca have serious consequences (Chudik & Pesaran, 2013). Cross-section dependence may arise due to omitted common effects, spatial effects or even interactions among socioeconomic networks. This results in misleading inference and even inconsistent estimators, depending on the extent of cross-section dependence. Also, since many unit root tests assume independence, cross-section dependence may result in size distortions.

Modelling cross-sectional dependence is not as easy as for time series. Unlike the latter, in which there is a natural flow of time, interactions between social entities are more varied in nature and structure. The earliest methodology dealing with cross-sectional dependence is the spatial approach. It involves the formulation of a distance metric to measure the extent of cross-sectional dependence between units. This methodology has applications in fields such as regional science and urban economics, which use the concept of economic distance (Sarafidis & Wansbeek 2012). An alternative approach to modelling CSD is the factor model. Chudik & Pesaran (2013) provide a comprehensive review focusing on factor models and the corresponding tests and estimators involved. The literature has extended since, and the focus of this paper is to extend the review to cover the current advances in the literature.

3.1.1 Outline

In this paper, we review the literature regarding factor models and the associated Common Correlated Effects (CCE) estimators (Pesaran, 2006). The following section starts with summarising the foundations in Chudik & Pesaran (2013) with additions from others in the literature, such as Pesaran (2006), Chudik, Pesaran & Tosetti (2011), Sarafidis & Wansbeek (2012) and a few others. The next section builds on the review with current additions to the literature, including software packages (Ditzen, 2018) written for the CCE estimators. The next section demonstrates the CCE estimators with an economic application and the final section concludes.

3.2 Cross-section Dependence, Factor Models and the CCE Estimators

This section covers the main points in Chudik & Pesaran (2013), with additions from the literature. This section starts out with reviewing models of cross-section dependence, followed by the relevant estimation procedures (the CCE estimators) and lastly with tests for cross-section dependence.

3.2.1 Modelling Cross-section Dependence

Consider a double-indexed process $\{z_{it}|i \in \mathbb{N}, t \in \mathbb{Z}\}$ with z_{it} defined on a suitable probability space with t being ordered (for example, time) and i an indexing set representing units from an unordered population. The process z_{it} fulfills the following assumptions:

CSD 1. *For each $t \in \mathbb{Z}$, $\mathbf{z}_t = (z_{1,t}, z_{2,t}, \ldots, z_{Nt})'$ has mean $E(\mathbf{z}_t = 0$ and variance $Var(\mathbf{z}_t) = \mathbf{\Sigma}_t$, an $N \times N$ symmetric, non-negative definite matrix. The (i, j)-th term of $\mathbf{\Sigma}_t$, $\sigma_{ij,t}$, satisfies $0 \leq \sigma_{ij,t} \leq \sigma_{ii,t} \leq K < \infty$, where the constant K is independent of N, for all $i = 1, 2, \ldots, N$*

Alternatively, the mean $E(\mathbf{z}_t)$ can be conditioned on an information set $\mathbf{\Omega}_{t-1}$ and the zero means condition can be relaxed:

$$\begin{aligned} E(\mathbf{z}_t) &= \boldsymbol{\mu} \quad \text{or} \\ E(\mathbf{z}_t|\mathbf{\Omega}_{t-1}) &= \boldsymbol{\mu} \end{aligned} \tag{3.1}$$

There are a few ways to measure the extent of cross-section dependence, such as by the largest eigenvalue of $\mathbf{\Sigma}_t$, $\lambda(\mathbf{\Sigma}_t)$; by matrix norms of $\mathbf{\Sigma}_t$; and more commonly in the literature by the behaviour of weighted cross-section averages

$$\begin{aligned} \bar{\mathbf{z}}_{wt} &= \sum_{i=1}^{N} w_{it}\mathbf{z}_{it} \\ &= \mathbf{w}_t'\mathbf{z}_t \end{aligned} \tag{3.2}$$

Here, the process $\mathbf{z}_t$ fulfils assumption (CSD 1) and $\mathbf{w}_t$ are weights fulfilling the following assumption:

CSD 2. *Let $\mathbf{w}_t = (w_{1t}w_{2t}\dots w_{Nt})'$ each $t \in T \subseteq \mathbb{Z}$ and $N \in \mathbb{N}$ be a vector of stochastic weights (or a vector of randomly distributed weights independent of $\mathbf{z}_t$). For any $t \in T$, the sequence of vectors $\{\mathbf{w}_t\}_{N=1}^{\infty}$ satisfies:*

$$\text{i)} \quad \begin{aligned} \|\mathbf{w}_t\| &= \sqrt{\mathbf{w}_t'\mathbf{w}_t} \\ &= O\left(\frac{1}{\sqrt{N}}\right) \end{aligned} \tag{3.3}$$

$$\text{ii)} \quad \frac{w_{jt}}{\|\mathbf{w}_t\|} = O\left(\frac{1}{\sqrt{N}}\right) \text{ uniformly in } j \in \mathbb{N} \tag{3.4}$$

(3.3) and (3.4) of assumption (CSD 2) are also known in finance as the granularity conditions. Chudik et. al (2011) defined the concepts of weak and strong cross-section dependence based on the limiting behaviour of $\bar{\mathbf{z}}_{wt}$ for each $t \in T$:

Definition 3.2.1. (Weak Cross-section Dependence) The process $\{\mathbf{w}_t\}$ is cross-sectionally weakly dependent (CWD) at a given $t \in T$ if for any sequence of vectors $\{\mathbf{w}\}_t$ satisfying the granularity conditions, we have

$$Var(\mathbf{w}_t'\mathbf{z}_t) \to 0 \text{ as } N \to \infty \tag{3.5}$$

Definition 3.2.2. (Strong Cross-section Dependence) The process $\{\mathbf{w}_t\}$ is cross-sectionally strongly dependent (CSD) if there exists a sequence of vectors satisfying the granularity conditions and a constant $K < \infty$ and independent of N such that for any sufficiently large N (in particular, as

$N \to \infty$),

$$Var(\mathbf{w}_t'\mathbf{z}_t) \geq K \geq 0 \tag{3.6}$$

The above definitions can be applied to variances conditioned on information sets just like assumptions (CSD 1) and (CSD 2). There is a relationship between weak and strong cross-section dependence and the asymptotic behaviour of the largest eigenvalue of $\mathbf{\Sigma}_t$, $\lambda(\mathbf{\Sigma}_t)$:

Proposition 3.2.1. *The process z_{it} is CWD at a point in time $t \in T$ if $\lambda(\mathbf{\Sigma}_t$ is bounded in N or increases at a rate slower than N.*

Proposition 3.2.2. *The process z_{it} is CSD at a point in time $t \in T$ if and only if for any N sufficiently large (and as $N \to \infty$, $\lambda(\mathbf{\Sigma}_t)/N \geq K > 0$.*

Proofs are given in Chudik & Pesaran (2013) and Cudik, Pesaran & Tosetti (2011). Of interest in the literature (and in the scope of this paper) are common factor models, a strong form of cross-section dependence.

3.2.2 The Common Factor Model

Consider the m-factor model for z_{it}

$$z_{it} = \gamma_{i1} f_{1t} + \gamma_{i2} f_{2t} + \ldots + \gamma_{im} f_{mt} + e_{it} \tag{3.7}$$

or compactly,

$$\mathbf{z}_t = \mathbf{\Gamma}\mathbf{f}_t + \mathbf{e}_t \tag{3.8}$$

where $\mathbf{f}_t$ is the $m \times 1$ vector of common factors, $\boldsymbol{\gamma}_i = (\gamma_{i1}\gamma_{i2} \ldots \gamma_{im})'$ is the corresponding vector of factor loadings and e_{it} idiosyncratic errors. The common factors affect all cross-section units to varying degrees (via the factor loadings). In the common factor model, correlation between any two cross-section units does not depend on distance or a distance metric, unlike the spatial models (In fact, spatial models can be a special case of factor models). Chudik & Pesaran (2013) gives examples of common effects: observed common factors that affect all households' and firms' consumption and investment decisions include interest rates and oil prices. Unobserved factors; unobserved effecs include aggregate demand and supply shocks. In multifactor models, interdependence often arises from the reaction of units to some external events.

The common factors and errors of z_{it} typically fulfill the following assumptions:

CF 1. *The* $m \times 1$ *vector* $\mathbf{f}_t$ *is a zero-mean covariance stationary process with absolutely summable autocovariances*

$\mathbf{f}_t$ *is distributed independently of* e_{is} *for all* i, t, s

$E(f_{lt}^2) = 1)$ *and* $E(f_{lt} f_{l's}) = 0$ *for all* $l, l', t, s \neq t$

CF 2. $Var(e_{it}) = \sigma_i^2 < K < \infty$, e_{it} *is mean-zero and* $E(e_{it} e_{is}) = 0$ *for all* $s \neq t$

Assumption (CF 1) is an identification condition and typically, it is not possible to separate $\mathbf{f}_t$ from $\mathbf{\Gamma}$. A model fufilling (CF 1) and (CF 2) is also known as an exact factor model. Under these assumptions, the covariance of $\mathbf{z}_t$ is given by

$$E(\mathbf{z}_t \mathbf{z}_t') = \mathbf{\Gamma}\mathbf{\Gamma}' + \mathbf{V} \tag{3.9}$$

where $\mathbf{V} = diag(\sigma_1^2, \sigma_2^2, \ldots, \sigma_N^2)$. Besides the exact factor model, the assumption that the idiosyncratic errors e_{it} are cross-sectionally independent can be relaxed to weakly cross-correlated ones. In this case, the factor model is known as an approximate factor model. The correlation patterns in the approximate factor model can be characterised by

$$\mathbf{e}_t \quad = \quad \mathbf{R}\boldsymbol{\epsilon}_t \tag{3.10}$$

with $\boldsymbol{\epsilon}_t = (\epsilon_{1t}\epsilon_{2t} \ldots \epsilon_{Nt})' \ \sim \ (\mathbf{0}, \mathbf{I}_N)$. In this case, $\mathbf{V} = \mathbf{R}\mathbf{R}'$, which is not diagonal if $\mathbf{R}$ is not diagonal. To distinguish the the factor specification from the cross-section dependence in the idiosyncratic errors, one may assume CWD in the idiosyncratic errors, for example, that $\mathbf{R}$ has bounded row and column sum matrix norms. Also, $\lim_{N \to \infty} N^{-1} \mathbf{\Gamma}\mathbf{\Gamma}$ is assumed to be a full-rank matrix.

Factors can be classified as weak or strong, depending on their convergence:

Definition 3.2.3. (Strong and Weak factors)

i) The factor f_{lt} is strong if $\lim_{N \to \infty} \frac{1}{N} \sum_{i=1}^{N} |\gamma_{it}| = K > 0$.

ii) The factor f_{lt} is weak if $\lim_{N \to \infty} \sum_{i=1}^{N} |\gamma_{it}| = K < \infty$

Given $\alpha_l \in (0, 1)$, the factor f_{lt} is said to be semi-strong (or semi-weak) if

$$\lim_{N \to \infty} \frac{1}{N^{\alpha_l}} \sum_{i=1}^{N} |\gamma_{it}| = K < \infty \tag{3.11}$$

It is useful to classify factors with $0 < \alpha_l < 1/2$ as semi-weak (and factors with $1/2 \leq \alpha_l < 1$ as semi-strong). In a multi-factor setting, the overall exponent α of cross-section dependence can be defined as $\alpha = \max(\alpha_1, \alpha_2, \ldots, \alpha_m)$. Examples of α in the literature include:

- large macro datasets (Bailey, Kapertanios & Pesaran 2012) where α ranges from 0.77 to 0.92.
- cross-country quarterly GDP, inflation, real equity prices, where $\alpha \approx 0.97$

Strong factors can be used to represent the effect of cross-section units that are dominant or pervasive, and do not vanish as $N \to \infty$. For example, a large city may play a dominant role in house prices nationally (Holly, Pesaran & Yamagata, 2011; Chudik & Pesaran, 2013).
Strong and weak factors can determine if the process z_{it} is CSD or CWD in the following theorem (see Chudik & Pesaran, 2013):

Theorem 3.2.1. *Consider the model (3.7) where assumptions (CF 1) and (CF 2) hold. Suppose further that there exists* $\alpha = \max(\alpha_1, \alpha_2, \ldots, \alpha_m)$ *for* $\alpha, \alpha_l \in [0, 1]$ *such that (3.11) is fulfilled for all* $l = 1, 2, \ldots, m$*. Then the following hold:*
i) $\{z_{it}\}$ *is CWD at* $t \in T$ *if* $\alpha < 1$ *(That is, the process* $\{z_{it}\}$ *is CWD if there are only weak, semi-weak or semi-strong factors in it).*
ii) $\{z_{it}\}$ *is CSD if and only if there exists at least one strong factor.*

Proofs are provided in Chudik, Pesaran & Tosetti (2011). Unlike strong factors, weak or semi-strong factors do not affect consistency, but does affect inference (see Chudik & Pesaran 2013 for examples).

3.2.3 The Common Correlated Effects Estimators

Consider the heterogeneous panel data model

$$y_{it} = \alpha_i' \mathbf{d}_t + \beta_i' \mathbf{x}_{it} + e_{it} \tag{3.12}$$

$$e_{it} = \gamma_i' \mathbf{f}_t + \epsilon_{it} \tag{3.13}$$

where $\mathbf{d}_t$ is a $n \times 1$ vector of observed common effects, including deterministics such as intercepts and seasonal dummies.
$\mathbf{x}_{it}$ is a $k \times 1$ vector of observed individual-specific regressors on the ith cross-section at time t.
$\boldsymbol{\gamma}_i = (\gamma_{1i}\gamma_{2i} \ldots \gamma_{mi})'$ is the associated $m \times 1$ vector of factor loadings. The number of factor loadings is fixed and it is assumed that $m << N$.
The idiosyncratic errors e_{it} are allowed to be CWD, for example, being generated by a spatial process or having a weak factor structure. The individiual specific errors assumed to be independent of $\mathbf{d}_t$ and $\mathbf{x}_{it}$.

We distinguish between the homogeneous slope case $\boldsymbol{\beta}_i = \boldsymbol{\beta}\ \forall i$ and the heterogeneous case $E(\boldsymbol{\beta}_i) = \boldsymbol{\beta}$. For the scope of this paper (and the CCE estimators), we assume $\boldsymbol{\beta}_i = \boldsymbol{\beta} + \boldsymbol{\nu}_i$, with $\boldsymbol{\nu}_i \sim \text{iid.}(\mathbf{0}, \boldsymbol{\Omega}_{\boldsymbol{\nu}})$ in the heterogeneous case.

When the $\mathbf{x}_{it}$ are strictly exogeneous and $\boldsymbol{\nu}_i$ are independent of the regressors and errors, the slopes $\boldsymbol{\beta}_i$ can be consistently estimated by mean group (MG) estimators and pooled estimators (see Chudik & Pesaran (2013); Coakeley, Fuertes & Smith (2006); Eberhardt (2012); Bai (2009) for more details on MG estimators). However, if the $\mathbf{x}_{it}$ are weakly exogeneous, only MG estimators are consistent (see Pesaran & Smith (1995) for more details).

For the CCE estimators (Pesaran, 2006), additionally the factors $\mathbf{f}_t$ are assumed to be correlated to $\mathbf{d}_t$ and $\mathbf{x}_{it}$ by

$$\mathbf{x}_{it} = \mathbf{A}_i'\mathbf{d}_t + \boldsymbol{\Gamma}_i'\mathbf{f}_t + \mathbf{v}_{it} \tag{3.14}$$

with $\mathbf{A}_i$ and $\boldsymbol{\Gamma}_i$ as $n \times k$ and $m \times k$ factor loading matrices respectively with fixed components. d_t, $\mathbf{f}_t$ are assumed to be covariance stationary. (3.12), (3.13) and (3.14) combine to give the overall factor model

$$\begin{aligned} y_{it} &= \boldsymbol{\alpha}_i'\mathbf{d}_t + \boldsymbol{\beta}_i'\mathbf{x}_{it} + \boldsymbol{\gamma}_i'\mathbf{f}_t + \epsilon_{it} \\ \mathbf{x}_{it} &= \mathbf{A}_i'\mathbf{d}_t + \boldsymbol{\Gamma}_i'\mathbf{f}_t + \mathbf{v}_{it} \end{aligned} \tag{3.15}$$

The CCE estimators approximate the linear combinations of the unobserved factors by cross-section averages of the dependent and explanatory variables, and then runs standard panel regressions augmented with those cross-section averages. As shown in Pesaran (2006); Chudik & Pesaran (2013) and Chudik, Pesaran & Tosetti (2011), the overall factor model (3.15) can be re-written in combined form

$$\mathbf{z}_{it} = \begin{pmatrix} y_{it} \\ \mathbf{x}_{it} \end{pmatrix} = \mathbf{B}_i'\mathbf{d}_t + \mathbf{C}_i'\mathbf{f}_t + \boldsymbol{\xi}_{it} \tag{3.16}$$

where

$$\begin{aligned} \boldsymbol{\xi}_{it} &= \begin{pmatrix} e_{it} + \boldsymbol{\beta}_i'\mathbf{v}_{it} \\ \mathbf{v}_{it} \end{pmatrix}, \\ \mathbf{B}_i &= \begin{pmatrix} \boldsymbol{\alpha}_i & \mathbf{A}_i \end{pmatrix} \begin{pmatrix} \mathbf{1} & \mathbf{0} \\ \boldsymbol{\beta}_i & \mathbf{I}_k \end{pmatrix}, \quad \mathbf{C}_i = \begin{pmatrix} \boldsymbol{\gamma}_i & \boldsymbol{\Gamma}_i \end{pmatrix} \begin{pmatrix} \mathbf{1} & \mathbf{0} \\ \boldsymbol{\beta}_i & \mathbf{I}_k \end{pmatrix} \end{aligned}$$

The weighted averages obtained from the weights w_i satisfying (3.3) and (3.4) are given by

$$\bar{\mathbf{z}}_{wt} = \bar{\mathbf{B}}'_w \mathbf{d}_t + \bar{\mathbf{C}}'_w \mathbf{f}_t + \bar{\boldsymbol{\xi}}_{wt} \tag{3.17}$$

where

$$\begin{aligned} \bar{\mathbf{z}}_{wt} &= \sum_{i=1}^{N} w_i \mathbf{z}_{it}, \\ \bar{\mathbf{B}}_w &= \sum_{i=1}^{N} w_i \mathbf{B}_i, \quad \bar{\mathbf{C}}_w = \sum_{i=1}^{N} w_i \mathbf{C}_i \quad \text{and} \\ \bar{\boldsymbol{\xi}}_{wt} &= \sum_{i=1}^{N} w_i \boldsymbol{\xi}_{it} \end{aligned}$$

If the rank condition

$$\text{rank}(\bar{\mathbf{C}}_w) = m \leq k + 1 \tag{3.18}$$

is fulfilled, we have

$$\mathbf{f}_t = (\bar{\mathbf{C}}_w \bar{\mathbf{C}}'{}_w)^{-1} \bar{\mathbf{C}}_w (\bar{\mathbf{z}}_{wt} - \bar{\mathbf{B}}_w \mathbf{d}_t) \tag{3.19}$$

If e_{it} and $\mathbf{v}_{it}$ are CWD processes, Pesaran and Tosetti (2011) showed that

$$\boldsymbol{\xi}_{wt} \xrightarrow{q.m.} \mathbf{0} \quad \text{as} \quad N \to \infty \tag{3.20}$$

which results in

$$\mathbf{f}_t - (\bar{\mathbf{C}}_w \bar{\mathbf{C}}'{}_w)^{-1} \bar{\mathbf{C}}_w (\bar{\mathbf{z}}_{wt} - \bar{\mathbf{B}}_w \mathbf{d}_t) \xrightarrow{q.m.} \mathbf{0} \quad \text{as} \quad N \to \infty \tag{3.21}$$

Thus, $\mathbf{f}_t$ can be well-approximated by a linear combination of the observed effects $\mathbf{d}_t$ and the cross-section averages of the dependent variable $\bar{y}_{wt}$ and the individual-specific regressors $\bar{\mathbf{x}}_{wt}$. Define

$$\bar{\mathbf{M}}_w = \mathbf{I}_T - \bar{\mathbf{H}}_w (\bar{\mathbf{H}}'_w \bar{\mathbf{H}}_w)^{-} \bar{\mathbf{H}}'_w \tag{3.22}$$

Where $\bar{\mathbf{H}}_w = (\mathbf{D}, \bar{\mathbf{Z}}_w)$, $\mathbf{D} = (\mathbf{d}_1, \mathbf{d}_2, \ldots, \mathbf{d}_T)'$ and $\bar{\mathbf{Z}}_w = (\bar{\mathbf{z}}_{w1}, \bar{\mathbf{z}}_{w2}, \ldots, \bar{\mathbf{z}}_{wT})'$. $(\bar{\mathbf{H}}'_w \bar{\mathbf{H}}_w)^{-}$ is the generalised inverse of the matrix $(\bar{\mathbf{H}}'_w \bar{\mathbf{H}}_w)$. There are two main versions of the CCE first introduced in the same paper (Pesaran 2006). The CCE Mean Group (CCEMG) estimator first estimates the slope in each individual cross section and then takes the mean of all the estimates as the CCEMG estimator. The CCE Pooled (CCEP), as its name indicates, pools all observations in a single estimation. The CCEP estimator gains efficiency when all the individual slopes are identical. In the CCEMG estimator, the individual-specific estimates are given by

$$\hat{\mathbf{b}}_i = (\mathbf{X}_i'\bar{\mathbf{M}}_w\mathbf{X}_i)^{-1}(\mathbf{X}_i'\bar{\mathbf{M}}_w\mathbf{y}_i) \tag{3.23}$$

And the CCEMG estimator itself is given by

$$\hat{\mathbf{b}}_{MG} = \frac{1}{N}\sum_{i=1}^{N}\hat{\mathbf{b}}_i \tag{3.24}$$

The CCEP is given by

$$\hat{\mathbf{b}}_P = \left(\sum_{i=1}^{N}\theta_i\mathbf{X}_i'\bar{\mathbf{M}}_w\mathbf{X}_i\right)^{-1}\left(\sum_{i=1}^{N}\theta_i\mathbf{X}_i'\bar{\mathbf{M}}_w\mathbf{y}_i\right) \tag{3.25}$$

As in Pesaran's (2006) paper, the weights θ_i are typically set to $\frac{1}{N}$, but if the individual variances σ_i vary, it is optimal to set $\theta_i = \frac{\sigma_i^{-2}}{\sum_{j=1}^{N}\sigma_j^{-2}}$, although the efficiency gains are small especially with small T.

The advantage of the CCE estimators is that they yield consistent estimates under a wide variety of conditions, even when the rank condition (3.18) is not fulfilled (see Pesaran, 2006; Chudik & Pesaran, 2013). Another advantage of the CCE estimators is that the number of factors need not be known beforehand.

Dynamic CCE

When strict exogeneity is relaxed, cross-section dependence becomes more complicated (Chudik & Pesaran, 2013). The dynamic CCE estimates the dynamic model with a lagged dependent variable

$$y_{it} = \lambda_i y_{i,t-1} + \boldsymbol{\beta}_i'\mathbf{x}_{it} + u_{it} \tag{3.26}$$

$$u_{it} = \boldsymbol{\gamma}_i'\mathbf{f}_t \tag{3.27}$$

Let

$$\boldsymbol{\zeta}_{it} = \begin{pmatrix} y_{i,t-1} \\ \mathbf{x}_{it} \end{pmatrix}, \quad \boldsymbol{\pi}_i = \begin{pmatrix} \lambda_i \\ \boldsymbol{\beta}_i \end{pmatrix}$$

and rewrite the dynamic model (3.26) as

$$y_{it} = \boldsymbol{\pi}_i' \boldsymbol{\zeta}_{it} + u_{it} \tag{3.28}$$

As in the static factor model (3.7), we distinguish between the heterogeneous case and the homogeneous case $\boldsymbol{\pi}_i = \boldsymbol{\pi}$ for all i. Our focus in this paper will be on the static factor model, and we would like to consider the dynamic CCE in future research.

3.2.4 Testing for Cross-Section Dependence

The CD-test (Pesaran, 2015) tests a Panel data model for strong and weak cross-section dependence. The CD-statistic is given by

$$CD = \sqrt{\frac{2T}{N(N-1)}} \left(\sum_{i=1}^{N-1} \sum_{j=i+1}^{N} \hat{\rho}_{ij} \right) \tag{3.29}$$

where $\hat{\rho}_{ij}$ are the estimated cross-correlations between the residuals $\hat{e}_{ij}$ of a preliminary estimator:

$$\hat{\rho}_{ij} = \hat{\rho}_{ji} = \frac{\sum_{t=1}^{T} \hat{e}_{it} \hat{e}_{jt}}{\sqrt{\sum_{t=1}^{T} \hat{e}_{it}^2 \sum_{t=1}^{T} \hat{e}_{it}^2}} \tag{3.30}$$

The preliminatry estimator can be least squares or a within groups estimator. Software for computing this test is available on Stata under the *xtcd2* package (Ditzen, 2018).

3.3 Effect of Outliers on the CCE estimators

We now examine the effects of the CCE estimators under the influence of outliers. We perform a simulation study using the original factor model in Experiment 1A of Pesaran (2006). The factor model is given by

$$y_{it} = \alpha_{1i} d_{1t} + \beta_{1i} x_{1it} + \beta_{2i} x_{2it} + \gamma_{1i} f_{1t} + \gamma_{2i} f_{2t} + \epsilon_{it} \tag{3.31}$$

$$x_{jit} = a_{1ji} d_{1t} + a_{2ji} d_{2t} + \gamma_{1ji} f_{1t} + \gamma_{3ji} f_{3t} + v_{jit} \qquad j = 1, 2 \tag{3.32}$$

For each $i = 1, 2, \ldots, N$, the common effects are modelled as

$$\begin{aligned}
&d_{1t} = 1 \\
&d_{2t} = 0.5 d_{2,t-1} + \epsilon_{dt} \qquad \epsilon_{dt} \sim \text{iid.} N(0, (1-0.5)^2) \\
&f_{jt} = 0.5 f_{j,t-1} + \epsilon_{fjt} \qquad \epsilon_{fjt} \sim \text{iid.} N(0, (1-0.5)^2) \qquad j = 1, 2, 3
\end{aligned} \tag{3.33}$$

The individual specific errors v_{jit} of x_{jit} are given by

$$\begin{aligned}
&v_{jit} = \rho_{vj} v_{ji,t-1} + \epsilon_{vjt} \\
&\epsilon_{vjt} \sim \text{iid.} N(0, 1 - \rho_{vj}^2) \\
&\rho_{vj} \sim \text{iid.} U(0.05, 0.95)
\end{aligned} \tag{3.34}$$

The factor loadings of the common observed effects $\mathbf{d}_t$ are generated as

$$\begin{aligned}
&\alpha_{1i} \sim \text{iid.} N(1, 1) \\
&a_{jli} \sim \text{iid.} N(0.5, 0.5) \qquad \text{for all combinations } j, l = 1, 2
\end{aligned} \tag{3.35}$$

The factor loadings of the unobserved effects $\mathbf{f}_t$ are generated as

$$\begin{gathered}
\gamma_{ji} \sim \text{iid.} N(1, 0.2) \qquad j = 1, 2 \\
\begin{pmatrix} \gamma_{11i} & \gamma_{12i} & \gamma_{13i} \\ \gamma_{21i} & \gamma_{22i} & \gamma_{23i} \end{pmatrix} \sim \text{iid.} \begin{pmatrix} N(0.5, 0.5) & 0 & N(0, 0.5) \\ N(0, 0.5) & 0 & N(0.5, 0.5) \end{pmatrix}
\end{gathered} \tag{3.36}$$

The individual-specific errors of y_{it} are generated as stationary AR(1) processes for $i = 1, 2, \ldots, N/2$ and as MA(1) processes for the rest (lines 62-70):

$$\begin{aligned}
&\varepsilon_{it} = \rho_{i\varepsilon} \varepsilon_{i,t-1} + \sigma_i (1 - \rho_{i\varepsilon}^2)^{1/2} \varsigma_{it} \qquad i = 1, 2, \ldots, N/2 \\
&\varepsilon_{it} = \sigma_i (1 + \theta_{i\varepsilon}^2)^{-1/2} (\varsigma_{it} + \theta_{i\varepsilon} \varsigma_{i,t-1}) \qquad i = N/2+1, N/2+2, \ldots, N \\
&\sigma_i^2 \sim \text{iid} U(0.5, 1.5) \\
&\rho_{i\varepsilon} \sim \text{iid} U(0.05, 0.95) \\
&\theta_{i\varepsilon} \sim \text{iid} U(0, 1)
\end{aligned} \tag{3.37}$$

3.4 Simulation Studies

We perturb the data-generating processes (DGPs) with both leverage points and vertical outliers. We use the above specification for each experiment. We break our Experiment into 4 sub-experiments 1A, 1B, 1C, and 1D. For each experiment, we introduce 4 levels of contamination, one for each sub-experiment. The levels of contamination are:

Experiments 1A: "Clean" (uncontaminated) data

Esperiments 1B: 5% of observations replaced with vertical outliers and leverage points

Esperiments 1C: 10% of observations replaced with vertical outliers and leverage points

Esperiments 1D: 20% of observations replaced with vertical outliers and leverage points

For Experiments B, C and D, the observations are randomly selected across the entire panel for replacement. The vertical outlier is $\tilde{y} \sim \text{iid.}N(0, 2)$ and the leverage points are distributed as $(\tilde{x}_1, \tilde{x}_2) \sim \text{iid.}(N(2, 2), N(2, 2))$. We then regress each DGP with both the CCEP and CCEMG estimators using the *xtdcce2* package in Stata (Ditzen, 2018). We record the biases and root mean squared errors (RMSEs).

3.4.1 Results

Tables 3.1 to 3.4 report the results for Experiments 1A to 1D. When the level of contamination was increased, we can see that both estimators fail, with biases and RMSEs inflating with more outliers present.

3.5 Conclusions

We can see that the CCE estimators are sensitive to outliers. In the next chapter, we consider a robust version of the Trimmed CCE estimators (Lee and Sul, 2022).

Table 3.1: Experiment 1A - Effects of outliers on the CCEMG and CCEP Estimators (β_1)

		CCEMG		CCEP	
N	T	Bias	RMSE	Bias	RMSE
20	20	-0.031	0.114	-0.029	0.995
30	20	0.030	0.096	0.030	0.840
30	30	0.029	0.076	0.028	0.695
50	20	0.031	0.080	0.032	0.723
50	30	0.032	0.066	0.032	0.609
50	50	0.030	0.055	0.030	0.519
100	20	-0.059	0.076	-0.068	0.770
100	30	-0.075	0.084	-0.107	1.031
100	50	-0.074	0.081	-0.107	1.018
100	100	-0.072	0.077	-0.105	0.997
200	20	-0.048	0.059	-0.060	0.628
200	30	-0.046	0.054	-0.059	0.601
200	50	-0.046	0.051	-0.059	0.582
200	100	-0.086	0.088	-0.114	1.060

Uncontaminated data

Table 3.2: Experiment 1B - Effects of outliers on the CCEMG and CCEP Estimators (β_1)

		CCEMG		CCEP	
N	T	Bias	RMSE	Bias	RMSE
20	20	-0.234	0.273	-0.340	0.373
30	20	-0.283	0.309	-0.425	0.448
30	30	-0.311	0.327	-0.410	0.425
50	20	-0.258	0.275	-0.404	0.419
50	30	-0.287	0.297	-0.388	0.398
50	50	-0.307	0.313	-0.368	0.374
100	20	-0.228	0.236	-0.366	0.373
100	30	-0.253	0.258	-0.350	0.354
100	50	-0.276	0.279	-0.336	0.338
100	100	-0.344	0.346	-0.380	0.382
200	20	-0.286	0.290	-0.453	0.457
200	30	-0.316	0.319	-0.433	0.436
200	50	-0.342	0.344	-0.416	0.418
200	100	-0.360	0.361	-0.399	0.399

5% of cross-section units are contaminated

Table 3.3: Experiment 1C - Small Sample Properties of the CCEMG and CCEP Estimators (β_1)

		CCEMG		CCEP	
N	T	Bias	RMSE	Bias	RMSE
20	20	-0.403	0.428	-0.506	0.527
30	20	-0.488	0.505	-0.623	0.637
30	30	-0.522	0.532	-0.610	0.619
50	20	-0.453	0.464	-0.594	0.603
50	30	-0.484	0.490	-0.575	0.581
50	50	-0.507	0.510	-0.558	0.562
100	20	-0.397	0.402	-0.528	0.532
100	30	-0.425	0.428	-0.513	0.516
100	50	-0.446	0.447	-0.497	0.499
100	100	-0.502	0.503	-0.529	0.530
200	20	-0.481	0.484	-0.638	0.641
200	30	-0.516	0.518	-0.621	0.622
200	50	-0.540	0.541	-0.602	0.603
200	100	-0.555	0.555	-0.587	0.587

10% of cross-section units are contaminated

Table 3.4: Experiment 1D - Small Sample Properties of the CCEMG and CCEP Estimators (β_1)

		CCEMG		CCEP	
N	T	Bias	RMSE	Bias	RMSE
20	20	-0.614	0.628	-0.670	0.681
30	20	-0.736	0.746	-0.814	0.822
30	30	-0.755	0.761	-0.800	0.806
50	20	-0.697	0.704	-0.779	0.784
50	30	-0.708	0.712	-0.756	0.760
50	50	-0.716	0.719	-0.740	0.742
100	20	-0.604	0.607	-0.683	0.685
100	30	-0.621	0.623	-0.668	0.669
100	50	-0.632	0.633	-0.655	0.656
100	100	-0.639	0.639	-0.645	0.645
200	20	-0.715	0.717	-0.805	0.806
200	30	-0.737	0.738	-0.793	0.793
200	50	-0.747	0.747	-0.777	0.777
200	100	-0.751	0.752	-0.765	0.765

20% of cross-section units are contaminated

References

Bai, J., 2009. Panel data models with interactive fixed effects. *Econometrica*, 77(4), pp.1229-1279.

Chudik, A. and Pesaran, M., 2013. Large Panel Data Models with Cross-Sectional Dependence: A Survey. *SSRN Electronic Journal*

Chudik, A. and Pesaran, M., 2015. Common Correlated Effects Estimation of Heterogenous Dynamic Panel Data Models with Weakly Exogenous Regressors. *SSRN Electronic Journal*

Chudik, A., Pesaran, M. and Tosetti, E., 2011. Weak and strong cross-section dependence and estimation of large panels. *Econometrics Journal*, 14, pp.C45-C90.

Coakley, J., Fuertes, A.M. and Smith, R., 2006. Unobserved heterogeneity in panel time series models. *Computational Statistics & Data Analysis*, 50(9), pp.2361-2380.

Ditzen, J., 2018. Estimating dynamic common-correlated effects in Stata. *The Stata Journal*, 18(3), pp.585-617.

Eberhardt, M., 2012. Estimating panel time-series models with heterogeneous slopes. *The Stata Journal*, 12(1), pp.61-71.

Holly, S., Pesaran, M.H. and Yamagata, T., 2011. The spatial and temporal diffusion of house prices in the UK. *Journal of urban economics*, 69(1), pp.2-23.

Pesaran, M., 2006. Estimation and Inference in Large Heterogeneous Panels with a Multifactor Error Structure. *Econometrica*, 74(4), pp.967-1012.

Pesaran, M.H., 2015. Testing weak cross-sectional dependence in large panels. *Econometric reviews*, 34(6-10), pp.1089-1117.

Pesaran, M.H. and Smith, R., 1995. Estimating long-run relationships from dynamic heterogeneous panels. *Journal of econometrics*, 68(1), pp.79-113.

Sarafidis, V. and Wansbeek, T., 2012. Cross-Sectional Dependence in Panel Data Analysis. *Econometric Reviews*, 31(5), pp.483-531.

Chapter 4

A Robust Trimmed Common Correlated Effects Estimator

4.1 Cross-Sectional Dependence in the Presence of Outliers

It is now well-known that least squares are not robust to outliers, having a breakdown point of 0. The CCE estimators of Pesaran (2006) are least squares estimates augmented with cross-section averages. This results in outliers at two levels: namely the least squares estimation; and the cross-section averages. There are two areas in this model which are affected by outliers. The first area in which outliers affect the estimates are the least squares estimation of the CCE factor model itself. The second area in which outliers affect the estimates is in the cross-section averages. The empirical mean is also well-known to be sensitive to outliers. Hence, a sufficiently large outlier among the cross-sections for each of the time series t is sufficient to distort the cross-section average $\bar{\mathbf{z}}_{wt}$ for that time series t. The same sufficiently large outlier could be a leverage point, a vertical outlier, or both in the main model, further distorting the estimates for the least squares estimation. In practice, outliers do not usually appear singly, often multiple outliers mask other outliers in the same sample, making results from traditional outlier detection methods less precise. In this paper, we define 'robustness' as resistance to distortions by outliers, in line with the robust statistics literature. In the wider econometrics literature, the term 'robustness' is often used to denote resistance to efficiency losses and consistency by heteroscedasticity, autocorrelation, and other aspects of the classical least squares assumptions. Two examples of work done for the CCE estimator regarding improved efficiencies include Newey and West (1987; quoted in Pesaran, 2006) and alternative variance estimator for small fixed-T panels by Joakim Westerlund, Yana Petrova and Milda Norkute (Westerlund, Petrova & Norkute, 2019). We will also use these variance estimators the next two chapters. For the rest of this paper, when we use the term 'robust', we shall refer to robustnstness to outliers. There are four types of outliers in panel data (Bramati and Croux, 2007). The outliers can be classified into block outliers (outlying cross-section units) and cell outliers (outlying observations within each cross-section unit). For each class, we have vertical outliers and leverage points. There is little work done on the robustness of the CCE estimator. Lee and Sul (2022) addressed the issue of leverage points in block outliers by developing the Trimmed CCE Mean Group and Pooled (TCCEMG and TCCEP) estimators. The TCCE estimators compute the sample variance of each cross-section unit and estimate a depth statistic for each sample variance in order to flag units with unusually large sample variances as outlying. In their paper, Lee and Sul (2022) used the Mahalanobis distance (Mahalanobis, 1936; Maesschalck, Jouan-Rimbaud & Massart, 2000 for a recent version) to compute the depth statistic. The Mahalanobis distance computes a standardised distance from each point to the centre of the data cloud, using the sample mean $\hat{\boldsymbol{\mu}}$ and sample covariance matrix $\hat{\boldsymbol{\Sigma}}$. It is well-known that the sample mean and covariance matrix are not robust to outliers, as only one sufficiently large observation will distort the mean and sample variance. The Mahalanobis distance detects a single outlier sufficiently well, but in samples with multiple outliers, the distance measure breaks down due to masking and swamping (Hardin and Rocke, 2005). The Minimum Covariance Determinant (MCD) is a robust answer to the Mahalanobis distance. It minimises the determinant of the covariance matrix of a subsample of size $h \leq n$, of the sample of size n. In this paper, we continue Lee and Sul's work by investigating the efficacy of the MCD and another robust estimator, the Minimum Volume Ellipsoid (MVE; Rousseeuw, 1984; see Van Aelst and Rousseeuw, 2009 for a comprehensive review), in the TCCE estimators. The next section discusses the mathematical foundations of the MCD and MVE estimators. Section (4.3) details simulation experiments evaluating the perfomance of the TCCE estimators with robust depth trimming. Section (4.3.1) discusses the results, and the final section concludes.

4.2 Mathematical Foundations

4.2.1 The TCCEMG

We begin by introducing the TCCEMG and TCCEP estimators of Lee and Sul (2022). We first define statistical depth (see Zuo and Serfling, 2000; Lee and Sul, 2020; etc. for more details): given a distribution P in $\mathbb{R}^d$, a depth function $D(x, P)$ is any function that provides an outward ordering of points $x \in \mathbb{R}^d$ from the centre of the distribution P. The Mahalanobis distance (MD) is a classical example of a depth function. It computes the standardised Euclidean distance from each observation

$\mathbf{x}_i$ to the sample mean $\bar{\mathbf{x}}$ of the sample:

$$MD_i = \sqrt{(\mathbf{x}_i - \bar{\mathbf{x}})'\boldsymbol{\Sigma}^{-1}(\mathbf{x}_i - \bar{\mathbf{x}})} \tag{4.1}$$

where $\boldsymbol{\Sigma}^{-1}$ is the sample covariance. To compute the depth statistic, the individual unit sample variances are first computed:

$$\boldsymbol{\Sigma}_i = \frac{1}{N}\sum_{i=1}^{N}(x_i - \mu_i)^2$$

and the vector of sample variances for the covariates of unit i are

$$\begin{aligned} \mathbf{v}_i &= (v_{1i}, v_{2i}, \ldots, v_{ki}) \\ &= diag(\boldsymbol{\Sigma}_i) \end{aligned} \tag{4.2}$$

The Mahalanobis Distance of the entire sample of variance vectors is then computed:

$$MD_i = (\mathbf{v}_i - \bar{\mathbf{v}})'\boldsymbol{\Sigma}^{-1}(\mathbf{v}_i - \bar{\mathbf{v}}) \tag{4.3}$$

where $\bar{\mathbf{v}}$ is the sample mean over all the sample variance vectors $\mathbf{v}_i$.

$$D_i = \frac{1}{1 + MD_i} \tag{4.4}$$

The Depth statistic is $\chi^2(1)$-distributed and units with sufficiently small D_i are flagged as outlying. In this paper, we follow the practice in Lee & Sul (2022) and trim the smallest 20% of depths as outliers. The TCCEMG estimator is then given by

$$\hat{\boldsymbol{\beta}}_{TCCEMG} = \frac{1}{n_G}\sum_{i\in G}\hat{\mathbf{b}}_i \tag{4.5}$$

where $n_G = |G|$ is the number of units in G, the set of units not trimmed by the depth statistic. $\hat{\mathbf{b}}_i$ is the usual individual CCEMG estimate

$$\hat{\mathbf{b}}_i = (\mathbf{X}_i'\bar{\mathbf{M}}_w\mathbf{X}_i)^{-1}(\mathbf{X}_i'\bar{\mathbf{M}}_w\mathbf{y}_i) \tag{4.6}$$

Similarly, the trimmed pooled CCE estimator (TCCEP) is

$$\boldsymbol{\beta}_{TCCEP} = \left(\frac{1}{n_G} \sum_{i \in G} \mathbf{X}_i' \bar{\mathbf{M}}_w \mathbf{X}_i \right)^{-1} \left(\frac{1}{n_G} \sum_{i \in G} \mathbf{X}_i' \bar{\mathbf{M}}_w \mathbf{y}_i \right) \tag{4.7}$$

with emphasis on the equal weights $\theta_i = \frac{1}{n_G}$, although any suitable weights may be used, both in the mean group and pooled versions. Thus, the Trimmed CCE estimators are simply the usual CCE estimators with outlying cross-section units trimmed. Consistency and asymptotic normality then follow from the original CCE estimators in Pesaran (2006). However, it is well-known in the literature that the MD suffers from the masking effect. This is because $\bar{\mathbf{x}}$ and $\boldsymbol{\Sigma}^{-1}$ are nonrobust, and a single sufficiently large outlier can distort the sample mean and inflate the sample covariance, giving other outliers small distances MD_i that escape detection (Rousseeuw and van Zomeren, 1990; Hubert, Debruyne and Rousseeuw, 2017). We therefore proceed to replace the Mahalanobis distance with robust location and scale estimates. We select the Minimum Covariance Determinant (MCD) and the Minimum Volume Ellipsoid (MVE).

4.2.2 The MCD estimator

The MCD (Rouseeuw, 1985; Hubert, Debruyne & Rouseeuw, 2018; Rousseeuw & Leroy, 1987; Maronna et. al, 2019) estimator is a strongly robust analogue to the Mahalanobis distance, being able to withstand as much as 50% contamination of the data. Formally, given a $N \times p$ data matrix $\mathbf{X}$ with rows $\mathbf{S} = \{\mathbf{X}_1, \mathbf{X}_2, \ldots, \mathbf{X}_N\}$, and a tuning constant h between $\lfloor (n+1+p)/2 \rfloor$ and N, the MCD estimator $(\hat{\boldsymbol{\mu}}_0, \hat{\boldsymbol{\Sigma}}_0)$ minimises the determinant of the sample covariance matrix over all subsets of of size h of the rows of $\mathbf{X}$:

$$(\hat{\boldsymbol{\mu}}_0, \hat{\boldsymbol{\Sigma}}_0) = \underset{S_h \subseteq \mathbf{S}}{\operatorname{argmin}}\ det(\hat{\boldsymbol{\Sigma}}) \tag{4.8}$$

where S_h is a subset of $\mathbf{S}$ of size h and $\hat{\boldsymbol{\Sigma}}$ is the sample covariance matrix corresponding to S_h, multiplied by a factor c_0 for consistency. Here, $\lfloor x \rfloor$ represents the largest integer smaller than or equal to x. The MCD estimator (as well as the MVE) Terms of Trades robustness with efficiency: the higher the breakdown point, the lower the efficiency. For example, a breakdown point of 50% results in an efficiency of only 6%. To retain efficiency, the MCD is often reweighted in practice, using suitable weights W_i to obtain the reweighted MCD estimator $(\hat{\boldsymbol{\mu}}_{MCD}, \hat{\boldsymbol{\Sigma}}_{MCD})$:

$$\hat{\boldsymbol{\mu}}_{MCD} = \frac{\sum_{i=1}^{N} W(d_i^2)\mathbf{x}_i}{\sum_{i=1}^{N} W(d_i^2)} \tag{4.9}$$

$$\hat{\boldsymbol{\Sigma}}_{MCD} = c_1 \frac{1}{N} \sum_{i=1}^{N} W(d_i^2)(\mathbf{x}_i - \hat{\boldsymbol{\mu}}_{MCD})(\mathbf{x}_i - \hat{\boldsymbol{\mu}}_{MCD})' \tag{4.10}$$

$d_i = \sqrt{(\mathbf{x}_i - \hat{\boldsymbol{\mu}}_0)'\hat{\boldsymbol{\Sigma}}_0^{-1}(\mathbf{x}_i - \hat{\boldsymbol{\mu}}_0)} \sim \chi^2_{p,0.975}$ and c_1 is another consistency factor. A typical set of weights used in software is $W(d^2) = I(d < \chi^2_{p,0.975}$. To illustrate the efficiency gains, for example, under a breakdown point of 50%, the efficiency improves to 45.5% for $p = 2$ and 82% for $p = 10$ (see Hubert & Debruyne, 2009; Hubert, Debruyne & Rousseeuw, 2017 for more details).

4.2.3 The MVE estimator

The MVE estimator (Van Aelst & Rouseeuw, 2009; Rousseeuw & Leroy, 1987; Maronna et.al., 2019) is another robust scale and location estimator. The MVE estimates the centre and scatter of the h most concentrated observations in the dataset by fitting the ellipsoid with minimal volume that covers at least h points of $\mathbf{X}$, with $\lfloor \frac{n}{2} + 1 \rfloor \leq h \leq N$. Formally, given $\mathbf{t}_N \in \mathbb{R}^p$ and $\mathbf{C}_N \in \text{PDS}(p)$, where PDS($p$) is the class of $p \times p$ positive definite symmetric matrices, the MVE estimator of scale and location is defined as

$$(\mathbf{t}_N, \mathbf{C}_N) = \underset{\mathbf{t} \in \mathbb{R}^p, \mathbf{C} \in \text{PDS}(p)}{\text{argmin}} \#\{i : (\mathbf{x}_i - \mathbf{t})'\mathbf{C}^{-1}(\mathbf{x}_i - \mathbf{t}) \leq c^2\} \leq h \tag{4.11}$$

where c is a consistency factor. For a multivariate normal distribution, $c = \sqrt{\chi^2_{p,h/n}}$. Maximal breakdown (of 50%) occurs when $h = \lfloor \frac{n+p+1}{2} \rfloor$. Like the MCD, gains in robustness are followed with losses in efficiency. Furthermore, the MVE has a relatively slow convergence rate to a non-Gaussian distribution at a rate of $n^{1/3}$. This leads to an asymptotic efficency of 0%. A reweighted MVE estimator $(\mathbf{t}_N^1, \mathbf{C}_N^1)$ is often used instead:

$$\mathbf{t}_N^1 = \frac{\sum_{i=1}^{N} W_i \mathbf{x}_i}{\sum_{i=1}^{N} W_i} \tag{4.12}$$

$$\mathbf{C}_N^1 = \frac{\sum_{i=1}^{N} W_i (\mathbf{x}_i - \mathbf{t}_N^1)(\mathbf{x}_i - \mathbf{t}_N^1)'}{\sum_{i=1}^{N} W_i} \tag{4.13}$$

where

$$W_i = \begin{cases} 1, & \text{RD}(\mathbf{x}_i) \leq \sqrt{\chi^2_{p,0.975}} \\ 0, & \text{otherwise} \end{cases} \tag{4.14}$$

and $\mathrm{RD}(\mathbf{x}_i) = \sqrt{(\mathbf{x}_i - \mathbf{t})'\mathbf{C}^{-1}(\mathbf{x}_i - \mathbf{t})}$ is the robust distance computed from the original MVE estimates obtained in (4.11).

4.2.4 Robust Depth statistics

Both the MCD and MVE estimators are affine equivariant, that is, their properties are preserved by linear transformations and scalings of data. In other words, estimates scale accordingly with changes in units. Both estimators are also available on Stata (Veradi, 2010) under the *robmv* command. In this paper, we replace the Mahalanobis distance MD_i with the MCD estimator and again with the MVE estimator to obtain robust estimates of depth in (4.4). The resulting asymptotics for the trimmed CCE estimators are otherwise similar, if not identical to the ones in Lee & Sul (2020; 2022). We therefore replace the MD in (4.4) with the MCD and the MVE estimates. As in(4.4), we compute the reweighted MCD estimate, MCD_i, for the sample variances $\mathbf{v}_i$. The resulting robust depth statistic computed from the MCD is

$$D_i^{MCD} = \frac{1}{1 + MCD_i} \tag{4.15}$$

Similarly, the robust depth statistic computed from the reweighted MVE is

$$D_i^{MVE} = \frac{1}{1 + MVE_i} \tag{4.16}$$

where MVE_i is the reweighted MVE estimate of $\mathbf{v}_i$. We then trim 20% of the units with smallest depth for a comparison. In the next section, we perform simulations to illustrate the performance of the MD, MCD and MVE estimators of depth on the TCCEMG and TCCEP estimators.

4.3 Performance of CCE estimators under outliers

In this section, we investigate the effect of the MD, MCD and MVE estimators on TCCE trimming. We follow the same model specification as Lee and Sul (2022). We perform four experiments, each with a different level of outlier contamination. For each experiment, we perform 3 sub-experiments A, B and C. We perform Experiment 1 and its sub-experiments 1A, 1B and 1C without any outliers. In Experiment 1A, we use the MD for the depth estimator D_i, which is identical to Lee and Sul's specification. We then use the MCD estimator for the depth statistic D_i^{MCD} in Experiment 1B. Finally, we use MVE_i and D_i^{MVE} in Experiment 1C. For each sub-experiment, we perform 2×14 sets of simulations: the TCCEMG and TCCEP estimations for $N = 20, 30, 50, 100, 200$; $T = 20, 30, 50, 100$;

$N \geq T$.

Monte Carlo Specification

The Monte Carlo specification is based on Lee and Sul's (2022) second Data-Generating Process (DGP) with depth-based trimming. The DGP is given by

$$\begin{aligned}
y_{it} &= \beta_{1i}x_{1it} + \beta_{2i}x_{2it} + a_iF_t + u_{it} \\
x_{jit} &= \rho x_{ji,t-1} + \lambda_{j,i}F_t + \epsilon_{jit}; \qquad j = 1, 2 \\
u_{it} &= \rho u_{i,t-1} + v_{it}
\end{aligned} \tag{4.17}$$

As with Lee and Sul, we set $\epsilon_{jit} \sim N(0, \sigma^2_{ji})$, where $\sigma_{ji} \sim \chi^2_1$ for $j = 1, 2$. The white noise errors v_{it} are set to $N(0, 1)$. Lee and Sul (2022) investigated two values of ρ (0 and 0.8), but reported little difference in the results except for the absolute values of the Mean Squared Errors (MSE). In all our experiments, we set $\rho = 0.5$ to represent the stationary nature of the data, in line with the assumptions of Pesaran (2006). As in Lee and Sul's experiments, the unobserved factor F_t is set as stationary AR(1), with $F_t = \rho F_t - 1 + e_t$, where $e_t \sim N(0, 1)$. Likewise, the parameter $\lambda_{2i} = \lambda_{1i} + \eta_i$, where $\lambda_{1i}, \eta_i \sim U(1, 2)$. Finally, we adopt the heterogeneous coefficients in Lee and Sul's paper: $\beta_{1i}, \beta_{2i} \sim N(1, 1)$ for each $i = 1, 2, \ldots, N$. We use this DGp with the above specifications in each of our experiments and sub-experiments.

Masking Outliers

We repeat the simulations of Experiment 1 by contaminating the generated samples with outlying cross-section units. Experiments 2, 3 and 4 are identical to Experiment 1, except that the data panels have been contaminated with outlying cross-section units. In each of experiments 2, 3 and 4, we contaminate, for each data panel, the first $h = \lfloor pN \rfloor$ cross-section units (or cells) with an outlying distribution. In order to capture the masking effect, we included a large variability in the outliers, which would overlap with the bulk of the data with high probability. Specifically, for $N = 1, 2, \ldots, h$, we replace $\beta_{1i} = \beta_{2i} = 5$ as in Lee and Sul (2022). For the same units, we then replace the σ_{ji}s with $\boldsymbol{\sigma}_i \sim |N(1, 5^2))|$. A similar situation is discussed in the data on milk bottles given by Daudin, Dauby and Trecourt and analyzed by Atkinson (Daudin, Duby, and Trecourt 1988; Atkinson 1994; all quoted in Hardin and Rocke, 2005). Hardin and Rocke (2005) also contains a similar example. In figure 3 of their paper, Hardin and Rocke showed two obvious outliers far away from the rest of the points in the distance-expected distance plots, but there is stilla need to emphasise a cutoff point. Lee and Sul(2022) trimmed the bottom 20% of depths. We follow a similar trimming scheme, in which we select a cutoff of $MD_i = MCD_i = MVE_i = 4$, which corresponds to $D_i^* = D_i^{*,MCD} = D_i^{*,MVE} = 0.2$. We trim all depths below this cutoff value. Figures 4.1, 4.2 and 4.3 illustrate the distribution of these outliers. Starting with figure 4.1, under 10% contamination, we see that units $N = 1, 5, 7, 9$ and 10 are simulated as actual outliers, but due to their (suffciently small) sample variances, these units escape detection by MD estimation. In contrast, for figure 4.2 under 10% contamination, we see that there are no masked outliers. The tolerance threshold for each depth estimator has been set to the same magnitude ($MD_i^* = MCD_i^* = MVE_i^* = 4$). Despite this, for the majority of the outliers, we have the corresponding $MCD_i, MVE_i > MD_i$. This illustrates the masking effect: the remaining outliers (with larger variances, marked as red) have distorted the sample mean and covariance matrix used in the computation of the Mahalanobis distances, hence causing the computed MD_i to be biased downwards, escaping outlier detection. Units $N = 1, 9$ by nature have sufficiently small variances, hence they have robust distances that are very similar to that of the bulk of the data. In practice, these observations would be counted as the bulk of the data (Note that the TCCE Depth trimming targets only the sample variances). These outliers are not necessarily masked; if the data contains no other outlying parameters (apart from the targeted sample variances), these observations would be fitted alongside the bulk of the data. The resulting estimates would still be robust then. The same effects can be seen in the MVE estimators (Figure 4.3) and for 20% contamination for all three depth estimators. Due to the affine equivariance of the MD, MCD and MVE, the transformations into D_i, D_i^{MCD} and D_i^{MVE} respectively preserve the above properties, hence the scatterplots for the depth statistics are similar. We will leave these scatterplots in the appendix. The level of contamination p is set to 5% for Experiment 2, 10% for Experiment 3 and 20% for Experiment 4. The next section compiles and discusses the results.

4.3.1 Results

Tables 4.1 to 4.4 display the small sample results for Experiments 1 to 4 respectively for the slope estimates $\beta_{1,TCCEMG}$ and $\beta_{2,TCCEP}$. Tables 4.5 to 4.8 report the same set of results for $\beta_{2,TCCEMG}$ and $\beta_{2,TCCEP}$. The pattern in these results are similar. We will leave these results in the appendix. Experiment 1 shows the estimates without any contamination. In Experiment 2, under 5% (mild) contamination, the bias of the MD-based estimates increase from $O(10^{-3})$ to $O(10^{-1})$, whereas the MCD- and MVE-based estimates have only increased from $O(10^{-3})$ to $O(10^{-2})$. The sizes for the MCD- and MVE-based slope estimates have also resisted much distortion, unlike the MD-based estimates. In Experiment 3, under 10% (moderate) contamination, the sizes of the MCD- and MVE-estimates are distorted more than in Experiment 2; however this distortion is nowhere as severe as in the MD-based

Figure 4.1: Mahalanobis distances by Cross-section

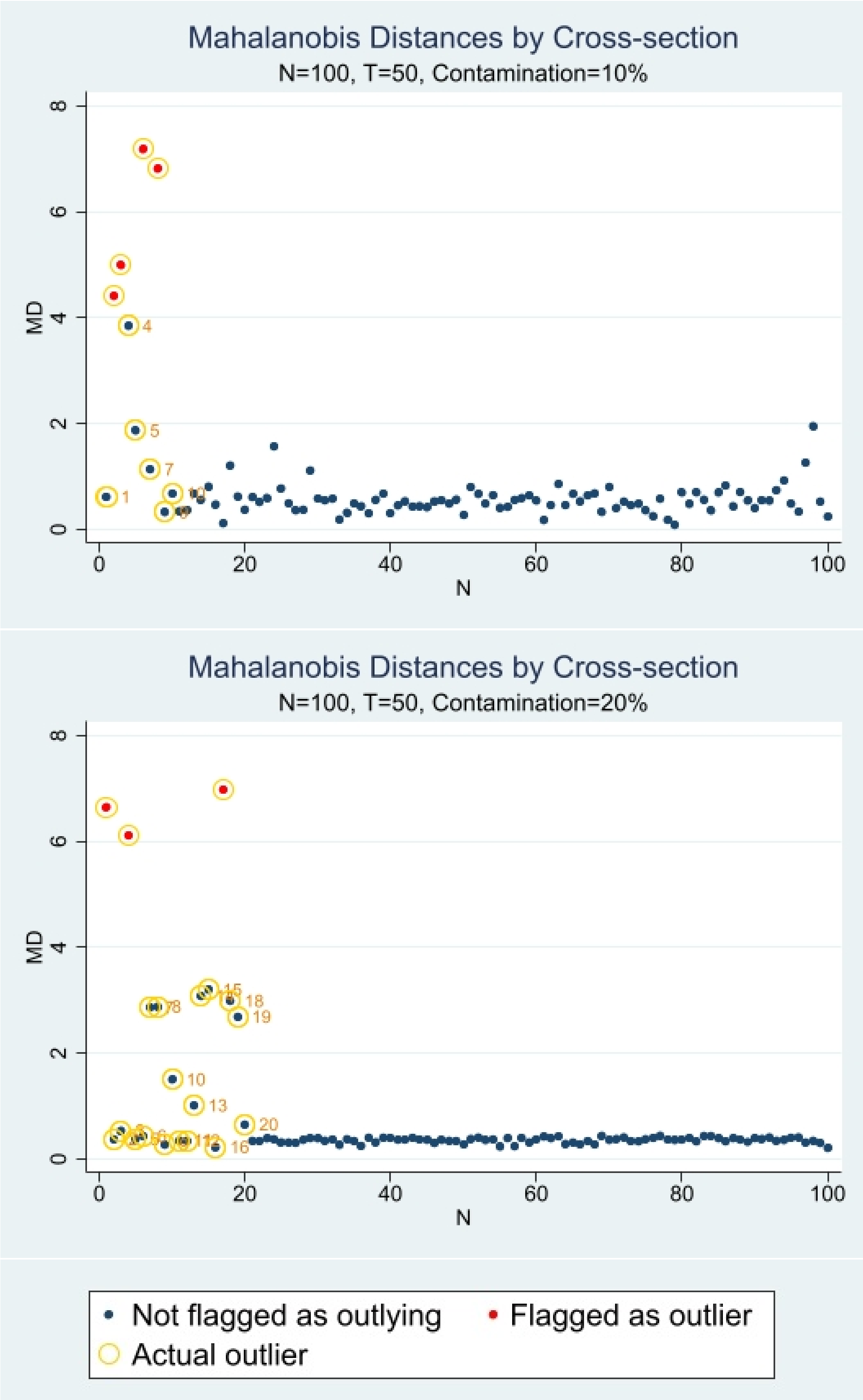


Figure 4.2: MCD distances by Cross-section

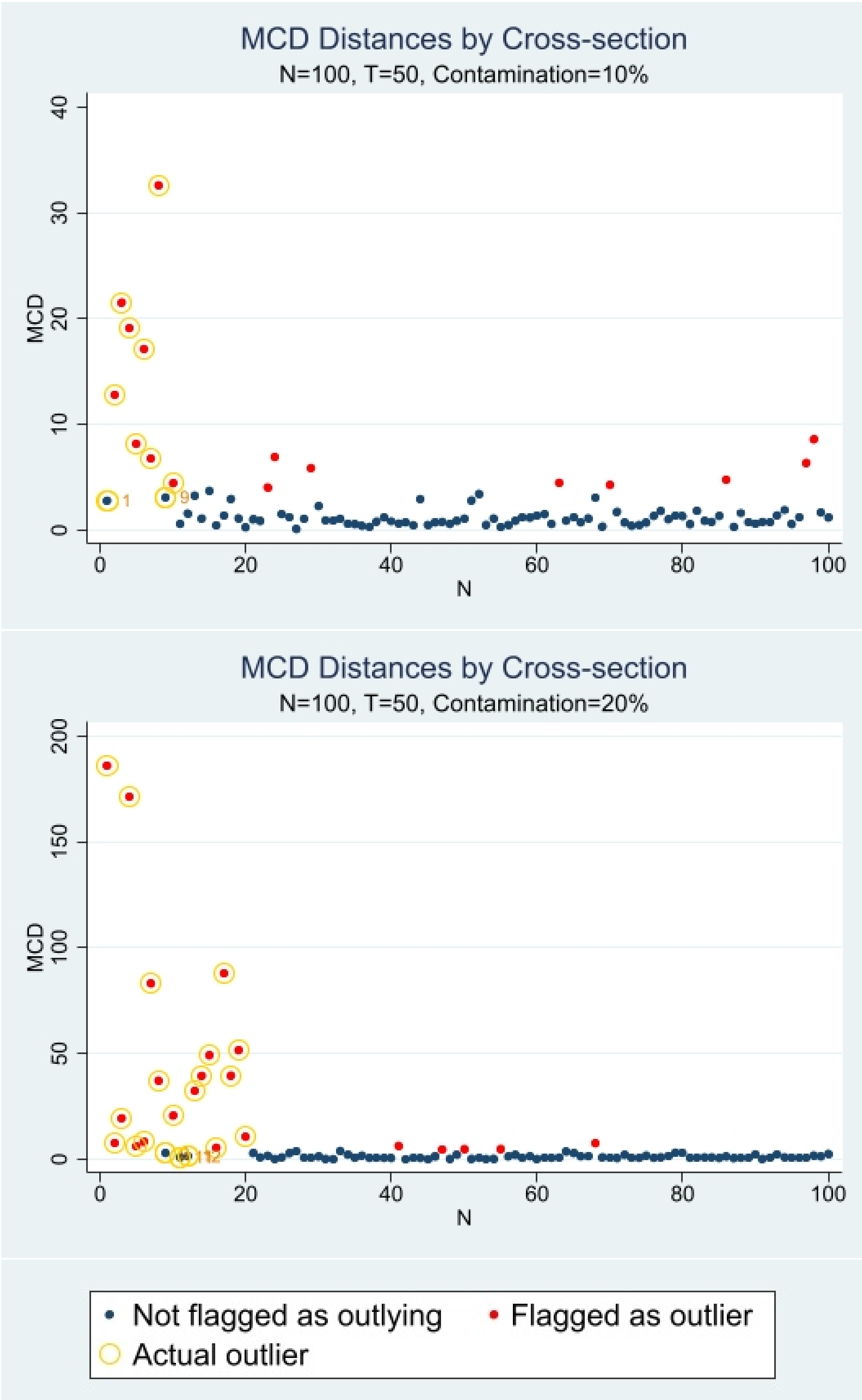
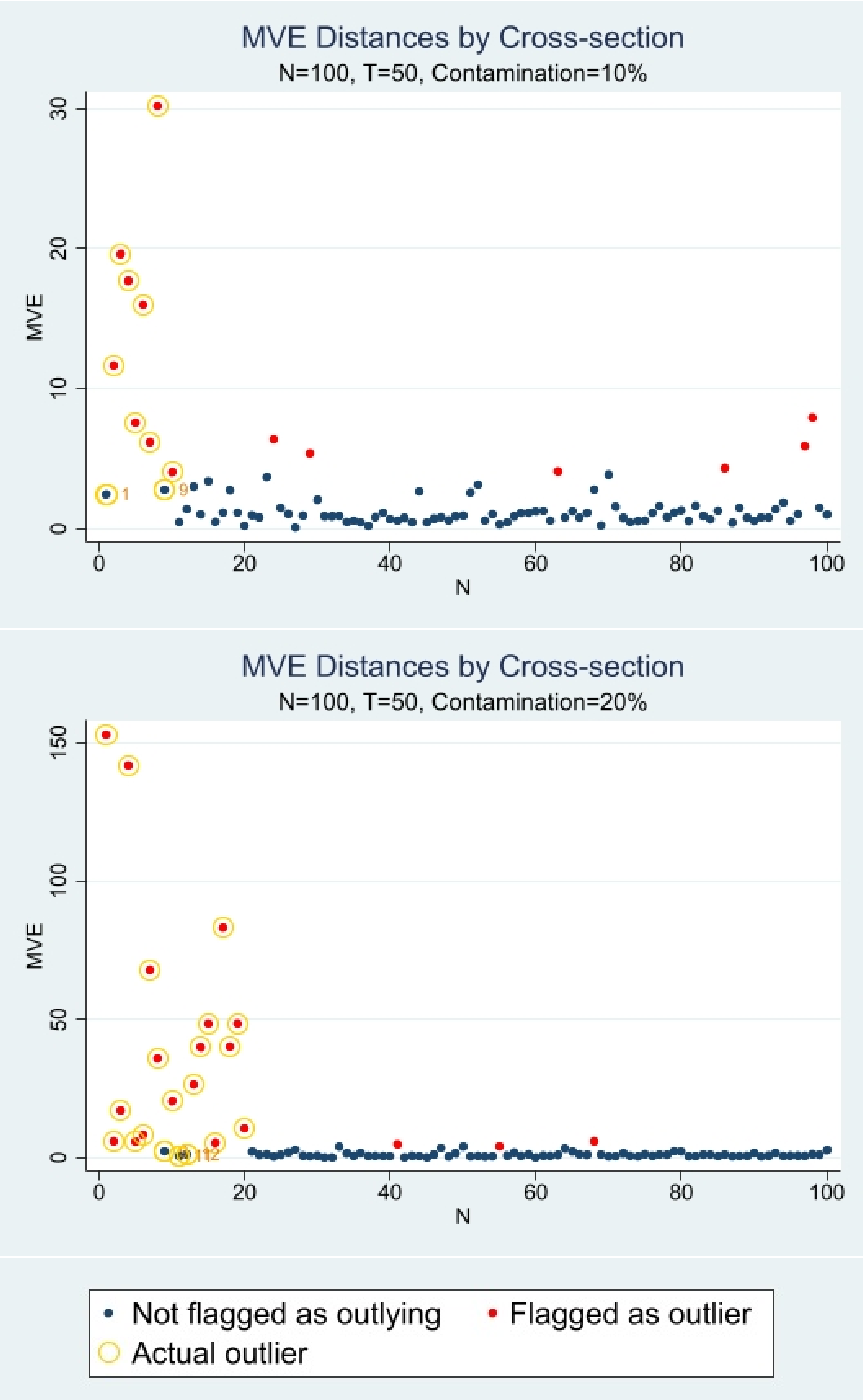


Figure 4.3: MVE distances by Cross-section



estimates. The biases for the MCD-and MVE-based estimates have remained steady at $O(10^{-2})$ for the TCCEMG, whereas the MD-based estimates have been severely distorted, rendering them unusable. This steadiness illustrates the boundedness of the influence function of the MCDand MVE estimators. In robust statistics, a bounded influence function implies that within the breakdown point, the estimates will not deviate beyond a certain finite value; see Hubert, Debruyne and Rousseeuw, 2018 for a review). The distortions for all 3 estimates (MD-, MCD- and MVE-based) are even greater in Experiment 4. However, the biases for the MCD- amd MVE-based estimates have resisted much distortion, with only 21.2% distortion at $(N, T) = (200, 100)$ for the MCD. Similarly, the sizes and RMSEs for the MCD- and MVE-based estimates have resisted much distortion, compared to the MD-based ones. A possible explanation for the gradual increase in distortions is due to non-extreme outliers. The TCCE estimators overall only detect outlying variances. The distribution of the outlying variances in the simulation allows for a fraction of outlying units with variances that fit into the bulk of the data (see Figures 4.1 to 4.3 for illustration). These outlying units are not masked by other outliers, rather their depth statistics match those of the bulk of the data. In practice, these outliers would be considered a fit for the data, if every other parameter were not outlying. Thus, if these experiments only dealt with outlying variances and not slope estimates, we would have obtained robust results for all experiments. Since $\beta_{1i} = \beta_{2i} = 5 \geq 1$ for all outlying observations, the slopes of the undetected outliers were not trimmed and caused the distortions in the estimates. It then follows that the MCD and MVE also captured more outliers than the traditional MD, resulting in less undetected outliers distorting the sample. This is again illustrates the masking effect in non-robust estimators, since the outliers would already distort the mean and sample variance estimates in the computation of the MD. We will deal with outlying slope estimates in a later chapter.

4.4 Conclusion

We have investigated the performance of the TCCEMG and TCCEP estimators under robust depth trimming. The MCD and MVE estimators are able to resist distortions from outliers and address the masking effect that affects non-robust (multivariate) estimators. The TCCE estimators only address outlying variances by design (see Lee and Sul, 2022). If there exist outlying cross-section units whose variances fit the bulk of the data whilst having extreme outliers in other parameters, these aberrant units naturally remain undetected and potentially cause distortions in the estimates. Nevertheless, the MCD and MVE estimators are still robust compared to the traditional MD. Considerations for future work include deriving the limiting distributions for the robust TCCE estimators. In the next two chapters, we apply the robust TCCE estimators to data in Economic growth.

Table 4.1: Experiment 1 - Effects of Outliers on Small Sample Properties of the CCEMG and CCEP Estimators (β_1)

TCCEMG

		MD			MCD			MVE		
N	T	Bias	RMSE	Size	Bias	RMSE	Size	Bias	RMSE	Size
20	20	0.015	0.117	0.096	0.013	0.135	0.101	0.012	0.132	0.096
30	20	0.006	0.077	0.079	0.009	0.086	0.079	0.007	0.084	0.079
30	30	0.007	0.072	0.080	0.009	0.082	0.085	0.008	0.081	0.081
50	20	0.004	0.047	0.079	0.005	0.052	0.081	0.004	0.051	0.080
50	30	0.005	0.043	0.067	0.005	0.048	0.072	0.005	0.047	0.069
50	50	0.004	0.043	0.068	0.004	0.047	0.069	0.004	0.046	0.069
100	20	0.001	0.023	0.068	0.002	0.025	0.068	0.000	0.025	0.067
100	30	0.000	0.022	0.067	0.000	0.024	0.063	0.001	0.024	0.064
100	50	-0.001	0.022	0.067	-0.002	0.023	0.065	-0.001	0.023	0.066
100	100	0.003	0.021	0.059	0.004	0.023	0.059	0.003	0.022	0.060
200	20	0.001	0.011	0.067	0.002	0.013	0.070	0.002	0.012	0.067
200	30	0.000	0.011	0.062	0.001	0.012	0.060	0.000	0.012	0.058
200	50	0.001	0.010	0.057	0.001	0.011	0.059	0.001	0.011	0.058
200	100	0.000	0.010	0.052	0.000	0.011	0.057	0.001	0.011	0.063

TCCEP

		MD			MCD			MVE		
N	T	Bias	RMSE	Size	Bias	RMSE	Size	Bias	RMSE	Size
20	20	0.010	0.240	0.094	0.009	0.243	0.102	0.005	0.237	0.097
30	20	-0.004	0.171	0.079	0.003	0.173	0.079	0.002	0.167	0.079
30	30	-0.001	0.167	0.077	0.002	0.163	0.083	0.003	0.164	0.085
50	20	-0.001	0.109	0.064	0.003	0.106	0.066	0.003	0.106	0.067
50	30	0.002	0.104	0.066	0.000	0.102	0.066	0.000	0.102	0.064
50	50	0.003	0.104	0.070	0.000	0.100	0.077	0.002	0.101	0.071
100	20	-0.002	0.058	0.054	0.001	0.055	0.052	-0.002	0.055	0.051
100	30	-0.003	0.058	0.064	0.000	0.053	0.062	0.000	0.054	0.065
100	50	-0.003	0.054	0.055	-0.002	0.051	0.058	-0.003	0.052	0.056
100	100	-0.001	0.053	0.055	0.001	0.049	0.057	0.001	0.049	0.056
200	20	0.000	0.030	0.057	0.001	0.028	0.057	0.002	0.028	0.052
200	30	0.000	0.028	0.048	0.000	0.026	0.053	-0.001	0.027	0.054
200	50	-0.002	0.028	0.057	-0.003	0.025	0.050	-0.002	0.026	0.052
200	100	-0.001	0.027	0.053	-0.002	0.025	0.056	0.000	0.026	0.058

Uncontaminated DGP

Table 4.2: Experiment 2 - Effects of Outliers on Small Sample Properties of the CCEMG and CCEP Estimators (β_1)

TCCEMG

		MD			MCD			MVE		
N	T	Bias	RMSE	Size	Bias	RMSE	Size	Bias	RMSE	Size
20	20	0.121	0.176	0.120	0.042	0.154	0.096	0.055	0.161	0.098
30	20	0.057	0.095	0.094	0.030	0.096	0.088	0.033	0.096	0.087
30	30	0.052	0.088	0.088	0.028	0.089	0.080	0.033	0.090	0.082
50	20	0.090	0.070	0.121	0.035	0.059	0.083	0.036	0.059	0.085
50	30	0.085	0.066	0.112	0.030	0.056	0.075	0.032	0.056	0.076
50	50	0.080	0.061	0.109	0.030	0.053	0.075	0.031	0.054	0.079
100	20	0.173	0.066	0.298	0.032	0.030	0.085	0.035	0.030	0.090
100	30	0.170	0.063	0.292	0.033	0.030	0.086	0.035	0.029	0.085
100	50	0.172	0.062	0.315	0.030	0.028	0.074	0.032	0.028	0.075
100	100	0.168	0.059	0.299	0.029	0.027	0.067	0.033	0.027	0.067
200	20	0.191	0.057	0.550	0.031	0.018	0.089	0.034	0.019	0.088
200	30	0.188	0.054	0.559	0.030	0.015	0.080	0.032	0.015	0.087
200	50	0.186	0.051	0.559	0.030	0.014	0.077	0.032	0.014	0.078
200	100	0.184	0.050	0.562	0.028	0.014	0.074	0.030	0.014	0.076

TCCEP

		MD			MCD			MVE		
N	T	Bias	RMSE	Size	Bias	RMSE	Size	Bias	RMSE	Size
20	20	0.149	0.423	0.091	0.033	0.274	0.101	0.050	0.300	0.094
30	20	0.066	0.241	0.070	0.023	0.193	0.075	0.028	0.194	0.075
30	30	0.064	0.225	0.077	0.024	0.182	0.084	0.031	0.184	0.084
50	20	0.174	0.302	0.053	0.036	0.128	0.067	0.036	0.130	0.069
50	30	0.166	0.285	0.055	0.034	0.122	0.067	0.036	0.122	0.066
50	50	0.152	0.274	0.057	0.026	0.115	0.071	0.028	0.117	0.070
100	20	0.562	0.819	0.056	0.043	0.069	0.056	0.049	0.073	0.056
100	30	0.545	0.769	0.060	0.043	0.067	0.068	0.046	0.069	0.071
100	50	0.552	0.765	0.062	0.033	0.063	0.065	0.039	0.065	0.064
100	100	0.541	0.737	0.054	0.034	0.061	0.060	0.042	0.064	0.057
200	20	0.775	1.000	0.311	0.043	0.037	0.058	0.053	0.040	0.059
200	30	0.785	1.015	0.320	0.040	0.034	0.057	0.047	0.037	0.059
200	50	0.768	0.963	0.324	0.045	0.034	0.067	0.052	0.036	0.064
200	100	0.760	0.952	0.323	0.036	0.032	0.061	0.041	0.034	0.063

5% of cross-section units are contaminated

Table 4.3: Experiment 3 - Effects of Outliers on Small Sample Properties of the CCEMG and CCEP Estimators (β_1)

TCCEMG

		MD			MCD			MVE		
N	T	Bias	RMSE	Size	Bias	RMSE	Size	Bias	RMSE	Size
20	20	0.332	0.327	0.186	0.081	0.179	0.099	0.107	0.194	0.104
30	20	0.345	0.257	0.248	0.076	0.122	0.097	0.086	0.125	0.101
30	30	0.349	0.256	0.262	0.077	0.116	0.095	0.089	0.121	0.100
50	20	0.406	0.259	0.471	0.073	0.075	0.095	0.083	0.078	0.103
50	30	0.404	0.246	0.488	0.072	0.070	0.091	0.079	0.072	0.096
50	50	0.409	0.252	0.510	0.072	0.069	0.094	0.079	0.071	0.098
100	20	0.457	0.263	0.838	0.069	0.040	0.115	0.077	0.041	0.125
100	30	0.458	0.258	0.859	0.067	0.037	0.108	0.073	0.038	0.114
100	50	0.458	0.255	0.879	0.065	0.035	0.102	0.070	0.036	0.103
100	100	0.459	0.253	0.898	0.068	0.034	0.096	0.074	0.035	0.103
200	20	0.471	0.253	0.983	0.069	0.022	0.150	0.076	0.023	0.172
200	30	0.471	0.248	0.992	0.066	0.020	0.134	0.073	0.022	0.158
200	50	0.472	0.248	0.996	0.065	0.019	0.121	0.071	0.020	0.138
200	100	0.466	0.240	0.995	0.061	0.018	0.116	0.066	0.019	0.127

TCCEP

		MD			MCD			MVE		
N	T	Bias	RMSE	Size	Bias	RMSE	Size	Bias	RMSE	Size
20	20	0.610	1.550	0.082	0.077	0.319	0.103	0.114	0.368	0.097
30	20	0.737	1.604	0.065	0.073	0.239	0.080	0.092	0.262	0.082
30	30	0.718	1.510	0.061	0.067	0.224	0.083	0.084	0.246	0.085
50	20	1.179	2.635	0.109	0.085	0.163	0.064	0.108	0.186	0.065
50	30	1.192	2.608	0.115	0.083	0.154	0.065	0.100	0.171	0.067
50	50	1.196	2.618	0.118	0.077	0.147	0.072	0.093	0.163	0.072
100	20	1.621	3.629	0.623	0.092	0.095	0.073	0.112	0.110	0.074
100	30	1.637	3.679	0.652	0.085	0.088	0.067	0.105	0.104	0.076
100	50	1.649	3.690	0.673	0.085	0.082	0.068	0.097	0.092	0.067
100	100	1.629	3.604	0.674	0.085	0.082	0.072	0.102	0.092	0.072
200	20	1.857	4.070	0.988	0.097	0.053	0.078	0.120	0.067	0.096
200	30	1.891	4.189	0.990	0.092	0.050	0.087	0.112	0.060	0.100
200	50	1.887	4.167	0.993	0.087	0.046	0.081	0.107	0.056	0.090
200	100	1.878	4.111	0.992	0.084	0.045	0.081	0.102	0.053	0.092

10% of cross-section units are contaminated

Table 4.4: Experiment 4 - Effects of Outliers on Small Sample Properties of the CCEMG and CCEP Estimators (β_1)

TCCEMG		MD			MCD			MVE		
N	T	Bias	RMSE	Size	Bias	RMSE	Size	Bias	RMSE	Size
20	20	0.861	1.015	0.561	0.170	0.246	0.110	0.230	0.295	0.129
30	20	0.854	0.923	0.776	0.181	0.194	0.129	0.218	0.221	0.148
30	30	0.857	0.907	0.796	0.162	0.169	0.131	0.194	0.191	0.151
50	20	0.900	0.927	0.970	0.170	0.121	0.151	0.207	0.143	0.188
50	30	0.906	0.940	0.980	0.168	0.115	0.153	0.200	0.134	0.189
50	50	0.907	0.928	0.984	0.164	0.117	0.154	0.193	0.129	0.184
100	20	0.944	0.959	1.000	0.172	0.077	0.239	0.199	0.089	0.295
100	30	0.943	0.953	1.000	0.162	0.070	0.214	0.188	0.082	0.265
100	50	0.951	0.962	1.000	0.160	0.067	0.220	0.183	0.078	0.267
100	100	0.952	0.961	1.000	0.155	0.064	0.209	0.180	0.075	0.261
200	20	0.958	0.955	1.000	0.161	0.050	0.375	0.187	0.060	0.457
200	30	0.960	0.956	1.000	0.156	0.046	0.363	0.180	0.056	0.443
200	50	0.961	0.956	1.000	0.151	0.044	0.347	0.172	0.051	0.405
200	100	0.960	0.951	1.000	0.148	0.041	0.332	0.169	0.050	0.406

TCCEP		MD			MCD			MVE		
N	T	Bias	RMSE	Size	Bias	RMSE	Size	Bias	RMSE	Size
20	20	2.246	7.680	0.414	0.164	0.446	0.101	0.270	0.674	0.101
30	20	2.378	7.712	0.613	0.199	0.393	0.087	0.276	0.537	0.099
30	30	2.416	7.807	0.653	0.175	0.344	0.090	0.243	0.473	0.100
50	20	2.710	8.813	0.939	0.214	0.285	0.083	0.294	0.422	0.100
50	30	2.728	8.855	0.956	0.211	0.270	0.091	0.290	0.408	0.112
50	50	2.725	8.804	0.963	0.194	0.247	0.084	0.264	0.358	0.098
100	20	2.926	9.391	1.000	0.235	0.189	0.112	0.317	0.301	0.148
100	30	2.928	9.368	1.000	0.223	0.173	0.117	0.295	0.272	0.151
100	50	2.944	9.447	1.000	0.214	0.163	0.121	0.281	0.251	0.152
100	100	2.937	9.395	1.000	0.203	0.150	0.110	0.276	0.244	0.145
200	20	3.045	9.706	1.000	0.237	0.129	0.216	0.323	0.221	0.296
200	30	3.046	9.702	1.000	0.225	0.116	0.216	0.303	0.198	0.295
200	50	3.046	9.691	1.000	0.217	0.109	0.208	0.283	0.172	0.275
200	100	3.050	9.702	1.000	0.212	0.102	0.209	0.286	0.177	0.290

20% of cross-section units are contaminated

References

Bramati, M. and Croux, C., 2007. Robust estimators for the fixed effects panel data model. *The Econometrics Journal*, 10(3), pp.521-540. *Computational Statistics*, 33(2), pp.675-708.

Ditzen, J., 2018. Estimating dynamic common-correlated effects in Stata. *The Stata Journal*, 18(3), pp.585-617.

Hardin, J. and Rocke, D.M., 2005. The distribution of robust distances. *Journal of Computational and Graphical Statistics*, 14(4), pp.928-946.

Hubert, M. and Debruyne, M., 2010. Minimum covariance determinant. *Wiley interdisciplinary reviews: Computational statistics*, 2(1), pp.36-43.

Hubert, M., Debruyne, M. and Rousseeuw, P.J., 2018. Minimum covariance determinant and extensions. *Wiley Interdisciplinary Reviews: Computational Statistics*, 10(3), p.e1421.

Lee, Y. and Sul, D., 2020. Depth-Weighted Estimation of Heterogenous Agent Panel Data Models. *Mimeo, Syracuse University.*

Lee, Y. and Sul, D., 2022. Trimmed mean group estimation. In Essays in Honor of M. Hashem Pesaran: Panel Modeling, Micro Applications, and Econometric Methodology (pp. 177-202). *Emerald Publishing Limited.*

Maronna, R.A., Martin, R.D., Yohai, V.J. and Salibián-Barrera, M., 2019. *Robust statistics: theory and methods (with R)*. John Wiley & Sons.

Pesaran, M.H., 2006. Estimation and inference in large heterogeneous panels with a multifactor error structure. *Econometrica*, 74(4), pp.967-1012.

Rousseeuw, P. and Hubert, M., 2011. Robust statistics for outlier detection. *WIREs Data Mining and Knowledge Discovery*, 1(1), pp.73-79.

Rousseeuw, P. and Leroy, A., 1987. Robust Regression And Outlier Detection. *Hoboken: John Wiley & Sons.* Rousseeuw, P., 1984. Least Median of Squares Regression. *Journal of the American Statistical Association*, 79(388), pp.871-880.

Van Aelst, S. and Rousseeuw, P., 2009. Minimum volume ellipsoid. *Wiley Interdisciplinary Reviews: Computational Statistics*, 1(1), pp.71-82.

Verardi, V. and Dehon, C., 2010. Multivariate outlier detection in Stata. *The Stata Journal*, 10(2), pp.259-266.

Westerlund, J., Petrova, Y. and Norkute, M., 2019. CCE in fixed-T panels. *Journal of Applied Econometrics*, 34(5), pp.746-761.

Zuo, Y. and Serfling, R., 2000. Structural properties and convergence results for contours of sample statistical depth functions. *Annals of Statistics*, pp.483-499.

Appendix

4.A Illustration of the Masking Effect of Outliers

Figures 4.4 to 4.6 display the plots of the log-sample variances of $\mathbf{x}_{1i}$ against $\mathbf{x}_{2i}$. We see that the MD has missed several outliers, whereas the MCD and MVE have detected them within a (relatively) generous tolerance band. The MCD flags more outliers than the MVE; otherwise, both are effective at capturing outliers that are not embedded within the bulk of the data cloud. Figures 4.7 to 4.9 Display the computed depths for the MD, MCD and MVE estimators for a generated sample of $(N, T) = (100, 50)$ for contamination levels of 10% and 20%. The effects between contamination levels are similar, except they are more pronounced in higher levels of contamination.

4.B Estimates and Small Sample Results of β_2

Tables 4.5 to 4.8 report the results for β_2. The results are similar, except for a slight upward estimated bias compared to β_1. This also occurs in Experiment 1, without outliers.

Figure 4.4: Outliers Flagged by MD

Figure 4.5: Outliers Flagged by MCD

Figure 4.6: Outliers Flagged by MVE

Figure 4.7: MD Depths by Cross-section

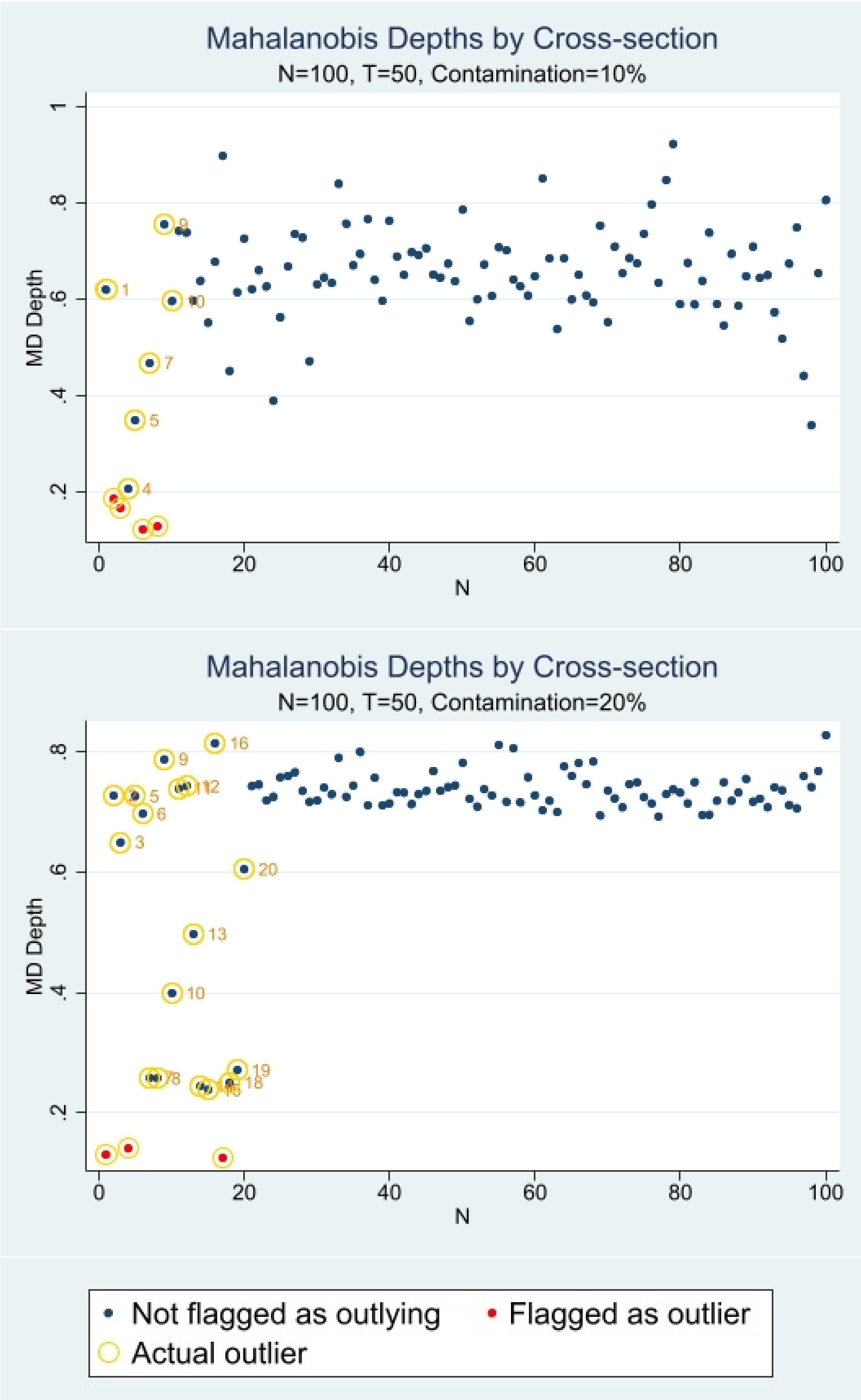


Figure 4.8: MCD Depths by Cross-section

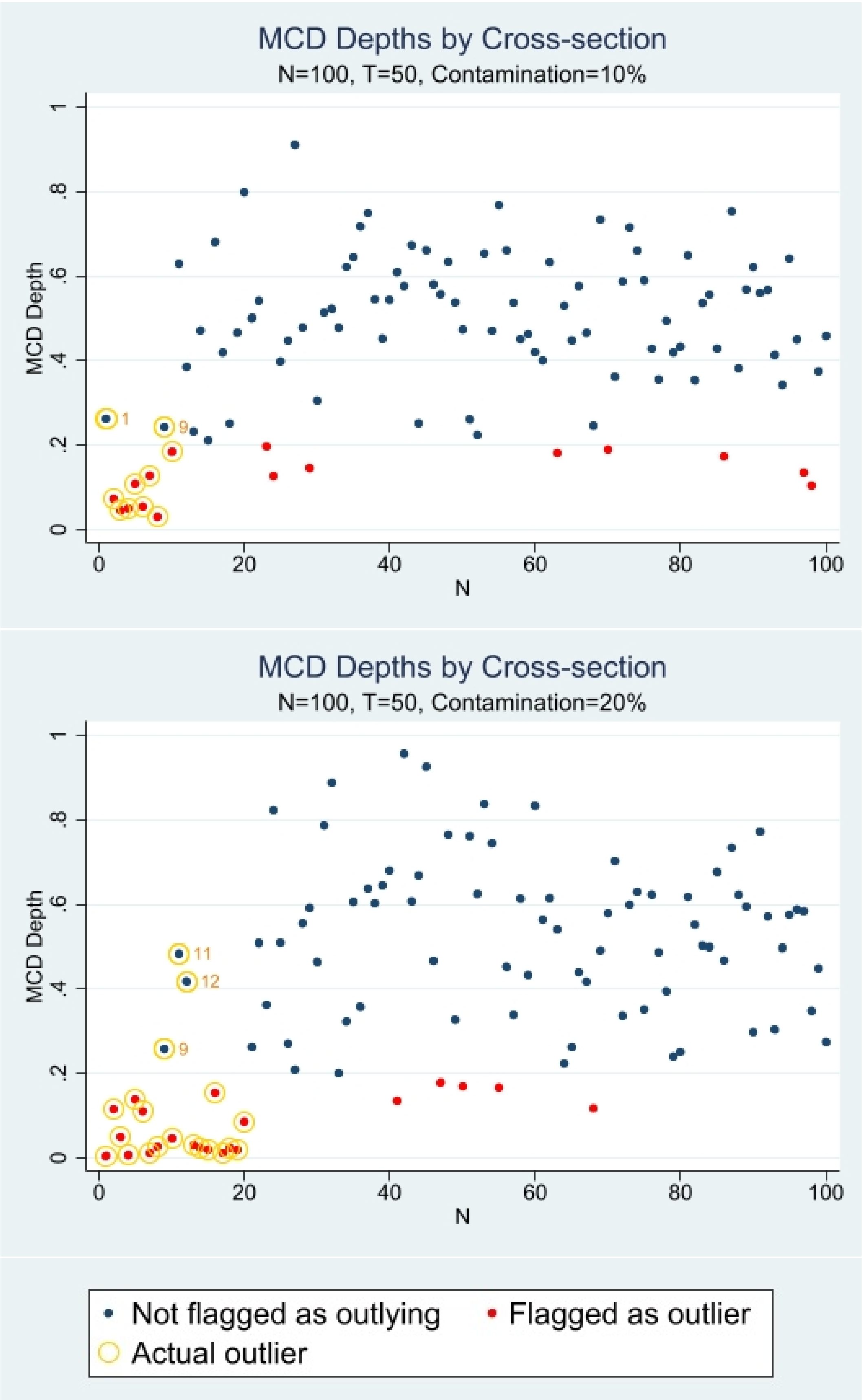


Figure 4.9: MVE Depths by Cross-section

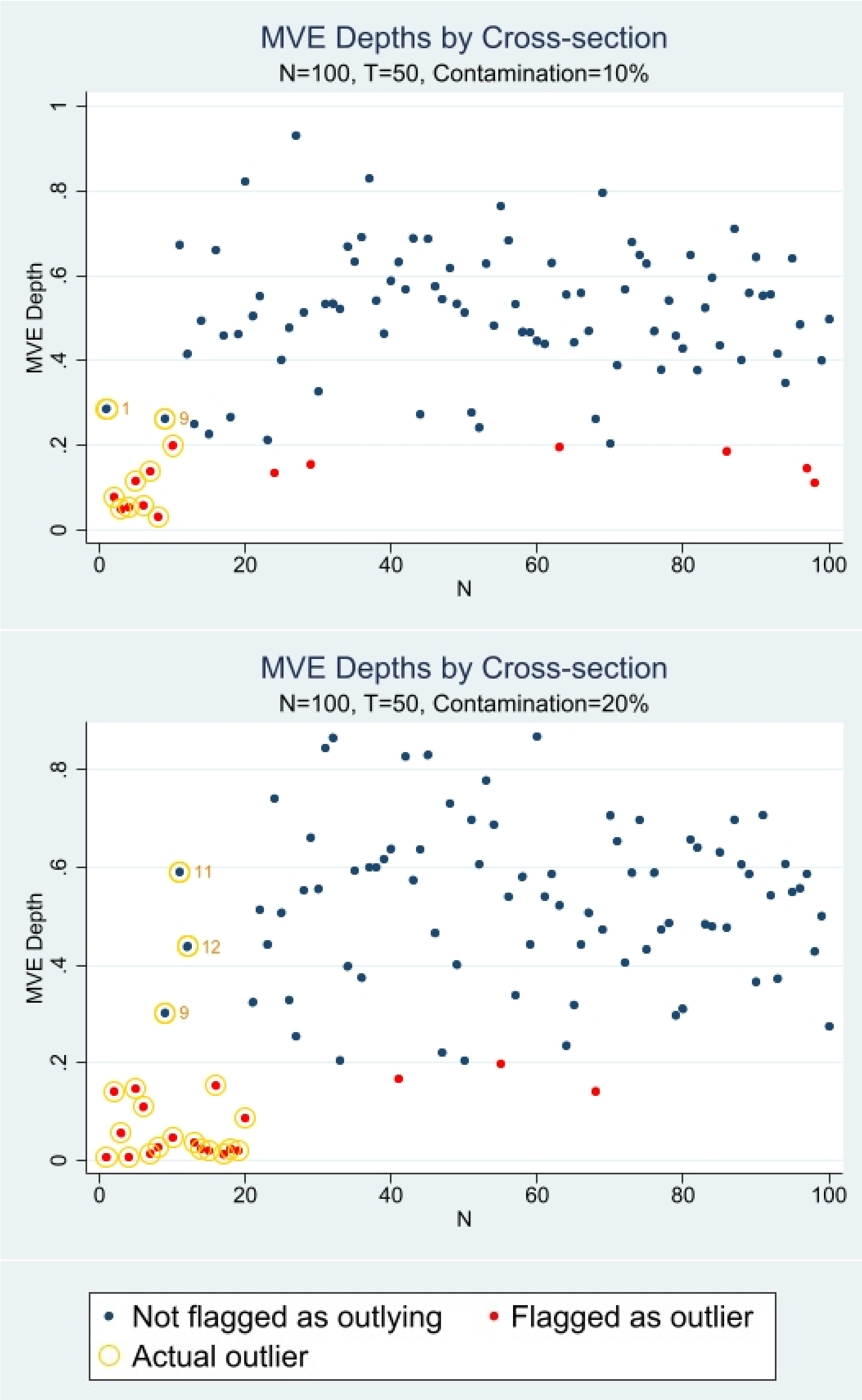


Table 4.5: Experiment 1 - Effects of Outliers on Small Sample Properties of the CCEMG and CCEP Estimators (β_2)

TCCEMG		MD			MCD			MVE		
N	T	Bias	RMSE	Size	Bias	RMSE	Size	Bias	RMSE	Size
20	20	0.016	0.107	0.091	0.013	0.122	0.094	0.014	0.121	0.095
30	20	0.009	0.074	0.090	0.008	0.084	0.089	0.008	0.082	0.092
30	30	0.007	0.070	0.073	0.007	0.078	0.077	0.008	0.076	0.075
50	20	0.004	0.045	0.077	0.004	0.049	0.080	0.004	0.049	0.081
50	30	0.008	0.043	0.071	0.007	0.047	0.071	0.008	0.046	0.072
50	50	0.003	0.041	0.065	0.002	0.045	0.066	0.003	0.045	0.063
100	20	0.001	0.022	0.071	0.002	0.025	0.070	0.001	0.024	0.072
100	30	0.003	0.022	0.067	0.003	0.024	0.068	0.003	0.023	0.066
100	50	0.001	0.021	0.062	0.000	0.023	0.059	0.001	0.022	0.060
100	100	0.002	0.021	0.061	0.003	0.023	0.062	0.003	0.022	0.062
200	20	0.000	0.011	0.072	0.000	0.012	0.071	0.000	0.012	0.066
200	30	0.003	0.011	0.062	0.003	0.012	0.061	0.003	0.011	0.059
200	50	0.002	0.010	0.059	0.002	0.011	0.057	0.002	0.011	0.057
200	100	0.002	0.011	0.066	0.001	0.011	0.058	0.001	0.011	0.059

TCCEP		MD			MCD			MVE		
N	T	Bias	RMSE	Size	Bias	RMSE	Size	Bias	RMSE	Size
20	20	0.009	0.242	0.101	0.004	0.249	0.106	0.005	0.244	0.104
30	20	-0.001	0.182	0.085	0.002	0.189	0.091	0.001	0.185	0.088
30	30	0.002	0.169	0.078	0.006	0.173	0.084	0.005	0.171	0.081
50	20	-0.003	0.115	0.071	-0.003	0.118	0.072	-0.004	0.116	0.072
50	30	0.005	0.109	0.076	0.004	0.115	0.072	0.004	0.113	0.071
50	50	0.000	0.105	0.069	0.001	0.109	0.068	-0.001	0.108	0.066
100	20	-0.001	0.060	0.058	-0.001	0.063	0.062	-0.001	0.063	0.065
100	30	-0.001	0.057	0.061	0.000	0.060	0.060	0.001	0.058	0.057
100	50	0.003	0.055	0.055	0.002	0.057	0.055	0.002	0.056	0.055
100	100	0.001	0.053	0.063	0.000	0.056	0.060	0.001	0.054	0.059
200	20	-0.001	0.031	0.055	-0.002	0.033	0.056	-0.002	0.032	0.053
200	30	0.001	0.029	0.049	0.001	0.030	0.050	0.001	0.030	0.048
200	50	0.001	0.029	0.061	0.001	0.030	0.057	0.000	0.029	0.060
200	100	0.004	0.028	0.053	0.003	0.029	0.054	0.003	0.029	0.050

Uncontaminated DGP

Table 4.6: Experiment 2 - Effects of Outliers on Small Sample Properties of the CCEMG and CCEP Estimators (β_2)

TCCEMG

		MD			MCD			MVE		
N	T	Bias	RMSE	Size	Bias	RMSE	Size	Bias	RMSE	Size
20	20	0.123	0.159	0.110	0.040	0.139	0.092	0.056	0.144	0.089
30	20	0.054	0.089	0.095	0.029	0.092	0.086	0.033	0.091	0.082
30	30	0.046	0.085	0.086	0.023	0.087	0.084	0.026	0.087	0.082
50	20	0.086	0.064	0.117	0.032	0.055	0.076	0.034	0.055	0.079
50	30	0.079	0.061	0.113	0.026	0.053	0.076	0.028	0.053	0.078
50	50	0.078	0.059	0.107	0.027	0.052	0.078	0.028	0.052	0.083
100	20	0.168	0.062	0.302	0.030	0.029	0.086	0.033	0.029	0.086
100	30	0.169	0.060	0.311	0.032	0.028	0.075	0.034	0.028	0.078
100	50	0.170	0.060	0.312	0.031	0.027	0.078	0.032	0.027	0.074
100	100	0.164	0.057	0.297	0.026	0.027	0.073	0.029	0.027	0.074
200	20	0.189	0.055	0.566	0.031	0.016	0.092	0.033	0.016	0.095
200	30	0.188	0.053	0.573	0.030	0.014	0.078	0.032	0.014	0.079
200	50	0.185	0.050	0.567	0.028	0.014	0.075	0.030	0.014	0.078
200	100	0.183	0.049	0.574	0.028	0.014	0.072	0.029	0.014	0.077

TCCEP

		MD			MCD			MVE		
N	T	Bias	RMSE	Size	Bias	RMSE	Size	Bias	RMSE	Size
20	20	0.187	0.513	0.094	0.042	0.299	0.111	0.069	0.329	0.111
30	20	0.081	0.271	0.073	0.034	0.223	0.083	0.040	0.222	0.083
30	30	0.064	0.239	0.076	0.026	0.201	0.087	0.030	0.202	0.085
50	20	0.194	0.355	0.050	0.048	0.152	0.061	0.052	0.155	0.063
50	30	0.185	0.335	0.051	0.040	0.145	0.067	0.043	0.146	0.067
50	50	0.172	0.307	0.060	0.038	0.141	0.076	0.043	0.141	0.077
100	20	0.581	0.866	0.058	0.058	0.094	0.059	0.065	0.098	0.055
100	30	0.574	0.841	0.058	0.064	0.089	0.070	0.066	0.090	0.066
100	50	0.590	0.866	0.061	0.054	0.084	0.052	0.057	0.084	0.055
100	100	0.586	0.863	0.052	0.052	0.080	0.064	0.056	0.082	0.063
200	20	0.809	1.096	0.301	0.070	0.054	0.061	0.075	0.055	0.060
200	30	0.797	1.056	0.306	0.064	0.049	0.059	0.070	0.050	0.061
200	50	0.794	1.032	0.311	0.061	0.046	0.057	0.066	0.047	0.062
200	100	0.787	1.021	0.298	0.057	0.044	0.059	0.060	0.045	0.062

5% of cross-section units are contaminated

Table 4.7: Experiment 3 - Effects of Outliers on Small Sample Properties of the CCEMG and CCEP Estimators (β_2)

TCCEMG

		MD			MCD			MVE		
N	T	Bias	RMSE	Size	Bias	RMSE	Size	Bias	RMSE	Size
20	20	0.332	0.306	0.203	0.084	0.167	0.111	0.112	0.180	0.114
30	20	0.345	0.244	0.275	0.081	0.114	0.101	0.091	0.118	0.106
30	30	0.333	0.229	0.266	0.068	0.107	0.082	0.079	0.110	0.086
50	20	0.403	0.243	0.522	0.074	0.071	0.102	0.083	0.073	0.103
50	30	0.400	0.236	0.523	0.072	0.068	0.101	0.079	0.070	0.103
50	50	0.399	0.232	0.523	0.065	0.062	0.073	0.072	0.065	0.085
100	20	0.452	0.250	0.866	0.066	0.037	0.110	0.074	0.038	0.119
100	30	0.455	0.251	0.881	0.067	0.036	0.117	0.073	0.037	0.120
100	50	0.452	0.244	0.890	0.065	0.034	0.087	0.070	0.035	0.096
100	100	0.450	0.241	0.905	0.062	0.033	0.088	0.068	0.034	0.098
200	20	0.470	0.248	0.986	0.069	0.021	0.145	0.076	0.022	0.164
200	30	0.472	0.246	0.994	0.067	0.020	0.146	0.074	0.021	0.150
200	50	0.467	0.239	0.995	0.062	0.019	0.128	0.068	0.020	0.144
200	100	0.464	0.236	0.997	0.061	0.018	0.112	0.066	0.019	0.126

TCCEP

		MD			MCD			MVE		
N	T	Bias	RMSE	Size	Bias	RMSE	Size	Bias	RMSE	Size
20	20	0.649	1.658	0.082	0.094	0.372	0.111	0.140	0.440	0.104
30	20	0.776	1.694	0.067	0.109	0.305	0.091	0.126	0.325	0.091
30	30	0.783	1.733	0.063	0.101	0.294	0.077	0.123	0.322	0.082
50	20	1.239	2.870	0.108	0.128	0.235	0.062	0.145	0.259	0.065
50	30	1.235	2.805	0.116	0.126	0.216	0.073	0.139	0.235	0.071
50	50	1.241	2.804	0.111	0.106	0.196	0.066	0.121	0.216	0.066
100	20	1.678	3.876	0.616	0.140	0.151	0.068	0.157	0.167	0.068
100	30	1.681	3.858	0.641	0.135	0.139	0.067	0.146	0.148	0.068
100	50	1.677	3.829	0.646	0.125	0.121	0.062	0.137	0.132	0.065
100	100	1.666	3.785	0.658	0.118	0.115	0.064	0.130	0.125	0.072
200	20	1.914	4.315	0.988	0.154	0.094	0.092	0.167	0.101	0.097
200	30	1.921	4.331	0.992	0.141	0.083	0.088	0.155	0.090	0.095
200	50	1.913	4.284	0.991	0.132	0.077	0.098	0.144	0.084	0.103
200	100	1.901	4.226	0.990	0.129	0.071	0.095	0.139	0.077	0.100

10% of cross-section units are contaminated

Table 4.8: Experiment 4 - Effects of Outliers on Small Sample Properties of the CCEMG and CCEP Estimators (β_2)

TCCEMG

		MD			MCD			MVE		
N	T	Bias	RMSE	Size	Bias	RMSE	Size	Bias	RMSE	Size
20	20	0.851	0.966	0.585	0.160	0.233	0.126	0.222	0.279	0.154
30	20	0.848	0.882	0.800	0.182	0.178	0.134	0.217	0.202	0.159
30	30	0.847	0.874	0.810	0.156	0.161	0.122	0.188	0.180	0.142
50	20	0.894	0.900	0.979	0.171	0.115	0.162	0.205	0.136	0.198
50	30	0.892	0.893	0.983	0.161	0.109	0.156	0.192	0.127	0.190
50	50	0.901	0.904	0.987	0.162	0.110	0.152	0.191	0.126	0.186
100	20	0.939	0.939	1.000	0.169	0.074	0.238	0.196	0.086	0.294
100	30	0.944	0.945	1.000	0.164	0.069	0.235	0.190	0.081	0.287
100	50	0.949	0.950	1.000	0.157	0.065	0.215	0.180	0.075	0.256
100	100	0.946	0.942	1.000	0.153	0.063	0.209	0.178	0.073	0.260
200	20	0.953	0.939	1.000	0.161	0.048	0.376	0.186	0.058	0.449
200	30	0.955	0.942	1.000	0.156	0.046	0.368	0.179	0.055	0.444
200	50	0.960	0.948	1.000	0.154	0.043	0.350	0.175	0.051	0.418
200	100	0.959	0.944	1.000	0.147	0.042	0.332	0.169	0.050	0.408

TCCEP

		MD			MCD			MVE		
N	T	Bias	RMSE	Size	Bias	RMSE	Size	Bias	RMSE	Size
20	20	2.289	7.849	0.419	0.200	0.538	0.108	0.318	0.787	0.110
30	20	2.412	7.891	0.599	0.266	0.532	0.087	0.348	0.725	0.095
30	30	2.450	8.014	0.634	0.230	0.464	0.083	0.303	0.631	0.091
50	20	2.728	8.928	0.935	0.299	0.436	0.082	0.387	0.628	0.099
50	30	2.734	8.917	0.948	0.284	0.410	0.081	0.360	0.568	0.094
50	50	2.747	8.949	0.955	0.280	0.387	0.086	0.353	0.541	0.102
100	20	2.943	9.495	1.000	0.348	0.352	0.124	0.415	0.452	0.160
100	30	2.946	9.489	1.000	0.329	0.310	0.128	0.392	0.411	0.164
100	50	2.958	9.538	1.000	0.306	0.281	0.115	0.364	0.374	0.150
100	100	2.952	9.482	1.000	0.298	0.263	0.124	0.361	0.362	0.157
200	20	3.048	9.732	1.000	0.360	0.258	0.268	0.426	0.341	0.336
200	30	3.053	9.746	1.000	0.342	0.233	0.265	0.400	0.310	0.321
200	50	3.053	9.735	1.000	0.334	0.216	0.266	0.391	0.285	0.339
200	100	3.059	9.766	1.000	0.311	0.193	0.247	0.364	0.260	0.314

20% of cross-section units are contaminated

Chapter 5

Labour Shares and Economic Growth: A Cross-Sectionally Dependent Perspective

5.1 Introduction and Review

Labour Share of Income

Labour share is the share of national income allocated to labour compensation (ILO, 2015). Interest in labour share has a long history. For comprehensive reviews of labour share, see Daron Acemoglu's *Introduction to Modern Economic Growth* (Acemoglu, 2008), Thomas Piketty's *Capital in the Twenty-First Century* (Piketty, 2014) and the papers from the International Labour Organisation (ILO). There is extensive empirical literature on labour share. Traditionally, labour share is assumed to be constant, but evidence from the 1980s onwards indicate that labour shares rise and fall with time (Piketty, 2014). Charpe et al (2019) investigated the behaviour of labour share using long historical data from the US, UK and France using wavelet analysis and revealed that labour shares rise and fall with long-run cycles of 30-50 years instead of business cycles (which are often modelled after). Since post-World War Two, labour shares have been falling from peaks in the 1950s (Piketty, 2014). Falling labour shares are often associated with rapid growth in labour productivity rather than dwindling average labour compensation. It is also associated with an increase in returns to capital relative to labour (ILO, 2015). Today, there exists many sources of data for labour share, such as the Penn World Tables (Feenstra, Inklaar and Timmer, 2015), which is computed from data adjusted for purchasing power parity (PPP). There are two major issues in modelling economic growth (and in turn, labour share). In particular, the endogeneity of technological change (see Acemoglu, 2008, p.101) is a major issue in the estimations. Techonolgical change is often unobserved, and affects numerous covariates of economic growth, such as human capital, price levels of capital and even savings rates, which in turn affect labour shares. Although technological progress can sometimes be approximated by country Research and Development (R&D) expenditure, such data is often not available, especially for developing countries and countries with socio-political instability. A common practice to bypass such a limitation is to use Method-of-Moments etimators, such as the Generalised Method of Moments (GMM) estimator (Arellano & Bond, 1990; Gagliardini et. al., 2005) or the System GMM (Sys-GMM) estimator (Blundell & Bond, 1998) that can use lags as instruments. However, alongside traditional least squares-based estimation, such practices assume that the individual units are not correlated. Cross-section dependence is more prevalent in data than we would like to admit. Yet, cross-section dependence is another major issue in panel data that is currently avoided by most practitioners. Together with cross-section dependence, the presence of outliers almost always distorts estimates and the results of significance tests.

Outline

The purpose of this paper is to examine the relationship between the labour share of income and economic growth in light of cross-section dependence and outliers. The next section reviews the econometrics and presents the methodology. The following section examines the data and estimation results in the light of cross-section dependence and outliers. The final section concludes.

5.2 Theory and Methodology

The Econometrics of Labour Share

The emprical framework is based on the neoclassical growth model

$$DY = f(Y, Y*) \tag{5.1}$$

Where DY is the growth rate of per capita output Y, and $Y*$ is the target level of per capita output (see Barro, 2000; Solow, 1956 for details). The discrete-time Solow model based on this framework is

$$\Delta y_{it} = \beta_0 + \beta_1 y_{i,t-1} + \epsilon_{it} \tag{5.2}$$

where y is the log of per capita output (see Acemoglu, 2008; Solow, 1956; Barro, 2000 for examples and details) and $\Delta y_{it} = y_{it} - y_{i,t-1}$ is the corresponding growth rate. Under this model, Barro (1991) and Barro & Sala-i-Martin (1992) obtained conditional convergence (negative β_1) for OECD countries, but a positive β_1 for worldwide data, implying non-covergence of growth rates for the entire world. This is often accounted for as cross-country differences and unobserved factors (see Acemoglu, 2008 for a detailed discussion). (5.2) can be augmented with individual effects β_{i0} to capture cross-country differences. The contemporary Barro growth regression developed from (5.2) is

$$\Delta y_{it} = \beta_{i0} + \beta_{i1} y_{i,t-1} + \mathbf{x}_{it}' \boldsymbol{\beta}_i + \epsilon_{it} \tag{5.3}$$

where $\mathbf{x}_{it}$ is a vector of covariates and individual effects β_{i0} and β_{1i} are captured. Here, we also allow for heterogeneous coefficients to account for the variability among countires. Applications of the Barro regression in literature are widespread. An example here includes Matthieu Charpe, Slim Bridji and Peter McAdam(Charpe et. al, 2019) in their paper on long run data and the cyclical nature of labour share. In addition to least squares regression on more than a century of data each from the UK, US and France, the paper includes wavelet analysis to reveal that labour share rises and falls with a long-run cycle of 30 to 50 years, instead of with the business cycle. We address three major issues in this model: the endogeneity of techonolgical change and unobserved covariates(Acemoglu, 2008); the presence of cross-section dependence in the data and the presence of outliers. We address cross-section dependence with the Common Correlated Effects (CCE; Pesaran, 2006) estimator. We trim outliers using an outlier-robust depth statistic.

Cross-Section Dependence

Due to the interconnectedness of today's economies, it is impossible to assume that cross-section dependence is not present in data. Most economic models assume closed economies, but terms of trade and globalisation already ensure that no economy is an island (apart from a marginal few isolated regimes or tribes). Furthermore, although data is widely available today, there are still factors that are unobserved, such as ability over education. Unobserved factors may simply be due to data that is not recorded, unavailable or highly confidential. Unobserved endogeneous factors potentially affect all covariates, but they can be captured in a factor model. The general equation factor model is given in Pesaran (2006) as

$$\begin{aligned} y_{it} &= \boldsymbol{\alpha}_i' \mathbf{d}_t + \boldsymbol{\beta}_i' \mathbf{x}_{it} + \boldsymbol{\gamma}_i' \mathbf{f}_t + e_{it} \\ \mathbf{x}_{it} &= \mathbf{A}_i' \mathbf{d}_t + \boldsymbol{\Gamma}_i' \mathbf{f}_t + \mathbf{v}_{it} \end{aligned} \tag{5.4}$$

where $\mathbf{d}_t$ is a $n \times 1$ vector of observed common effects, including deterministics such as intercepts and seasonal dummies, $\mathbf{x}_{it}$ is a $k \times 1$ vector of observed individual-specific regressors on the ith cross-section at time t and $\mathbf{f}_t$ is a $m \times 1$ vector of unobserved effects. The CCE Mean Group (CCEMG) estimator is given, as in Pesaran (2006):

$$\hat{\mathbf{b}_i} = (\mathbf{X}_i'\bar{\mathbf{M}}_w\mathbf{X}_i)^{-1}(\mathbf{X}_i'\bar{\mathbf{M}}_w\mathbf{y}_i) \tag{5.5}$$

$$\hat{\boldsymbol{\beta}}_{CCEMG} = \frac{1}{N}\sum_{i=1}^{N}\hat{\mathbf{b}_i} \tag{5.6}$$

$\bar{\mathbf{M}}_w$ is the annihilator matrix that proxies the unobserved effects (see Pesaran, 2006, for more details). The CCE Pooled (CCEP) estimator is

$$\hat{\mathbf{b}}_P = \left(\sum_{i=1}^{N}\theta_i\mathbf{X}_i'\bar{\mathbf{M}}_w\mathbf{X}_i\right)^{-1}\left(\sum_{i=1}^{N}\theta_i\mathbf{X}_i'\bar{\mathbf{M}}_w\mathbf{y}_i\right) \tag{5.7}$$

The CCE estimators produce consistent estimates in the presence of strong cross-section dependence. (5.4) can include the Barro regression (5.3), which we will do in the following subsection.

The Model

We begin with a Barro growth regression with individual and time effects

$$\Delta y_{it} = a_i + \delta_t + \beta_{1i}labsh_{i,t-1} + \boldsymbol{\beta}_i'\mathbf{x}_{it} + e_{it} \tag{5.8}$$

where y_{it} is the log per capita output, $labsh$ is the labour share of income as a percentage of GDP and $\mathbf{x}_{it}$ is a vector of covariates. To account for strong cross-section dependence, we fit the errors e_{it} with a strong cross-sectional dependence structure

$$e_{it} = \boldsymbol{\gamma}_i'\mathbf{f}_t + \epsilon_{it} \tag{5.9}$$

where $\mathbf{f}_t$ is the vector of unobserved common effects as in (5.4) and idiosyncratic error ϵ_{it}. Assuming that δ_t is a mean-zero stationary process with absolute summable autocovariances, we can include the time effect δ_t in the vector $\mathbf{f}_t$ and rewrite (5.8) as

$$\Delta y_{it} = a_i + \beta_{1i}labsh_{i,t-1} + \boldsymbol{\beta}_i'\mathbf{x}_{it} + \boldsymbol{\gamma}_i'\mathbf{f}_t + \epsilon_{it} \tag{5.10}$$

The techonological change A_i (see Acemoglu, 2008; Piketty, 2014 for details) is unobserved and correlated with savings rates for capital, which in turn affects the labour share. Hence, we assume that one of the components (without loss of generality (WLOG), the first component) of $\mathbf{f}_t$ is

$$A_i = \gamma_{1i} f_{1t} \tag{5.11}$$

Together with (5.10), we estimate this factor model with the CCE estimators. However, the CCE estimators are not robust to outliers. As the CCE is a LS-based estimator, it takes only one suffciently large outlier to distort estimates. Furthermore, outliers are almost always present in data. The next subsection addresses the issue of outliers.

Outliers

Lee and Sul (Lee & Sul, 2022) proposed the Trimmed CCE (TCCE) estimators for both the MG and Pooled versions of the CCE. In their paper, Lee and Sul trimmed cross-section units that have unusually large sammple variances based on a depth statistic (for a review, see Lee & Sul, 2020). To compute the depth statistic, the individual unit sample variances are first computed:

$$\mathbf{\Sigma}_i = \frac{1}{N}\sum_{i=1}^{N}(x_i - \mu_i)^2$$

and the vector of sample variances for the covariates of unit i are

$$\begin{aligned} \mathbf{v}_i &= (v_{1i}, v_{2i}, \ldots, v_{ki}) \quad (5.12) \\ &= diag(\mathbf{\Sigma}_i) \quad (5.13) \end{aligned}$$

(WLOG, when we refer to the covariates $\mathbf{x}_{it}$ here, we can regard the first component as $x_{1,it} = labsh_{i,t-1}$). The depth statistic D_i for cross-section i is computed as

$$D_i = \frac{1}{1 + MD_i} \tag{5.14}$$

where MD_i the Mahalanobis Distance (Mahalanobis, 1936; Maesschalck, Jouan-Rimbaud & Massart, 2000 for a recent version) of the entire sample of variance vectors:

$$MD_i = (\mathbf{v}_i - \bar{\mathbf{v}})'\mathbf{\Sigma}^{-1}(\mathbf{v}_i - \bar{\mathbf{v}}) \tag{5.15}$$

where $\bar{\mathbf{v}}$ is the sample mean over all the sample variance vectors $\mathbf{v}_i$. The Depth statistic is $\chi^2(1)$-distributed and units with sufficiently small D_i are flagged as outlying. In this paper, we follow the practince in Lee & Sul (2020) and trim the smallest 20% of depths as outliers. The TCCEMG estimator is then given by

$$\hat{\boldsymbol{\beta}}_{TCCEMG} = \frac{1}{n_G} \sum_{i \in G} \hat{\mathbf{b}}_i \tag{5.16}$$

where $n_G = |G|$ is the number of units in G, the set of units not trimmed by the depth statistic. $\hat{\mathbf{b}}_i$ is the usual individual CCEMG estimate as in (5.5). Similarly, the trimmed pooled CCE estimator (TCCEP) is

$$\boldsymbol{\beta}_{TCCEP} = \left(\frac{1}{n_G} \sum_{i \in G} \mathbf{X}_i' \bar{\mathbf{M}}_w \mathbf{X}_i \right)^{-1} \left(\frac{1}{n_G} \sum_{i \in G} \mathbf{X}_i' \bar{\mathbf{M}}_w \mathbf{y}_i \right) \tag{5.17}$$

with emphasis on the equal weights $\theta_i = \frac{1}{n_G}$, although any suitable weights may be used, both in the mean group and pooled versions. Thus, the Trimmed CCE estimators are simply the usual CCE estimators with outlying cross-section units trimmed. However, the Mahalanobis distance is not robust to outliers. Furthermore, it suffers from the masking effect (see Rousseeuw & Leroy, 1987 for example). We therefore propose to replace the Mahalanobis distance with robust location and scale estimates. We select the Minimum Covariance Determinant (MCD; Hubert, Debruyne and Rousseeuw, 2018) and the Minimum Volume Ellipsoid (MVE; Van Aelst and Rousseeuw, 2009).

5.3 Data and Results

Data and Choice of variables

For the control variables, we selected $\mathbf{x}_{it} = (agffgrowth_{i,t-1}, FixedCapitalGrowth_{i,t-1})$, where $agffgrowth$ is the growth rate of the output of agriculture, forestry and fishing, $FixedCapitalGrowth$ is the growth rate of fixed capital formation. y_{it} is the log of output per capita (in 2015 US dollars). Data for the labour share of income, $labsh$, was obtained from Penn World Tables version 10.01. All other data was obtained from World Bank's World Development Indicators (Beegle, et al. (2023)). We used data available from 1990 to 2019 for a panel of 62 countries. These include OECD, emerging markets and developing countries.

Cross-Section Dependence

The CD-test (Pesaran, 2015) tests residuals for cross-section dependence. The CD-statistic is computed as

$$CD = \sqrt{\frac{2T}{N(N-1)}} \sum_{i=1}^{N-1} \sum_{j=i+1}^{N} \hat{\rho}_{ij} \tag{5.18}$$

where $\hat{\rho}_{ij}$s are the correlations between (preliminary) residual estimates:

$$\hat{\rho}_{ij} = \hat{\rho}_{ji} = \frac{\sum_{t=1}^{T} \hat{e}_{it}\hat{e}_{jt}}{\sqrt{\sum_{t=1}^{T} \hat{e}_{it}}\sqrt{\sum_{t=1}^{T} \hat{e}_{jt}}} \tag{5.19}$$

Using preliminary within-groups residuals, we obtained a CD-test result of $CD = 25.745$, indicating the presence of strong cross-section dependence in the data. This is expected in theory and econometrics. As mentioned in the previous section, endogeneous techonological progress possibly affects all covariates: contemporary agricutural processes utilises much machinery for tremendously increased yields; similarly more advanced trawlers dredge more live fish than traditional boats. Furthermore, institutions and firms generally utilise technological progress to increase capital gains (Acemoglu, 2008). Without available worldwide data, we henceforth account for techonological progress as an unobserved factor in (5.4). Other possible factors include fertility rates and education, but we shall treat them as unobserved factors in $\mathbf{f}_t$ for now. We start by performing the following regressions: (1) Within groups (FE); (2) CCEMG; (3) CCEP (3); (4)CCEPnw and (5)CCEPwpn. CCEPnw is the CCEP estimator with the alternate Newey-West variance estimator (see Pesaran, 2006). Similarly, CCEPwpn is the CCEP estimator with an alternative variance estimator by Joakim Westerlund, Yana Petrova and Milda Norkute (Westerlund, Petrova & Norkute, 2019).

Estimations and Results

Table (5.3) presents the regression results from the standard Within-Groups estimator and the basic CCE estimators. From the within-goups residuals, we obtained a CD-statistic of 25.745, which corresponds to a p-value close to zero and much smaller than a 1% significance level. Thus, by rejecting the null of the CD-test, we indicate the presence of strong cross-sectional dependence, despite the within-groups estimates showing high significance levels. On the other hand, the CCE estimators have not fared well in this table, obtaining barely any significant results. In the following tables, we treat the data for outliers and perform trimmed regressions. Table (5.4) presents the same regressions with Lee & Sul's (2022) trimming by Mahalanobis Distance (MD). The results are similar to the standard regressions. Tables (5.5) and (5.6) present the same regressions under MCD trimming and MVE trimming respectively, showing significant results and global convergence for economic growth rates: negative and significant β_1, the coefficient of $y_{i,t-1}$, of approximately -0.05. These regressions, however, exclude the flagged outliers that potentially do not fit the bulk of the data. Under non-robust trimming (and other non-robust outlier diagnostics), some outliers can be masked by others, making the former appear to be non-outlying. This is known as the masking effect (see Rousseeuw & Leroy, 1987 for examples). These masked outliers, being undetected, continue to affect the estimation results, as also illustrated in our results. Under robust trimming, coefficient on labour share has on economic growth is smaller than under non-robust trimming or without any treatment for outliers or cross-section dependence. The results are in line with Charpe et. al (2019), with labour share having a negative effect for periods of 30 years or less. Our results seem to indicate that the negative effect is much greater than it would seem under standard regressions, perhaps mitigated by unobserved effects or outliers.

Table 5.1: Summary Statistics for GDP Growth - World

	mean	sd	min	max
Argentina	0.015	0.056	-0.119	0.093
Australia	0.016	0.013	-0.017	0.037
Austria	0.015	0.016	-0.040	0.036
Barbados	0.003	0.029	-0.060	0.057
Benin	0.016	0.018	-0.014	0.055
Botswana	0.024	0.040	-0.094	0.101
Brazil	0.010	0.028	-0.061	0.065
Bulgaria	0.020	0.055	-0.136	0.082
Burundi	-0.014	0.035	-0.093	0.037
Cameroon	0.001	0.035	-0.105	0.042
Chile	0.033	0.026	-0.022	0.093
China	0.085	0.025	0.024	0.136
Colombia	0.020	0.023	-0.058	0.059
Costa Rica	0.026	0.020	-0.021	0.068
Cyprus	0.019	0.034	-0.063	0.065
Denmark	0.013	0.018	-0.054	0.050
Egypt	0.023	0.016	-0.011	0.053
Eswatini	0.024	0.033	-0.010	0.175

Fiji	0.018	0.029	-0.036	0.080
Finland	0.014	0.033	-0.085	0.060
France	0.011	0.013	-0.034	0.032
Gabon	-0.005	0.034	-0.111	0.034
Guatemala	0.015	0.011	-0.014	0.043
Guinea	0.019	0.022	-0.034	0.079
Honduras	0.012	0.025	-0.060	0.044
India	0.046	0.020	-0.010	0.071
Indonesia	0.035	0.036	-0.144	0.066
Iraq	0.055	0.218	-0.650	0.540
Jordan	0.007	0.030	-0.049	0.081
Kenya	0.008	0.023	-0.040	0.052
Korea Rep	0.045	0.033	-0.058	0.107
Kyrgyz Republic	0.004	0.071	-0.201	0.087
Malaysia	0.037	0.034	-0.097	0.072
Mauritania	0.005	0.044	-0.064	0.150
Mauritius	0.038	0.014	0.009	0.071
Mexico	0.011	0.029	-0.078	0.052
Mongolia	0.031	0.056	-0.103	0.152
Morocco	0.024	0.036	-0.069	0.108
Namibia	0.017	0.030	-0.040	0.106
Netherlands	0.016	0.018	-0.042	0.043
New Zealand	0.014	0.021	-0.058	0.052
Norway	0.015	0.017	-0.030	0.047
Panama	0.039	0.030	-0.014	0.100
Paraguay	0.014	0.031	-0.043	0.096
Peru	0.028	0.036	-0.070	0.102
Rwanda	0.037	0.120	-0.475	0.375
Saudi Arabia	0.009	0.043	-0.057	0.113
Senegal	0.009	0.022	-0.037	0.044
Sierra Leone	0.010	0.085	-0.223	0.210
South Africa	0.006	0.022	-0.046	0.043

Sweden	0.015	0.024	-0.052	0.051
Thailand	0.036	0.037	-0.087	0.096
Togo	0.006	0.056	-0.168	0.124
Tunisia	0.025	0.021	-0.030	0.056
Turkiye	0.031	0.045	-0.071	0.095
Zimbabwe	0.00019	0.084	-0.185	0.181

Table 5.2: Summary Statistics for Labour Share of Income - World

	mean	sd	min	max
Argentina	0.432	0.0780	0.310	0.544
Australia	0.596	0.0158	0.569	0.626
Austria	0.593	0.0254	0.551	0.633
Barbados	0.751	0.0000	0.751	0.751
Benin	0.621	0.0133	0.603	0.650
Botswana	0.295	0.0296	0.278	0.383
Brazil	0.548	0.0239	0.492	0.581
Bulgaria	0.472	0.0434	0.386	0.536
Burundi	0.701	0.0920	0.606	0.837
Cameroon	0.510	0.0146	0.504	0.555
Chile	0.459	0.0327	0.384	0.516
China	0.577	0.0171	0.548	0.606
Colombia	0.481	0.0161	0.448	0.505
Costa Rica	0.577	0.0039	0.575	0.589
Cyprus	0.534	0.0374	0.484	0.601
Denmark	0.637	0.0127	0.616	0.672
Egypt	0.376	0.0275	0.314	0.417
Eswatini	0.612	0.0000	0.612	0.612
Fiji	0.521	0.0448	0.489	0.605
Finland	0.595	0.0373	0.559	0.708
France	0.619	0.0043	0.607	0.624
Gabon	0.312	0.0398	0.275	0.377
Guatemala	0.520	0.0256	0.483	0.546

Guinea	0.309	0.0580	0.258	0.420
Honduras	0.585	0.0137	0.562	0.613
India	0.561	0.0637	0.481	0.680
Indonesia	0.452	0.0091	0.439	0.466
Iraq	0.219	0.0709	0.090	0.322
Jordan	0.485	0.0151	0.451	0.502
Kenya	0.595	0.0426	0.520	0.686
Korea Rep	0.527	0.0239	0.496	0.573
Kyrgyz Republic	0.594	0.0878	0.480	0.805
Malaysia	0.334	0.0369	0.306	0.396
Mauritania	0.460	0.0543	0.404	0.517
Mauritius	0.467	0.0463	0.425	0.550
Mexico	0.386	0.0243	0.356	0.434
Mongolia	0.400	0.0312	0.328	0.460
Morocco	0.498	0.0099	0.474	0.508
Namibia	0.579	0.0517	0.511	0.660
Netherlands	0.614	0.0237	0.581	0.669
New Zealand	0.550	0.0127	0.528	0.571
Norway	0.518	0.0301	0.454	0.571
Panama	0.392	0.0621	0.296	0.465
Paraguay	0.461	0.0233	0.417	0.514
Peru	0.436	0.0104	0.420	0.460
Rwanda	0.741	0.0000	0.741	0.741
Saudi Arabia	0.296	0.0336	0.199	0.327
Senegal	0.487	0.0024	0.481	0.497
Sierra Leone	0.537	0.0063	0.527	0.550
South Africa	0.557	0.0298	0.513	0.603
Sweden	0.542	0.0214	0.505	0.597
Thailand	0.658	0.0085	0.637	0.671
Togo	0.806	0.0000	0.806	0.806
Tunisia	0.510	0.0205	0.472	0.552
Turkiye	0.400	0.0237	0.375	0.466

Zimbabwe	0.533	0.0000	0.533	0.533

Table 5.3: Standard Regressions

	FE	CCEMG	CCEMGnw	CCEMGwpn	CCEP
Fixed Capital Growth	0.0294***	0.105***	0.0353	0.0353**	0.0353*
	(5.58)	(6.27)	(1.06)	(3.07)	(2.54)
agffgrowth	0.178***	0.183***	0.145***	0.145***	0.145***
	(14.34)	(4.15)	(3.65)	(5.30)	(4.30)
$y_{i,t-1}$	-0.0107***	-0.0482***	-0.0830***	-0.0830***	-0.0830***
	(-5.23)	(-6.84)	(-3.59)	(-5.87)	(-11.19)
labsh	-0.119***	-0.418**	-0.123	-0.123	-0.123
	(-3.87)	(-2.96)	(-1.02)	(-1.85)	(-1.47)
Constant	0.164***	0.120	0.0480	0.0480	0.0480
	(6.86)	(1.35)	(0.40)	(0.40)	(0.40)
Observations	1622	1622	1622	1622	1622

t statistics in parentheses

CD = 25.745, p = 0

* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Table 5.4: TCCEMG(Lee and Sul, 2020)

	FE	CCEMG	CCEMGnw	CCEMGwpn	CCEP
Fixed Capital Growth	0.0521***	0.118***	0.0485	0.0485**	0.0485*
	(8.75)	(6.43)	(1.28)	(2.88)	(2.43)
agffgrowth	0.151***	0.165***	0.116**	0.116***	0.116**
	(13.11)	(3.53)	(2.77)	(4.18)	(3.12)
$y_{i,t-1}$	-0.00380*	-0.0445***	-0.0715***	-0.0715***	-0.0715***
	(-2.09)	(-6.87)	(-5.63)	(-10.66)	(-7.82)
labsh	-0.115***	-0.354*	-0.0942	-0.0942	-0.0942
	(-4.21)	(-2.48)	(-0.76)	(-1.38)	(-0.97)
Constant	0.107***	-0.00901	2.10e-16	2.10e-16	2.10e-16
	(4.82)	(-0.10)	(0.00)	(0.00)	(0.00)
Observations	1508	1	1	1	1

t statistics in parentheses

* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Outliers

Table 5.7 lists the outliers flagged by the estimators MD, MCD and MVE. From this table, we identify eight outlying countries which have been detected by all three estimators. For the last three coluns, we

Table 5.5: TCCEMG(MCD)

	FE	CCEMG	CCEMGnw	CCEMGwpn	CCEP
Fixed Capital Growth	0.163***	0.139***	0.145***	0.145***	0.145***
	(16.99)	(6.92)	(3.93)	(5.25)	(3.90)
agffgrowth	0.135***	0.180**	0.115**	0.115***	0.115***
	(11.62)	(2.85)	(3.05)	(5.77)	(3.75)
$y_{i,t-1}$	-0.00616***	-0.0371***	-0.0525***	-0.0525***	-0.0525***
	(-3.40)	(-5.62)	(-4.21)	(-8.32)	(-6.19)
labsh	-0.0811*	-0.129	-0.200*	-0.200***	-0.200***
	(-2.29)	(-1.48)	(-2.47)	(-4.23)	(-3.59)
Constant	0.111***	-0.0462	3.72e-16	3.72e-16	3.72e-16
	(4.35)	(-0.33)	(0.00)	(0.00)	(0.00)
Observations	1131	1	1	1	1

t statistics in parentheses

* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Table 5.6: TCCEMG(MVE)

	FE	CCEMG	CCEMGnw	CCEMGwpn	CCEP
Fixed Capital Growth	0.156***	0.135***	0.143***	0.143***	0.143***
	(16.74)	(6.92)	(3.88)	(4.89)	(3.89)
agffgrowth	0.132***	0.168**	0.114**	0.114***	0.114***
	(11.62)	(2.95)	(2.94)	(4.88)	(3.45)
$y_{i,t-1}$	-0.00623***	-0.0368***	-0.0539***	-0.0539***	-0.0539***
	(-3.43)	(-5.47)	(-4.35)	(-7.49)	(-5.59)
labsh	-0.0842*	-0.0990	-0.194*	-0.194***	-0.194**
	(-2.49)	(-1.10)	(-2.27)	(-3.92)	(-3.27)
Constant	0.113***	0.00729	-9.11e-16	-9.11e-16	-9.11e-16
	(4.48)	(0.05)	(-0.00)	(-0.00)	(-0.00)
Observations	1160	1	1	1	1

t statistics in parentheses

* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

indicate with a 1 if the estimator has flagged the country as an outlier, and 0 otherwise. These outliers are not masked by other outliers in the non-robust MD: Singapore; Sudan; Mauritania; Morocco; Zimbabwe; China; Ecuador and Iraq. Why would these countries be outlying? For example, in Acemoglu (2008), countries like Singapore have experienced tremendous economic growth post World War Two, whereas countries like Nigeria have not, depsite being geographically more advantageous. In the same book, Acemoglu cites institutions as a possible reason. Another example is that Iraq could possibly contain numerous structural breaks due to decades of war. Clearly, a nonrobust estimator may not detect all the outliers (see Hubert, Debruyne & Rousseeuw, 2018 for more examples). For all three estimators MD, MCD and MVE, no OECD countries were flagged as outliers. We shall leave such discussions for another time.

Table 5.7: List of Outliers

Country	Class	MD	MCD	MVE	Out(MD)	Out(MCD)	Out(MVE)
Singapore	AE	4.749	13.339	13.263	1	1	1
Burundi	D	3.261	15.985	4.082	0	1	1
Guinea	D	2.476	9.823	3.807	0	1	0
Sierra Leone	D	2.903	9.735	7.409	0	1	1
Sudan	D	4.404	16.380	9.353	1	1	1
Mauritania	D	5.227	17.893	14.963	1	1	1
Dominican Republic	D	2.361	10.021	4.667	0	1	1
Panama	EM	2.508	9.806	8.394	0	1	1
Saudi Arabia	EM	1.906	5.678	2.411	0	1	0
Eswatini	EM	2.992	6.837	7.330	0	1	1
Honduras	EM	2.095	6.754	3.310	0	1	0
India	EM	2.554	4.764	4.048	0	1	1
Indonesia	EM	2.474	4.697	2.958	0	1	0
Kyrgyz Republic	EM	3.174	11.516	9.517	0	1	1
Mongolia	EM	3.007	8.015	7.281	0	1	1
Morocco	EM	6.221	9.409	10.097	1	1	1
Zimbabwe	EM	4.374	15.381	8.755	1	1	1
Argentina	EM	2.534	12.109	3.463	0	1	0
Botswana	EM	2.864	11.534	5.383	0	1	1
Bulgaria	EM	2.231	5.839	5.833	0	1	1
China	EM	4.425	10.901	8.634	1	1	1
Ecuador	EM	5.537	21.455	10.598	1	1	1
Iraq	EM	7.066	24.793	28.472	1	1	1
Malaysia	EM	3.220	11.225	9.786	0	1	1
Namibia	EM	1.661	4.368	2.526	0	1	0
Thailand	EM	1.610	6.032	4.671	0	1	1

Key
AE - Advanced Economy
EM - Emerging Market
D - Developing Country
Out(T) - Outliers flagged by estimator T

5.4 Conclusions

The CCE estimators produce consistent estimates in the presence of strong cross-section dependence, but they are very sensitive to outliers. The current outlier diagnostic in the literature trims out entire outlying cross-section units. This detects outliers on the aggregate level, but does not account for outliers within the individual units, unless they are sufficiently large or numerous enough at the aggregate level. Outlier diagnostics exist within the literature; furthermore robust outlier diagnostics exist (see Maronna et. al., 2019, for example) to detect and deal with outliers within individual units. Finally, outliers need not always be trimmed or eliminated. Outliers may contain additional information about the data, and may contribute useful information to the existing framework. Therefore, it would greatly contribute to the literature if one-step robust CCE estimators could be developed. As for the economic growth literature, perhaps there are additional factors to the framework that can allow for global convergence.

References

Acemoglu, D., 2008. *Introduction to modern economic growth.* Princeton university press.

Arellano, M. and Bond, S., 1991. Some tests of specification for panel data: Monte Carlo evidence and an application to employment equations. *The review of economic studies*, 58(2), pp.277-297.

Barro, R.J., 2000. Inequality and Growth in a Panel of Countries. *Journal of economic growth*, 5, pp.5-32.

Barro, R.J. and Sala-i-Martin, X., 1992. Convergence. *Journal of political Economy*, 100(2), pp.223-251.

Beegle, K. et al. (2023) *World Bank Open Data.* Available at: https://data.worldbank.org/ (Accessed: 28 July 2023). Blundell, R. and Bond, S., 1998. Initial conditions and moment restrictions in dynamic panel data models. *Journal of econometrics*, 87(1), pp.115-143.

Charpe, M., Bridji, S. and McAdam, P., 2020. Labor share and growth in the long run. *Macroeconomic Dynamics*, 24(7), pp.1720-1757.

De Maesschalck, R., Jouan-Rimbaud, D. and Massart, D.L., 2000. The mahalanobis distance. Chemometrics and intelligent laboratory systems, 50(1), pp.1-18.

Feenstra, Robert C., Robert Inklaar and Marcel P. Timmer (2015), The Next Generation of the Penn World Table *American Economic Review*, 105(10), 3150-3182, available for download at www.ggdc.net/pwt

Gagliardini, P., Trojani, F. and Urga, G., 2005. Robust GMM tests for structural breaks. *Journal of Econometrics*, 129(1-2), pp.139-182.

Hubert, M., Debruyne, M. and Rousseeuw, P.J., 2018. Minimum covariance determinant and extensions. *Wiley Interdisciplinary Reviews: Computational Statistics*, 10(3), p.e1421.

Lee, Y. and Sul, D., 2020. Depth-Weighted Estimation of Heterogenous Agent Panel Data Models. *Mimeo, Syracuse University.*

Lee, Y. and Sul, D., 2022. Trimmed mean group estimation. In Essays in Honor of M. Hashem Pesaran: Panel Modeling, Micro Applications, and Econometric Methodology (pp. 177-202). *Emerald Publishing Limited.*

International Labour Office, 2015. World employment and social outlook: Trends 2015. *Geneva:*

International Labour Organization.

Mahalanobis, P.C., 1936. Mahalanobis distance. *In Proceedings National Institute of Science of India* (Vol. 49, No. 2, pp. 234-256).

Maronna, R.A., Martin, R.D., Yohai, V.J. and Salibián-Barrera, M., 2019. *Robust statistics: theory and methods (with R)*. John Wiley & Sons.

Pesaran, M.H., 2006. Estimation and inference in large heterogeneous panels with a multifactor error structure. *Econometrica*, 74(4), pp.967-1012.

Piketty, T., 2014. *Capital in the twenty-first century*. Harvard University Press.

Rousseeuw, P. and Leroy, A., 1987. *Robust regression and outlier detection*. Hoboken, NJ: Wiley-Interscience.

Samuelson, P.A. and Solow, R.M., 1956. A complete capital model involving heterogeneous capital goods. *The Quarterly Journal of Economics*, 70(4), pp.537-562.

Van Aelst, S. and Rousseeuw, P., 2009. Minimum volume ellipsoid. *Wiley Interdisciplinary Reviews: Computational Statistics*, 1(1), pp.71-82.

Westerlund, J., Petrova, Y. and Norkute, M., 2019. CCE in fixed-T panels. *Journal of Applied Econometrics*, 34(5), pp.746-761.

Chapter 6

Institutions and Economic Growth: A Cross-sectionally Dependent Perspective with Additional Insights on Women's Studies

6.1 Part I - Institutions

6.1.1 Introduction

It is known in the literature that institutions have a profound effect on economic growth. Douglass North (1989) laid a solid framework for studying the effects of institutions on economic growth. In his paper, North stated that neoclassical theories of economic growth are functions of savings and population, which would hold in a closed economy without any institutions. In such a world, there would be no transaction costs incurred. Thus, institutions can either facilitate or hinder transactions by controlling the costs for the various groups of individuals in an economy. Hence, institutions indirectly affect wealth accumulation and economic growth. By the turn of the century, Dani Rodrik (2000), in his influential paper, reflected that the research climate changed from "do institutions matter?" to "which institutions matter and how does one acquire them?". In the same paper, Rodrik also listed five key aspects of quality institutions, namely strong adherence to and enforcement of property rights; regulatory institutions; institutions for macroeconomic stabilisation; institutions for social insurance and institutions for conflict management. Rodrik also detailed the various ways in which such institutions can be acquired to facilitate economic growth. Overall, Rodrik concluded that processing and aggregating local knowledge is key to effective economic growth, and participatory political systems excel at that. Currently, there is a wide body of literature covering the effects of institutions on economic growth. For example, after North and Rodrik, Valeriani and Peluso (2011) briefly reviewed the literature for the past twenty years and concluded that institutional quality has a positive significant effect on economic growth, which is in line with the consensus. They tested three indicators for institutional quality using a pooled panel and a within-groups estimation with regional dummies and over 181 countries. Although highly rudimentary, Valeriani and Peluso indicated that it had a degree of robustness due to having similar results under different economic conditions. Ulubasoglu and Doucouliagos (2004) performed a simultaneous systems with error components approach and also found overall positive and significant indirect and direct effects of institutional quality on economic growth. Many other papers that address econometric issues have found positive significant and positive results as well. This paper adds to the literature by looking at institutions and economic growth from a cross-sectionally dependent perspective, taking into account the presence of outliers. We showcase current estimation methods available to address the issues in the econometrics for economic growth, and we propose outlier-robust alternatives to the current methods. This paper is organised into two parts. The first part investigates the effect of government institutions on economic growth. The second part investigates the effect of gender disparity on economic growth. Each part is organised into sections as follows: An introduction and review; theory and methodology; data; results and discussion; and each part wraps up with a summary. 6.3 concludes the entire paper.

6.1.2 Institutions and Economic Growth

We follow the definition of an institution laid down by Lin and Nugent (1995) as "a set of humanly devised behavioural rules that govern and shape the interactions of human beings, in part by helping them to form expectations of what other people will do." (see also, Rodrik, 2000; Valeriani et. al., 2008) In neoclassical theories, the principal determinants of economic growth are population and savings (North, 1989). Additionally, in the literature, Robert Solow (Samuelson and Solow, 1956) lists labour, capital and technological progress as the key determinants in his flagship Solow Model. Mankil, Romer and Weil (1992) augmented the Solow model to include human capital as a determinant of Economic growth. If the above ingredients are sufficient, countries with rich resources and plenty of talented people should be wealthier and enjoy ever-increasing economic growth. But we do not see that happening. To illustrate, consider the economic progression of some first-world and third-world countries post-World War II. In his influential book, Daron Acemoglu (2008; pp. 11-12, 29-30, 872 and more) gave the example of Nigeria and India starting out at similar levels of income per capita, but India experienced little growth until the 1980s. From the 1980s till the 2000s, India experienced relatively rapid economic growth, but did not attain the wealth levels of advanced economies like the United States. On the other hand, Nigeria experienced a GDP contraction such that it was poorer in 2000 than in 1960. In the same book (Acemoglu, 2008), Acemoglu contrasted this with countries like South Korea and Singapore, which constantly experienced rapid growth postwar. Furthermore, Nigeria was geographically more advantageous than both those eastern countries, being richer in natural resources than both South Korea and Singapore. Acemoglu also posited that economists previously attributed this to failure on Nigeria's part to accumulate physical and human capital, as well as failing to provide a politically stable landscape that encourages investment in capital and technology. Why did countries like Nigeria fail to take advantage of their natural resources and education? Acemoglu then stated four fundamental forces of economic growth: luck and the existence of multiple equilibria; geography; institutions; and culture. Acemoglu also pointed out that institutions may be the most important determinant of the four. An alternative explanation to the unfortunate state of Nigeria was put forth by Sachs and Warner (1995), in which they found that resource-rich countries with high export-to-GDP ratios had lower growth rates than resource-scarce countries. In the decade that followed, many paper have studied this hypothesis, often known as the "natural resource curse hypothesis" (see Brunnschweiler, 2008 for a review). However, in the same paper, Brunnschweiler also remarked in her paper that "institutions may however play a critical role in determining the economic performance of resource-rich economies". Rodrik (2000) also stated that in the standard Arrow-Debreu model with complete and contingent markets extending indefinitely into the future, the model assumptions indicate that there exist institutions that establish and protect property rights and enforce contracts, in order to make even "perfect" markets function. According to Douglass North (1989), the neoclassical theories hold only if we live in a closed economy with zero transaction costs (and no governments or institutions). Institutions, therefore, either facilitate or hinder transaction costs. Following Rodrik (2000), what institutions promote economic growth while causing others to stagnate? Going back to North (1989), the costs of transacting in an economy are influenced by its size and complexity, the existence of institutions (government) to specify property rights and enforce contracts, and by its social norms. Similarly, Rodrik (2000) lists five types of market-supporting institutions: property rights; regulatory institutions; institutions for macroeconomic stabilization; institutions for social insurance; and institutions of conflict management. The general consensus in the literature is that the reason why advanced economies are more prosperous than third-world countries is that the institutions in advanced economies promote secure environments that promote investment, where the property rights and the rule of law is adhered to and enforced. Contracts are upheld and enforced, and the uncertainties associated with interacting with unknown economic agents are minimised or eliminated. Rodrik (2000) found that democracy is a meta-institution that builds quality institutions and participatory democracies promote higher-quality growth. To measure the quality of these market-supporting institutions, a common empirical practice is to aggregate relevant measured variables into indices. In the next section, we use six such indicators from the World Bank that proxy institutional quality.

6.1.3 The Econometrics of Institutional Quality

We now look at the econometrics of institutions in the literature. The standard model used is the Barro regression model (see Acemoglu, 2008 for example):

$$\Delta y_{it} = \beta_0 + \beta_1 y_{i,t-1} + \epsilon_{it} \tag{6.1}$$

where y is the log of per capita output, $\Delta y_{it} = y_{it} - y_{i,t-1}$ is economic growth rate and ϵ_{it} is the idiosyncratic error. The panel setting of this equation accounts for cross-country effects. (6.1) can be augmented with individual effects β_{i0} to capture cross-country differences. We also consider heterogeneous effects for all coefficients:

$$\Delta y_{it} = \beta_{i0} + \beta_{i1} y_{i,t-1} + \mathbf{x}'_{it}\boldsymbol{\beta}_i + \epsilon_{it} \tag{6.2}$$

Much of the literature is based on standard estimation methods. The common issues are endogeneity, omitted variable biases, cross-section dependence and the presence of outliers. Most empirical studies on panel data, including those on institutions and economic growth, tent to ignore the dependence of errors across units except for spatial models. However, with an increased availability of data and the widespread application of panel studies to aggregated macroeconomic data, it has been recognised that even after conditioning for individual specific regressors, individual units in general need not be cross-sectionally independent (Chudik and Pesaran, 2013). For more details on cross-section dependence, we refer to Alexander Chudik and Hashem Pesaran's (2013) review paper. Cross-sectional correlations may arise due to omitted common effects, spatial effects, or could arise as a result of interactions within socioeconomic networks. Depending on the extent of cross-section dependence, the consequences can be serious. If the cross-sectional correlation is strong, conventional fixed- or random effects estimators can have inconsistent estimators and misleading inferences. This is exacerbated if the source of the cross-sectional correlations are correlated with the regressors. Furthermore, unit root tests can have substantial size distortions (Chudik and Pesaran, 2013). In many empirical studies, economists often use a simultaneous equation or Seemingly Unrelated Regressions approach (SURE) or a Generalised Method of Moments approach (GMM; Arellano and Bond, 1990). Ulubasoglu and Doucouliagos (2004) attempted to address the cross-sectional correlations by using a simultaneous equation framework. They estimated a Solow model augmented with human capital (Mankiw, Romer and Weil, 1992) and an index for institutional qualiy. Simultaneously, they estimated each input of the Solow model with the institutional quality index as a regressor, reflective of the understanding in the literature that institutions influence the accumulation of knowledge, human and physical capital, and much more. Ulubasoglu and Doucouliagos (2004) found that institutional quality had positive and significant indirect and direct effects on economic growth. In the literature, institutional quality is generally positively correlated with economic growth: naturally, the better the institutional quality, the greater the economic growth. Endogeneity is another issue not only in institutions, but in the overall economic growth literature. The input, technological progress and innovation, is known to be endogenous in the literature. It is often proxied by Research and Development (R&D) spending as an instrumental variable. By definition, technological innovation refers to inventions that increase output, in contrast to progress, which makes improvements on existing inventions to further improve output levels (Acemoglu, 2008). Technological progress also affects the quality of capital and labour inputs, thus having both direct and indirect effects on economic growth in standard models like the augmented Solow Model (see Acemoglu, 2008 for details). Economists often use 2-or 3-stage Least Squares (2SLS; 3SLS), GMM or System-GMM estimators to address this endogeneity. A third issue is Omitted Variable Bias (OMV), which is a ubiquitous problem in empririical studies in general, due typically to the inavailability of data or a missspecified model. Catrinescu et. al. (2008) examined the effects of institutions on remittances and acknowledged the potential OMV remittances bring, that remittances are more likely to contribute to longer-term growth when the remittance-receiving countries' political and economic policies and institutions create incentives for financial and business investment and savings from remittances. Catrinescu and her team tackled the OMV by interacting institutional indices with the remittance variable in their regressions and also found generally positive and significant effects that remittances contribute to economic growth, when institutional quality is high. We will address the issue of OMV in the next section with a factor model framework. The fourth issue we will be looking at is the presence of outliers. Outliers are often ignored in empirical studies or treated with nonrobust outlier detection procedures. Traditional outlier detection methods, such as the Z-score and the Mahalanobis distance, suffer from masking effects of multiple outliers. Leave-one-out regressions are effective in weeding out lone outliers from a sample, but also suffer from the masking effect when other outliers are present. It is possible to try excluding different combinations of subsets to obtain estimates with the smallest residuals, and given the processing power of today's computers, calculating such combinations is not an issue with (relatively) small sample sizes. Outliers can be fundamentally modelled as a mixed distribution (see, among others, Davies and Gather, 1993; Maronna et. al., 2019):

$$F(x) = (1 - \epsilon)G(x) + \epsilon H(x) \tag{6.3}$$

where $F(x)$ is the mixed distribution of the sample, given the true distribution $G(x)$ and the contaminated distribution $H(x)$. The breakdown point is a measure of performance of robust estimators: it is the smallest percentage of outliers in a sample that can cause the estimates to deviate arbitrarily far away from their true values (see Rousseeuw and Leroy, 1987 for details). The maximum breakdown point of any estimator is 50%. Robust estimators with the highest breakdown points include the median-based Least Median of Squares (LMS; Rousseeuw, 1984; Rousseeuw and Leroy, 1987) and the Least Trimmed Squares estimator (LTS; Rousseeuw and Leroy, 1987). The tradeoff is that the LMS has an inherent breakdown point of 50%. The LTS can be tuned with a breakdown point of 50% simply by setting $h = n/2$. However, the LMS suffers from a slow convergence rate of $n^{-1/3}$ (see Rousseeuw and Leroy, 1987), which renders it relatively very inefficient. The LTS suffers from a tradeoff between efficiency and breakdown point. For this reason, the LMS and LTS are often used as diagnostic tools. MM-estimators (see Maronna et. al., 2019) were developed to harness the high breakdown points of estimators like the LMS and LTS while preserving relative efficiency. The standard OLS estimator is known to have a finite-sample breakdown point of $1/n$, or an asymptotic breakdown point of 0%: it takes only one suffciently large outlier to arbitrarily distort the slope estimates (for examples, see Rousseeuw and Leroy, 1987). Likewise, the CCE estimators are computed with least squares. There is further potential to distort the estimates as cross-section averages are used in the slope and variance estimations (similarly, the mean has a zero breakdown point, rendering it very nonrobust). Lee and Sul (2022) attempted to address the issue of outliers by trimming outlying cross-section units with unusually high sample variances. The outlier diagnostic used to detect the outlying individual unit sample variances was the Mahalanois distance, developed by P.C. Mahalanobis (1936). It is a ubiquitous tool in multivariate outlier detection. It computes the standardised distance of each and every data point from the (Euclidean) centre of the k-dimensional data cloud. Formally, the Mahalanobis distance

$$MD_i = (\mathbf{x}_i - \bar{\mathbf{x}})'\mathbf{\Sigma}^{-1}(\mathbf{x}_i - \bar{\mathbf{x}}) \tag{6.4}$$

The statistic is χ^2-distributed, and outlying observations $\mathbf{x}$ that are too far from the centre of the data cloud are flagged as outliers (typically, the critical value for a significance level of 97.5% on the χ^2_N distribution is used). The Mahalanobis distance has been ubiquitous in outlier detection diagnostics in many fields. However, the Mahalanobis distance suffers from a zero breakdown point and the masking effect. In this paper, we will adopt an outlier treatment approach similar to Lee and Sul's (2020;2022). We will estimate a Barro regression model with the CCE estimators trimmed with robust analogues of the Mahalanobis distance. We will look at the details in the next section.

6.1.4 Theory and Methodology

We begin with the Barro regression model (6.2) with individual effects and heterogeneous coefficients. To address the issues of endogeneity and cross-section dependence, we will transform this model into a factor model and estimate it with the Common Correlated Effects (CCE) estimators (Pesaran, 2006). The CCE factor model in Hashem Pesaran's (2006) paper encompasses several models, including the mean group (MG) estimator. We will transform the Barro regression model into a factor model. We start by adding strong cross-section dependence in the error terms (see Chudik and Pesaran, 2013 for more details).

The Model

The general factor model in Pesaran (2006) is given by

$$\begin{aligned} y_{it} &= \boldsymbol{\alpha}_i'\mathbf{d}_t + \boldsymbol{\beta}_i'\mathbf{x}_{it} + \boldsymbol{\gamma}_i'\mathbf{f}_t + e_{it} \\ \mathbf{x}_{it} &= \mathbf{A}_i'\mathbf{d}_t + \boldsymbol{\Gamma}_i'\mathbf{f}_t + \mathbf{v}_{it} \end{aligned} \tag{6.5}$$

where $\mathbf{d}_t$ is a $n \times 1$ vector of observed common effects, including deterministics such as intercepts and seasonal dummies, $\mathbf{x}_{it}$ is a $k \times 1$ vector of observed individual-specific regressors on the ith cross-section at time t and $\mathbf{f}_t$ is a $m \times 1$ vector of unobserved effects. $\mathbf{f}_t$ covers unobserved variables like technological change or innovation, which are known to be correlated with both economic growth and its inputs. The CCE Mean Group (CCEMG) estimator in Pesaran (2006) is

$$\begin{aligned} \hat{\mathbf{b}_i} &= (\mathbf{X}_i'\bar{\mathbf{M}}_w\mathbf{X}_i)^{-1}(\mathbf{X}_i'\bar{\mathbf{M}}_w\mathbf{y}_i) \\ \hat{\boldsymbol{\beta}}_{CCEMG} &= \frac{1}{N}\sum_{i=1}^{N}\hat{\mathbf{b}_i} \end{aligned} \tag{6.6}$$

$\bar{\mathbf{M}}_w$ is the annihilator matrix that proxies the unobserved effects (see Pesaran (2006) for more details). The CCE Pooled (CCEP) estimator is

$$\hat{\mathbf{b}}_P = \left(\sum_{i=1}^{N} \theta_i \mathbf{X}_i' \bar{\mathbf{M}}_w \mathbf{X}_i \right)^{-1} \left(\sum_{i=1}^{N} \theta_i \mathbf{X}_i' \bar{\mathbf{M}}_w \mathbf{y}_i \right) \tag{6.7}$$

(6.5) can include the Barro regression equation. We consider the Barro growth regression (6.2) with individual and time effects

$$\Delta y_{it} = a_i + \delta_t + \beta_{0i} y_{i,t-1} + \beta_{1i} I_{i,t} + \boldsymbol{\beta}_i' \mathbf{x}_{it} + e_{it} \tag{6.8}$$

where y_{it} is the log per capita output, I_{it} is the indicator of institutional quality and $\mathbf{x}_{it}$ is a vector of covariates. To account for strong cross-section dependence, we fit the errors e_{it} with a strong cross-sectional dependence structure

$$e_{it} = \boldsymbol{\gamma}_i' \mathbf{f}_t + \epsilon_{it} \tag{6.9}$$

where $\mathbf{f}_t$ is the vector of unobserved common effects as in (6.5) and idiosyncratic error ϵ_{it}. Assuming that the time effects δ_t in (6.8) are mean-zero stationary processes with absolute summable autocovariances, we can include the time effect δ_t in the vector $\mathbf{f}_t$ and rewrite (6.8) as

$$\Delta y_{it} = a_i + \beta_{1i} I_{i,t} + \boldsymbol{\beta}_i' \mathbf{x}_{it} + \boldsymbol{\gamma}_i' \mathbf{f}_t + \epsilon_{it} \tag{6.10}$$

We can write (6.10) as

$$\Delta y_{it} = a_i + \mathbf{b}_i \mathbf{X}_{it} + \boldsymbol{\gamma}_i' \mathbf{f}_t + \epsilon_{it} \tag{6.11}$$

with

$$\begin{aligned} \mathbf{b}_i &= (\beta_{0i}, \beta_{1i}, \beta_i')' \\ \mathbf{X}_{it} &= (y_{i,t-1}, I_{it}, \mathbf{x}_{it}')' \end{aligned} \tag{6.12}$$

Then $\hat{\mathbf{b}}_i$ is estimated exactly as in (6.6) for CCEMG and (6.7) for the CCEP. However, the asymptotics of the CCE estimators requite that both the cross-section and time dimensions (N and T respectively) are large. Since we are working with a relatively small time dimension (about 20 years' annual data), we aim to get consistent results with inference by considering results from two additonal variance estimators, the Newey-West type estiamtor (Newey and West, 1987; quoted and described in Pesaran, 2006); and the fixed-T variance estimators employed in Westerlund, Petrova and Norkute (2019). Both these modules are available in Stata under the *xtdcce2* package by Jan Ditzen (Ditzen, 2018). However, the CCE estimators are not robust to outliers. As the CCE is a LS-based estimator, it takes only one suffciently large outlier to distort estimates, resulting in a breakdown point of $1/NT$ (or an asymptotic breakdown point of 0%). Furthermore, outliers are almost always present in data (see Rousseeuw and Hubert, 2011; among many others). We discuss the data and the results, and address the issue of outliers in the following sections.

6.1.5 Data

We assess six indicators of institutional quality from the World Governance Indicators (WGI) project (Kaufmann, Kraay and Mastruzzi, 2010): Control of Corruption, Government Effectiveness, Political Stability, Regulatory Quality, Rule of Law, and Voice of Accountability. In the WGI project, these indicators are based on several hundred variables obtained from 31 different data sources, capturing governance perceptions as reported by survey respondents, non governmental organizations, commercial business information providers, and public sector organizations worldwide. The data is then organised into six clusters corresponding to the six broad dimensions of governance listed above. We use data from 1996 to 2019 for a panel of 142 countries in total. The indicators are defined accordingly in the WGI project:

Control of Corruption

Captures perceptions of the extent to which public power is exercised for private gain, including both petty and grand forms of corruption, as well as "capture" of the state by elites and private interests.

Government Effectiveness

Captures perceptions of the quality of public services, the quality of the civil service and the degree of its independence from political pressures, the quality of policy formulation and implementation, and the credibility of the government's commitment to such policies.

Political Stability

Captures perceptions of the likelihood that the government will be destabilized or overthrown by unconstitutional or violent means, including politically motivated violence and terrorism.

Regulatory Quality

Captures perceptions of the ability of the government to formulate and implement sound policies and regulations that permit and promote private sector development.

Rule of Law

Captures perceptions of the extent to which agents have confidence in and abide by the rules of society, and in particular the quality of contract enforcement, property rights, the police, and the courts, as well as the likelihood of crime and violence.

Voice of Accountability

Captures perceptions of the extent to which a country's citizens are able to participate in selecting their government, as well as freedom of expression, freedom of association, and a free media.

Table 6.1: Summary Statistics - Economic Growth and Institutions (World)

	Mean	Standard Deviation	10th Percentile	Median	90th Percentile
GDP growth	0.043	0.092	-0.033	0.040	0.127
Control of Corruption	-0.034	1.023	-1.148	-0.330	1.732
Government Effectiveness	0.031	0.987	-1.115	-0.160	1.616
Political Stability	-0.112	0.951	-1.381	-0.045	1.081
Regulatory Quality	0.067	0.947	-1.082	-0.108	1.512
Rule of Law	-0.043	0.999	-1.183	-0.263	1.634
Voice of Accountability	-0.042	0.969	-1.325	-0.070	1.343

The six indicators from the WGI project were compiled with data from 31 data sources on survey respondents and measured on a standard normal scale, with the majority of observations falling between -2.5 to 2.5. We note that the years 1997 and 1999 do not contain data as the WGI project initially started with biannual data. From 2000 onwards, all data is compiled annually. Table 6.1 presents summary statistics for GDP growth and the six indicators. In the WGI project paper, Kaufmann, Kraay and Mastruzzi indicated that the indicators are correlated, citing potentially strongly complementary dimensions. As quoted by the authors (Kaufmann, Kraay and Mastruzzi, 2010):

> "One might reasonably think for example that better accountability mechanisms lead to less corruption, or that a more effective government can provide a better regulatory environment, or that respect for the rule of law leads to fairer processes for selecting and replacing governments and less abuse of public office for private gain"

By simply computing the sample correlation matrix of the six indicators (and the other regressors) reveals strong correlations between the variables, even after treatment with outliers. To test for cross-section dependence, we first obtain within group residual estimates for each set of regressions for use in CD-tests. These CD-tests for each set of regressions have resulted in CD-statistics that reject the null (of no strong cross-section dependence), indicating the presence of strong cross-section

Table 6.2: Summary Statistics - Economic Growth and Institutions (by Economic Region)

	Mean	Standard Deviation	10th Percentile	Median	90th Percentile
OECD Countries					
GDP growth	0.030	0.045	-0.014	0.027	0.082
Control of Corruption	1.175	0.820	0.071	1.309	2.171
Government Effectiveness	1.213	0.616	0.309	1.334	1.911
Political Stability	0.668	0.735	-0.446	0.877	1.363
Regulatory Quality	1.214	0.484	0.522	1.254	1.800
Rule of Law	1.182	0.673	0.311	1.287	1.902
Voice of Accountability	1.105	0.460	0.642	1.152	1.557
Emerging Markets					
GDP growth	0.047	0.104	-0.037	0.047	0.138
Control of Corruption	-0.359	0.670	-1.166	-0.423	0.628
Government Effectiveness	-0.215	0.654	-1.056	-0.210	0.620
Political Stability	-0.332	0.866	-1.408	-0.270	0.808
Regulatory Quality	-0.181	0.708	-1.138	-0.174	0.683
Rule of Law	-0.364	0.684	-1.200	-0.403	0.560
Voice of Accountability	-0.342	0.806	-1.430	-0.300	0.674
Developing Countries					
GDP growth	0.049	0.093	-0.036	0.048	0.136
Control of Corruption	-0.720	0.466	-1.258	-0.760	-0.039
Government Effectiveness	-0.806	0.504	-1.382	-0.816	-0.344
Political Stability	-0.535	0.837	-1.849	-0.349	0.386
Regulatory Quality	-0.659	0.502	-1.240	-0.688	-0.161
Rule of Law	-0.721	0.534	-1.395	-0.728	-0.174
Voice of Accountability	-0.568	0.640	-1.411	-0.549	0.204

dependence. Hence, Fixed Effects results are expected to be biased and inconsistent. The results for the CD-tests are reported in each table of estimation results (tables 6.3 to 6.6; also in the appendix for the full set of results). Thus the factor model (6.5) and the CCE estimators are suitable for treating the multicollinearity and cross-section dependence. We first perform a standard fixed-effects panel regression. We then compare it with the rest of the regressions in the set. Secondly, we perform four regressions for the standard CCE estimators: CCEMG estimation; CCEP estimation; CCEP with Newey-West variance estimator; and CCEP with Westerlund et. al. (2019) variance estimator. The third set of estimators contain the robustified CCE estimators: Lee and Sul's (2022) TCCEMG estimators; TCCEMG estimators replaced with MCD depths; and TCCEMG estimators replaced with MVE depths. For each dataset, we record the list of outliers in tables 6.1.7. We will discuss the results and examine the outliers in the next section. From the global estimation results, we suspect that the global economic arena is possibly more diverse than just a dichotomous mix of distributions as in (6.3). The literature has shown that results are consistent and reliable only for countries with similar characteristics (see, for example, Acemoglu, 2008; Seguino, 2000). For this reason, we further split our sample into three different socioeconomic regions for comparison: OECD countries; emerging markets; and developing countries. We list the countries in the entire sample in the appendix. We then perform the same set of estimations on each region and report the results in tables 6.4 (for OECD countries), 6.5 (for emerging markets), and 6.6 (for developing countries). Table 6.2 presents summary statistics for each region. The values tell a clearer picture: the mean values for the indices of OECD countries are highly positive (with a score around 1); the emerging markets tend to have lower mean scores (approx. -0.3, on average); and the developing countries have the lowest scores. This clearly supports North's (1989) and Rodrik's (2000) claims that institutions truly matter. The current literature asks, "which institutions matter?" (Rodrik, 2000, for example). We therefore examine these institutions in the results.

6.1.6 Results

Estimation Results

Tables 6.3 to 6.6 list the main estimation results. We report select treated and untreated results. For the untreated results, we report the Fixed Effects estimator (FE, in the tables); the CCEP estimator (CCEP in the table); the CCEP estimator with Newey-West variance estimator (CCEPnw); and the CCEP with Westerlund et. al. (2019) variance estimator (CCEPwpn). The untreated results serve as a comparison to the outlier-trimmed results. For each of the outlier trimming techniques (MD-, MCD- and MVE-trimming), we report the CCEPnw and CCEPnw results (listed as TCCEP-r, MD-r and

MVE-r, where r = nw or wpn). In general, we did not get any significant results with the CCEMG estimator, but that will be outside the scope of our paper. However, we list the full results for all the estimation techniques in the appendix. We find that splitting the countries into the different economic regions yield mostly similar results for Regulatory Quality. We can thus assume that those government indicators apply to all countries at an aggregate level.

Regulatory Quality

On a global level, we find near-unaninmous significant positive results for Regulatory Quality. This is in line with the literature, that economic agents' confidence in institutions' abilities to implement reasonable policies and regulations results in better economic growth. We noticed that we obtained larger effects when the data is controlled for cross-section dependence (CCEPnw and CCEPwpn compared with FE). We also obtain slightly larger effects when the CCEP estimators are trimmed with MCD and MVE estimators.

Political Stability

We find positive and significant results with political stability only with Lee and Sul's TCCEMG estimator. We also find weakly significant results with the MVE-trimmed estimator augmented with the Newey-West variance estimator. The effects captured by these two estimators are otherwise similar (at about +0.07). We note that the MCD-trimmed results are closer to the untreated results in magnitude. This indicates that outliers detected by the MCD trim that the MD or MVE trim did not detect might have heavily influenced the variances and magnitude. We will look at the influence of the outliers in a later section. On the other hand, political stability is a very crucial (positive and significant) factor in OECD countries. In other words, agents in OECD countries place a premium on political stability in order to maximise capital accumulation.

Control of Corruption and Government Effectiveness

We find overall insignificant effects for Control of Corruption and Government Effectiveness, contrary to what the literature would expect. We also find strangely negative results for Government effectiveness overall, as well as negative and positive effects for the OECD countries. As these are surveys of the public opinion, the authors (Kaufmann, Kraay and Mastruzzi, 2010) acknowledged the inherent bias incurred when recording these observations. We also note that robust trimming has not addressed this issue. This inherent bias is probably significant in most observations (essentially more than half of the entire sample), thus exceeding the breakdown point of any robust estimator. Estimating or partialling out such biases could also be a possible future research topic. In a similar context, as Barro and Lee (1994) put it,

> "It would, of course, be nice to have a direct measure of corruption"

An alternative explanation to the negative and signficant results of Government effectiveness in OECD countries (and negative but insignificant in other countries) is that democracy is found to retard growth in some papers (see, for example, Barro, 1996). Barro (1996) attributed this apparent hindrance to growth as the tendency of demcratic regimes to enforce rich-to-poor redistributions of wealth, including land reform policies. It is perhaps the pressures of interest groups present in democratic systems that knowingly or unknowingly hinder economic growth in these cases. On the other hand, an increasing income gap can also retard economic growth in the long run, as will be demonstrated in the next part of this paper. A balance of these issues can be a possible topic for future research.

Rule of Law and Voice of Accountability

We only find weakly significant results for Rule of Law in developing countries. Apart form survey bias, we note that even in corrupt regimes, agents find some way to secure trust (Rose-Ackerman, 2001). While this means that in spite of the presence (or absence) of law enforcement, agents find their own ways to maximise their utilities. Similarly, for voice of accountability, agents find a way to profit whether or not they have a voice.

Also, in line with literature, we find overall significant and positive results for human capital for all countries. We also find significant negative but stationary results for $y_{i,t-1}$, indicating conditional convergence for economic growth (see Acemoglu, 2009; Charpe et. al., 2019).

Table 6.3: Estimation Results - World

	FE	CCEP	CCEPnw	CCEPwpn	TCCEPnw	TCCEPwpn	MCD-nw	MCD-wpn	MVE-nw	MVE-wpn
Control of Corruption	0.0426	-0.500	-0.0206	-0.0206	0.0573	0.0573	-0.0523	-0.0523	0.0312	0.0312
	(1.31)	(-0.90)	(-0.63)	(-0.47)	(1.83)	(1.28)	(-1.22)	(-0.88)	(0.72)	(0.53)
Government Effectiveness	-0.0907**	0.878	-0.0430	-0.0430	0.0259	0.0259	-0.00766	-0.00766	0.00476	0.00476
	(-2.72)	(1.25)	(-1.37)	(-0.98)	(0.76)	(0.56)	(-0.18)	(-0.14)	(0.11)	(0.07)
Political Stability	0.0309	0.115	0.0325	0.0325	0.0701***	0.0701**	0.0250	0.0250	0.0677*	0.0677
	(1.74)	(0.23)	(1.49)	(1.22)	(3.41)	(2.63)	(0.95)	(0.80)	(2.14)	(1.56)
Regulatory Quality	0.147***	-0.158	0.173***	0.173***	0.0964**	0.0964*	0.131**	0.131**	0.184**	0.184***
	(4.76)	(-0.25)	(5.03)	(4.03)	(2.77)	(2.50)	(3.05)	(2.69)	(3.23)	(3.51)
Rule of Law	-0.0510	-0.0277	0.105*	0.105	0.0482	0.0482	0.194**	0.194*	0.181*	0.181
	(-1.38)	(-0.03)	(2.58)	(1.88)	(0.93)	(0.61)	(3.09)	(2.08)	(2.44)	(1.58)
Voice of Accountability	0.00647	2.930	-0.0545	-0.0545	-0.0412	-0.0412	-0.0339	-0.0339	0.0242	0.0242
	(0.23)	(1.04)	(-1.81)	(-1.45)	(-1.34)	(-0.97)	(-1.05)	(-0.82)	(0.49)	(0.34)
Fixed Capital Growth	0.0237***	-0.0970	-0.0255	-0.0255	-0.0220*	-0.0220	-0.0346*	-0.0346	-0.0347*	-0.0347
	(3.33)	(-1.27)	(-1.89)	(-1.61)	(-2.07)	(-1.42)	(-2.35)	(-1.71)	(-2.35)	(-1.82)
Human Capital	0.316***	0.108	1.216***	1.216***	0.993***	0.993***	1.084***	1.084***	1.265***	1.265**
	(9.12)	(0.03)	(5.20)	(3.74)	(5.82)	(3.93)	(4.48)	(3.93)	(4.30)	(3.10)
$y_{i,t-1}$	-0.0854***	-0.219	-0.497***	-0.497***	-0.396***	-0.396***	-0.472***	-0.472***	-0.445***	-0.445***
	(-13.51)	(-0.52)	(-15.95)	(-11.85)	(-15.41)	(-10.24)	(-14.00)	(-11.47)	(-13.30)	(-9.81)
Constant	1.109***	-1.729	6.25e-14	6.25e-14	7.51e-15	7.51e-15	-2.04e-14	-2.04e-14	-7.27e-14	-7.27e-14
	(13.03)	(-0.26)	(0.00)	(0.00)	(0.00)	(0.00)	(-0.00)	(-0.00)	(-0.00)	(-0.00)
Observations	2976	2919	2919	2919	2520	2520	1617	1617	2037	2037

t statistics in parentheses

* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

CD = 79.985, p = 0.000

Table 6.4: Estimation Results - OECD

	FE	CCEP	CCEPnw	CCEPwpn	TCCEPnw	TCCEPwpn	MCD-nw	MCD-wpn	MVE-nw	MVE-wpn
Control of Corruption	0.150***	2.524	0.0351	0.0351	0.0833*	0.0833	-0.00428	-0.00428	0.00245	0.00245
	(3.48)	(0.94)	(0.87)	(0.66)	(2.32)	(1.69)	(-0.12)	(-0.09)	(0.06)	(0.05)
Government Effectiveness	-0.0619	-0.223	0.0214	0.0214	-0.0333	-0.0333	-0.164**	-0.164*	-0.150**	-0.150***
	(-1.19)	(-0.19)	(0.48)	(0.38)	(-0.71)	(-0.63)	(-2.68)	(-2.15)	(-2.64)	(-3.67)
Political Stability	0.0684**	-0.840	0.0261	0.0261	0.0409	0.0409	0.0751**	0.0751**	0.0957***	0.0957***
	(2.63)	(-1.67)	(1.23)	(1.20)	(1.96)	(1.90)	(3.29)	(2.60)	(3.56)	(3.79)
Regulatory Quality	0.0545	0.582	0.196***	0.196***	0.201***	0.201***	0.237***	0.237**	0.127*	0.127*
	(1.07)	(1.25)	(4.33)	(4.21)	(3.95)	(3.43)	(3.87)	(3.22)	(2.27)	(2.33)
Rule of Law	-0.0364	-0.664	-0.0546	-0.0546	-0.0946	-0.0946	-0.0206	-0.0206	0.0319	0.0319
	(-0.62)	(-0.32)	(-0.87)	(-0.70)	(-1.40)	(-1.19)	(-0.30)	(-0.21)	(0.42)	(0.37)
Voice of Accountability	0.0138	-0.747	0.0803	0.0803	-0.0537	-0.0537	0.148	0.148	0.0627	0.0627
	(0.23)	(-0.71)	(1.18)	(0.99)	(-0.59)	(-0.42)	(1.56)	(1.15)	(0.76)	(0.89)
Fixed Capital Growth	0.137***	0.127	0.0214	0.0214	0.0580**	0.0580*	0.0500*	0.0500	0.0624**	0.0624
	(8.84)	(1.31)	(0.87)	(0.58)	(2.83)	(2.32)	(2.08)	(1.62)	(2.75)	(1.74)
Human Capital	0.638***	2.687	1.762***	1.762***	2.077***	2.077***	2.387***	2.387***	2.834***	2.834***
	(6.76)	(0.59)	(5.92)	(6.24)	(4.57)	(6.41)	(6.97)	(5.58)	(7.24)	(6.59)
$y_{i,t-1}$	-0.122***	-1.498***	-0.547***	-0.547***	-0.561***	-0.561***	-0.587***	-0.587***	-0.567***	-0.567***
	(-7.95)	(-5.60)	(-15.12)	(-16.89)	(-8.87)	(-12.53)	(-15.03)	(-11.98)	(-11.41)	(-21.83)
Constant	1.554***	-0.211	-3.21e-14	-3.21e-14	-1.66e-13	-1.66e-13	-2.50e-13	-2.50e-13	-9.65e-16	-9.65e-16
	(7.37)	(-0.03)	(-0.00)	(-0.00)	(-0.00)	(-0.00)	(-0.00)	(-0.00)	(-0.00)	(-0.00)
Observations	798	798	798	798	609	609	504	504	504	504

t statistics in parentheses

* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

CD = 32.729, p = 0.000

Table 6.5: Estimation Results - Emerging Markets

	FE	CCEP	CCEPnw	CCEPwpn	TCCEPnw	TCCEPwpn	MCD-nw	MCD-wpn	MVE-nw	MVE-wpn
Control of Corruption	0.0838	0.147	-0.0360	-0.0360	0.00144	0.00144	0.151**	0.151*	0.0432	0.0432
	(1.93)	(0.39)	(-0.91)	(-0.72)	(0.03)	(0.03)	(3.05)	(2.41)	(0.77)	(0.57)
Government Effectiveness	-0.122**	-1.510	-0.0108	-0.0108	0.00463	0.00463	-0.0712	-0.0712	-0.0673	-0.0673
	(-2.88)	(-1.42)	(-0.30)	(-0.22)	(0.13)	(0.09)	(-1.76)	(-1.44)	(-1.59)	(-1.24)
Political Stability	0.0222	-0.0175	0.0815***	0.0815**	0.0612*	0.0612	0.106**	0.106**	0.0839*	0.0839*
	(0.87)	(-0.03)	(3.30)	(2.78)	(2.12)	(1.78)	(3.00)	(2.82)	(2.55)	(2.34)
Regulatory Quality	0.179***	0.619	0.114**	0.114*	0.135**	0.135*	-0.0280	-0.0280	-0.0496	-0.0496
	(4.57)	(1.39)	(2.96)	(2.25)	(2.87)	(2.53)	(-0.49)	(-0.39)	(-0.81)	(-0.63)
Rule of Law	-0.0519	-0.0184	0.0844	0.0844	0.0577	0.0577	0.0892	0.0892	0.198**	0.198*
	(-1.07)	(-0.03)	(1.73)	(1.22)	(0.97)	(0.70)	(1.38)	(1.05)	(2.76)	(2.10)
Voice of Accountability	0.0197	-0.231	-0.0348	-0.0348	-0.0760	-0.0760	-0.0691	-0.0691	-0.0836	-0.0836
	(0.54)	(-0.25)	(-0.94)	(-0.70)	(-1.48)	(-1.10)	(-1.50)	(-1.41)	(-1.57)	(-1.23)
Fixed Capital Growth	0.0547***	-0.0431	0.0205	0.0205	0.00651	0.00651	0.0197	0.0197	-0.0123	-0.0123
	(5.82)	(-0.48)	(1.28)	(0.90)	(0.45)	(0.36)	(1.00)	(0.90)	(-0.58)	(-0.52)
Human Capital	0.365***	-0.418	1.841***	1.841***	1.350***	1.350*	1.700***	1.700**	1.748***	1.748**
	(6.40)	(-0.24)	(7.07)	(4.40)	(3.48)	(2.45)	(4.15)	(2.63)	(4.30)	(2.86)
$y_{i,t-1}$	-0.0791***	-0.931***	-0.517***	-0.517***	-0.554***	-0.554***	-0.484***	-0.484***	-0.535***	-0.535***
	(-9.58)	(-5.38)	(-16.34)	(-11.47)	(-15.42)	(-11.29)	(-11.26)	(-8.48)	(-12.39)	(-12.12)
Constant	1.049***	-0.434	-3.24e-14	-3.24e-14	5.31e-14	5.31e-14	3.32e-14	3.32e-14	-2.02e-14	-2.02e-14
	(9.18)	(-0.08)	(-0.00)	(-0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(-0.00)	(-0.00)
Observations	1678	1659	1659	1659	1407	1407	987	987	1071	1071

t statistics in parentheses

* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

CD = 56.576, p = 0.000

Table 6.6: Estimation Results - Developing Countries

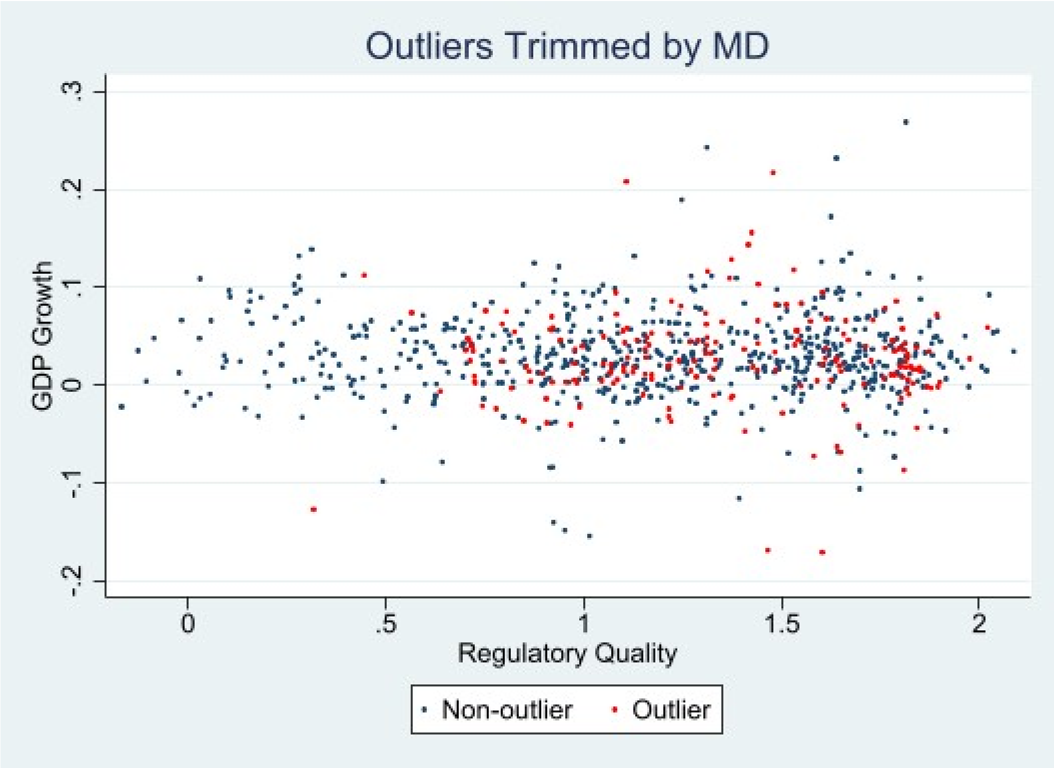
	FE	CCEP	CCEPnw	CCEPwpn	TCCEPnw	TCCEPwpn	MCD-nw	MCD-wpn	MVE-nw	MVE-wpn
Control of Corruption	-0.0208	-0.354	-0.0965	-0.0965	-0.105	-0.105	-0.329***	-0.329*	-0.104	-0.104
	(-0.30)	(-1.08)	(-1.37)	(-0.98)	(-1.56)	(-1.48)	(-3.45)	(-2.30)	(-1.27)	(-1.11)
Government Effectiveness	-0.0563	-0.236	-0.115	-0.115	-0.199	-0.199	-0.208*	-0.208	-0.0991	-0.0991
	(-0.66)	(-0.64)	(-1.20)	(-1.14)	(-1.96)	(-1.86)	(-2.03)	(-1.84)	(-1.01)	(-0.84)
Political Stability	0.0871*	0.164	0.103	0.103	0.0395	0.0395	0.196**	0.196*	0.149**	0.149*
	(2.51)	(0.50)	(1.76)	(1.44)	(0.57)	(0.48)	(2.75)	(2.21)	(2.63)	(2.16)
Regulatory Quality	-0.00212	-0.0366	0.154	0.154	0.296**	0.296**	0.143	0.143	0.465***	0.465***
	(-0.03)	(-0.07)	(1.88)	(1.50)	(2.82)	(2.96)	(1.17)	(1.40)	(4.58)	(4.43)
Rule of Law	-0.104	0.120	0.0174	0.0174	0.0798	0.0798	0.0801	0.0801	-0.194	-0.194
	(-1.26)	(0.23)	(0.16)	(0.17)	(0.67)	(0.61)	(0.54)	(0.70)	(-1.54)	(-1.16)
Voice of Accountability	-0.0670	0.629	0.0862	0.0862	0.0898	0.0898	0.140	0.140	0.196*	0.196*
	(-1.13)	(1.32)	(1.45)	(1.07)	(1.45)	(1.17)	(1.42)	(1.15)	(2.28)	(2.16)
Fixed Capital Growth	-0.0691***	0.0264	-0.0217	-0.0217	-0.0286	-0.0286	-0.0563	-0.0563	-0.0524**	-0.0524
	(-4.84)	(0.28)	(-1.31)	(-0.91)	(-1.05)	(-0.76)	(-1.86)	(-1.44)	(-3.08)	(-1.92)
Human Capital	0.416***	0.785	0.978***	0.978*	0.953**	0.953	0.903**	0.903*	1.089***	1.089**
	(6.59)	(0.41)	(3.80)	(2.48)	(2.64)	(1.86)	(2.94)	(2.06)	(4.18)	(3.19)
$y_{i,t-1}$	-0.139***	-0.861**	-0.650***	-0.650***	-0.602***	-0.602***	-0.646***	-0.646***	-0.670***	-0.670***
	(-7.96)	(-3.30)	(-16.00)	(-13.18)	(-14.05)	(-14.89)	(-19.00)	(-19.49)	(-16.99)	(-12.26)
Constant	1.776***	4.665	5.50e-14	5.50e-14	6.73e-14	6.73e-14	6.20e-14	6.20e-14	6.83e-16	6.83e-16
	(8.17)	(0.40)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)
Observations	626	588	588	588	483	483	399	399	441	441

t statistics in parentheses

* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

CD = 2.563, p = 0.010

Figure 6.1: Outliers Flagged by MD, OECD countries



Outliers

Table 6.1.7 lists the outliers flagged by the various estimators by economic region. We record the outliers flagged in each of the set of estimations (separated by region). We note that the MCD estimator has flagged the most outliers out of the three. Although the MVE estimator flagged less outlying countries than the MCD, both robust estimators still detect more outliers than the classical MD estimator. We also note that all outliers flagged by MD were also detected by MCD and MVE. In addition, the MVE and MCD reveal outliers not found by the non-robust MD. This demonstrates the masking effect of outliers, that other outliers can be masked or made to look like non-outliers by other outlying observations. Due to the diverse (highly heterogenously distributed) global economy, we will focus on the outliers by region. We still report the outliers for the global regressions in the appendix. Figures 6.1 to 6.3 display the respective outliers marked in red for all OECD observations. It is easy to visually inspect data for outliers if there is only one regressor. In practice, this is not applicable, as illustrated by the scatterplots. Even for the nonrobust MD, the outliers are embedded in the data cloud. The MCD and MVE reveal outliers that were undetected by MD. These outliers were masked by the outliers that were detectable by all three estimators.

Figure 6.2: Outliers Flagged by MCD, OECD countries

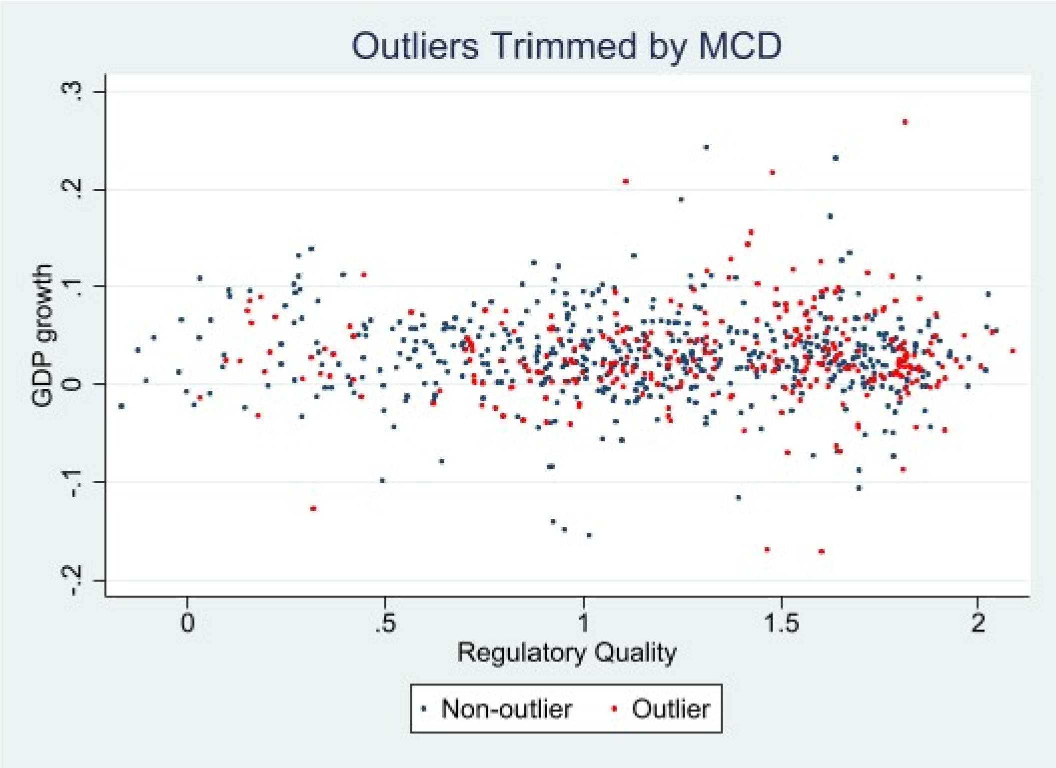
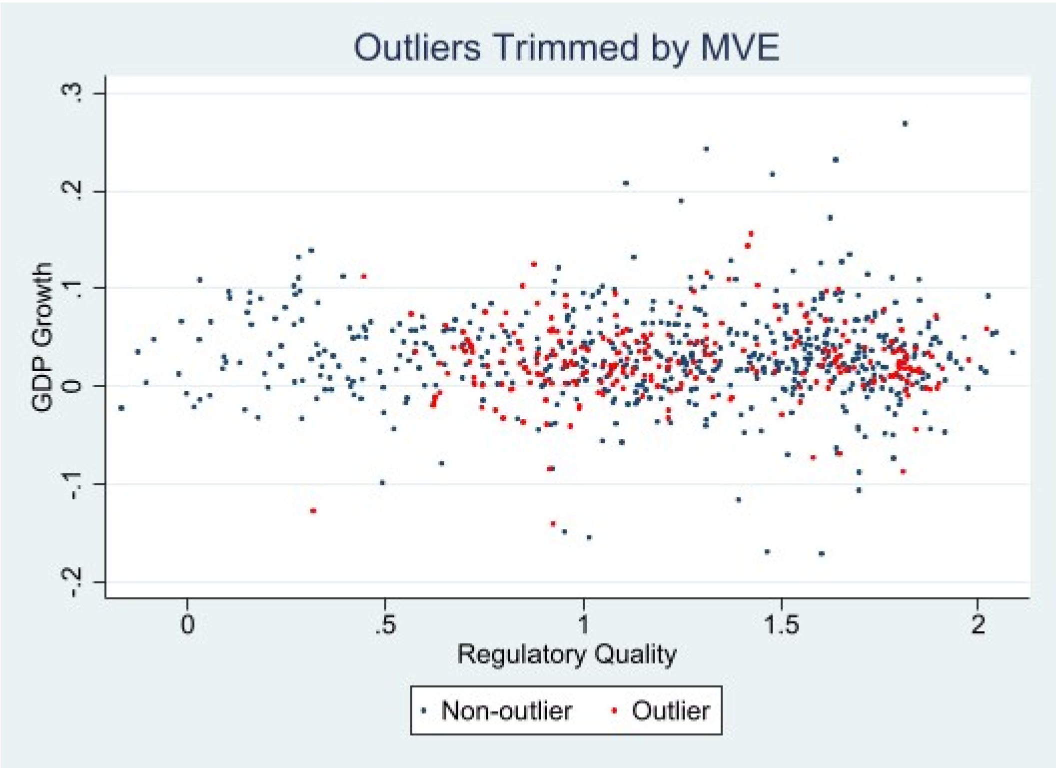


Figure 6.3: Outliers Flagged by MVE, OECD countries



6.1.7 Summary

We have examined six governance indicators. We accounted for the issues of endogeneous and unobseved variables such as technological progress, and of collinearity and cross-section dependence by transforming the Barro regression into a panel data model. We addressed the suceptibility of the CCE estimators with outlier-robust trimming and with alternative variance estimators. We have demonstrated that outliers can be masked by other outliers. We find that Regulatory Quality is a common denominator for economic agents to place their confidence in investing in an economy. We also find that the global economy is very heterogeneous, and further techniques are needed to further distinguish between the distributions. The common practice in the literature is to divide the global economy into several regions, however this is also slightly naive as the global economy is becoming increasingly interconnected. This interconnectedness also further increases the complexity of sharing and use of techonology, which further complicates its endogeneity. The task of dividing the countries into tighter regions is a potential area of research. Even within sub-divided regions, we can still expect outliers. In the next part, we will look at women's studies and economic growth.

Table 6.7: Outliers (Institutions) - By Economic Region

Country	GDP Growth	MD	MCD	MVE
OECD Countries				
Switzerland	0.036	0	1	0
Chile	0.144	1	1	1
Germany	0.049	0	1	1
Finland	0.034	1	1	1
France	0.041	1	1	1
United Kingdom	0.067	1	1	1
Ireland	0.015	0	1	1
Iceland	-0.084	0	0	1
Israel	0.053	1	1	1
Italy	0.012	1	1	1
Republic of Korea	0.070	1	1	1
Mexico	0.086	0	1	1
Norway	0.082	1	1	1
Portugal	0.041	0	1	1

Sweden	0.052	1	1	1
Emerging Markets				
Albania	0.113	0	1	0
Argentina	0.137	1	1	1
Azerbaijan	0.184	0	1	1
Bulgaria	0.037	0	1	1
Belarus	0.100	0	1	0
Belize	0.118	1	1	1
Bolivia	0.106	0	1	1
Brunei Darussalam	0.091	0	1	0
Côte d'Ivoire	0.132	0	0	1
D.R. of the Congo	0.114	0	1	1
Costa Rica	0.019	0	1	1
Gambia	0.041	1	1	1
Guatemala	0.053	0	1	1
Jamaica	0.024	0	1	1
Jordan	0.058	0	1	0
Kazakhstan	0.174	0	1	1
Kenya	0.100	0	1	1
Kuwait	0.122	1	1	1
Republic of Moldova	0.126	1	1	1
Maldives	0.050	0	1	0
Myanmar	0.181	0	1	1
Mauritius	0.025	1	1	1
Paraguay	0.145	1	1	1
Saudi Arabia	0.204	0	0	1
El Salvador	0.151	0	1	1
Syrian Arab Republic	0.088	1	1	1
Tajikistan	0.176	1	1	1
Trinidad and Tobago	0.144	0	1	1
Tunisia	0.060	1	1	1

Ukraine	0.033	0	1	1
Uruguay	0.138	0	1	0
Viet Nam	0.128	1	1	1
South Africa	0.041	0	1	1
Zimbabwe	0.091	1	1	1
Developing Countries				
Benin	0.032	0	1	0
Guinea	0.117	1	1	0
Liberia	0.029	0	1	1
Lesotho	-0.102	1	1	1
Mozambique	0.022	1	1	1
Mauritania	0.193	0	1	1
Niger	0.136	0	1	1
Nepal	0.105	0	0	1
Rwanda	0.094	1	1	1
Sudan	0.143	1	1	0
Senegal	0.049	0	1	0
Sierra Leone	0.043	0	0	1

1 = flagged as outlier by above diagnostic; 0 otherwise

6.2 Part II - Women's Studies and Economic Growth

6.2.1 Introduction and Review

There is a vast literature on women's studies and economic growth. Some focus on female education attainment, while others focus on female labour participation. Others even look at women's rights. The literature on gender linked to economic growth has yielded mixed results in the literature. Some report positive results, whereas others have reported negative results, including prominent figures in economic growth like Robert Barro (Barro and Lee, 1994; Barro, 1996). Barro (1994) attributed the negative effect of female education on economic growth to the gender gap in educational attainment as a measure of "backwardness". Barro referred to a backward country as one that is early in its stages of economic growth. Neoclassical models have shown that economies in the early stages of modern economic growth tend to experience larger growth rates in contrast to more advanced economies, as growth rate is diminishing over time, holding the long-run level of wealth constant. This is known as the conditional convergence effect (Barro, 1996; see also Charpe et. al., 2019; Acemoglu, 2008 for more examples). Thus, countries with larger gender gaps tend to be the economies that were also very early on their journey to economic growth, hence the apparent controversy. In our paper, we will look at Stephanie Seguino's study of Gender Inequality and Economic Growth (Seguino, 2000) from a cross-section dependent (and endogeneous) perspective. In her paper, Seguino examined the effects of gender wage gaps in a set of semi-industrialized and export-oriented economies in which women provide the bulk of labor in the export sector. Similar to Barro, her findings showed that for these economies, a larger gender wage gap was positively correlated with GDP growth. Similar to Barro, Steph attributed the controversial results to the exploitation of women in the export industry at the cost of increased profitability. Seguino posited that it was not gender gaps that directly promoted economic growth, but that the pay gap in the export industry was positively correlated with increased profitability (and consequently wealth accumulation) at the cost of increased income inequality. She also explained that in other literature, income inequality leads to political conflict or instability that hampers economic growth and investment. In our paper, we note that political stability is significant and positive, but the effect is most pronounced in OECD countries. It is highly likely that democracy is (by its own nature) reliant upon political stability in order for its institutions to be productive. On the other hand, Seguino stated that the income inequality shouldered by women in the export sector actually stimulated exports and raised profit expectations (as well as profits). Unfortunately, the women were less inclined to protest this inequality sufficiently to hinder growth (and thus force institutional change). Seguino concluded that due to the positive growth, patriarchal systems tended to reinforce social norms that favour men, reducing political resistance and the costliness of gender inequality. Our purpose in this paper is to examine gender pay differences in the light of cross-section dependence. We show that controlling for cross-section dependence and unobserved factors results in a negative and significant aggregate gender pay gap with respect to economic growth. The next section outlines the methodology. Section (6.2.3) outlines the data. Section (6.2.4) discusses the results and section (6.2.5) summarises.

6.2.2 Methodology

We adopt the structural model used in Seguino (2000):

$$\Delta y_{it} = \phi + \sum_{i=1}^{N} \lambda_i + \alpha_1 WGAP_{it} + \alpha_2 \Delta k_{it} + \alpha_3 HKF_{it} + \alpha_4 HKM_{it} + \epsilon_{it} \tag{6.13}$$

where k the log of capital stock, HKF and HKM are female and male human capital respectively, ϕ is technological change and ϵ is the idiosyncratic term, assumed normal. We account for heterogeneity in the model and the change of technological progress with time:

$$\Delta y_{it} = \phi_t + \sum_{i=1}^{N} \lambda_i + \alpha_{1i} WGAP_{it} + \alpha_{2i} \Delta k_{it} + \alpha_{3i} HKF_{it} + \alpha_{4i} HKM_{it} + \epsilon_{it} \tag{6.14}$$

As technological progress is unobserved and known to be endogenous, we assume a factor model structure as in (6.5) with

$$\begin{aligned} \Delta y_{it} &= \mathbf{a}_i' \mathbf{d}_t + \mathbf{b}_i' \mathbf{X}_i t + \boldsymbol{\gamma}_i' \mathbf{f}_t + \epsilon_{it} \\ \mathbf{X}_{it} &= \mathbf{A}_i' \mathbf{d}_t + \boldsymbol{\Gamma}_i' \mathbf{f}_t + \mathbf{v}_i t \end{aligned} \tag{6.15}$$

with

$$\mathbf{b}_i = (\alpha_{1i}, \alpha_{1i}, \alpha_{1i}, \alpha_{1i})' \tag{6.16}$$

$$\mathbf{X}_{it} = (WGAP_{it}, \Delta k_{it}, HKF_{it}, HKM_{it})' \tag{6.17}$$

$$\mathbf{a}_i = (\lambda_1, \lambda_2, \ldots, \lambda_N)' \tag{6.18}$$

$$\mathbf{f}_t = (\phi_t, \boldsymbol{\tau}_t')' \tag{6.19}$$

where $\mathbf{d}_t$ is a vector of ones, and $\boldsymbol{\tau}_t$ is a vector of unobserved effects. As before, this model accounts for multicollinearity, cross-section dependence and endogeneity. We also perform the same set of estimations: FE; CCEMG; CCEP, CCEPnw; CCEPwpn; and the trimmed estimators TCCEMG; TCCEP; TCCEPnw; and TCCEPwpn. For the TCCE estimators, we also adopt the classical MD, the robust MCD and MVE estimators for trimming. We present the data and results in the next two sections.

6.2.3 Data

Data on HKF and HKM are obtained from the Global Development Lab (Smits and Permanyer, 2019) for observations from year 2000 to 2021. This database is an ongoing project, with data available since the year 2000 for human capital and related indices. Data on GDP growth was transformed from real GDP output values from the Penn World Tables (Feenstra, Inklaar and Timmer, 2015) for dates 1960 to 2019. This is also an ongoing, with data available until 2019. Data for the gender paygap was computed from income indices for males and females from the GDL from 2000 to 2021. Data from both sources were available for over 200 countries and regions, after trimming countries with more than 3 missing values between years 2000 and 2019, we obtain a panel of 142 countries with data covering years 2000 to 2019. For the gender paygap variable, we compute it as the ratio of male income index to female income index

$$WGAP_{it} = \frac{IncM_{it}}{IncF_{it}} \tag{6.20}$$

where $IncM$ and $IncF$ are the income indices for men and women respectively. This ratio has the convenience of being nonnegative, having a value of 1 for gender equality, a gender pay advantage for women if $WGAP < 1$ and a disadvantage if $WGAP > 1$. Tables 6.8 display summary statistics for GDP growth, human capital for both genders and the associated gender pay gap. Worlwide and in all regions, the Gender Pay Gap, $WGAP_{it} > 1$, confirming the general disparity in pay between men and women. We leave full summary statistics by country in the appendix. We note that the Gender Pay Gap is above 1 for each country. We will discuss the results in the next section.

Table 6.8: Summary Statistics - Women's Studies (World)

	Mean	SD	Min	Median	Max
GDP Growth	0.043	0.083	-0.923	0.040	0.643
Female Human Capital	0.661	0.175	0.220	0.678	0.954
Male Human Capital	0.703	0.153	0.304	0.719	0.975
Gender Pay Gap	1.180	0.229	0.871	1.121	5.581
Fixed Capital Growth	0.048	0.243	-6.775	0.051	2.877

Table 6.9: Summary Statistics - Women's Studies (by Region)

	Mean	SD	Min	Median	Max
OECD Countries					
GDP Growth	0.028	0.042	-0.171	0.025	0.269
Female Human Capital	0.856	0.065	0.603	0.875	0.954
Male Human Capital	0.879	0.061	0.671	0.899	0.975
Gender Pay Gap	1.097	0.041	1.028	1.090	1.311
Fixed Capital Growth	0.032	0.101	-0.510	0.033	0.550
Emerging Markets					
GDP Growth	0.047	0.092	-0.923	0.046	0.580
Female Human Capital	0.660	0.117	0.255	0.676	0.882
Male Human Capital	0.703	0.091	0.445	0.717	0.899
Gender Pay Gap	1.219	0.284	0.993	1.142	5.581
Fixed Capital Growth	0.049	0.277	-6.775	0.057	2.877
Developing Countries					
GDP Growth	0.050	0.087	-0.545	0.050	0.643
Female Human Capital	0.451	0.116	0.220	0.442	0.890
Male Human Capital	0.514	0.106	0.304	0.509	0.902
Gender Pay Gap	1.170	0.112	0.871	1.137	1.713
Fixed Capital Growth	0.063	0.242	-1.513	0.062	2.057

6.2.4 Results

Estimation Results

Tables 6.10 to 6.13 report select results for each region. As before, we report FE, CCEP, CCEPnw, CCEPwpn, and the robust variants TCCEPnw, TCCEPwpn, TCCEPnw and TCCEPwpn with the MCD estimator, and the TCCEP estimators with the MVE estimator. We see that globally, the coefficient for Gender Pay Gap, $WGAP$ is negative and significant, indicating that the greater the disparity, the lower the economic growth rate. Breaking up the results by region, we find insignificant results with MVE and TCCEP (MD), whereas the MCD-TCCEP reports a negative and significant result. We note the the MCD flags more outliers than the MD and MVE; this negative and significant result is reflective of the countries selected by the MCD estimator. We can conclude that those countries have the most similar characteristics. We also find that for emerging markets, lowering the gender pay gap is crucial to increased economic growth, whereas we find positive but insignificant results for developing countries. We also note that these countries are the least politically stable in general, and thus they would be reflective of Barro's (1994) idea of being fledgling economies, with large gender disparities not only in wages, but also in education and human capital between the genders. Perhaps, these countries are improving gradually, and as their institutions improve (see Sako, 2004 for example), they will eventually attain the levels of wealth enjoyed by their more advanced neighbours. The authors of the GDl also acknowledge that aggregation loses information at the micro level (Smits and Permanyer, 2019). Seguino's paper focused on only the export sector; perhaps panel regressions by sector will reveal more insights. These would be potential research areas.

Table 6.10: Estimation Results (Women's Studies) - World

	FE	CCEP	CCEPnw	CCEPwpn	TCCEPnw	TCCEPwpn	MCD-nw	MCD-wpn	MVE-nw	MVE-wpn
Gender Pay Gap	-0.0741***	-0.493	-0.192**	-0.192***	-0.203**	-0.203***	-0.350***	-0.350***	-0.292***	-0.292***
	(-5.55)	(-1.55)	(-3.00)	(-4.70)	(-2.89)	(-4.90)	(-4.77)	(-7.25)	(-3.71)	(-6.95)
Female Human Capital	-0.262	-0.0000328	-0.692	-0.692*	-0.583	-0.583	-0.920	-0.920*	-0.587	-0.587
	(-1.28)	(-0.00)	(-1.78)	(-2.04)	(-1.44)	(-1.56)	(-1.92)	(-2.04)	(-1.33)	(-1.46)
Male Human Capital	0.138	0.929	1.839***	1.839***	2.045***	2.045***	2.146***	2.146***	1.857***	1.857***
	(0.56)	(0.82)	(4.55)	(5.40)	(4.58)	(5.51)	(4.64)	(6.83)	(4.20)	(5.95)
Fixed Capital Growth	0.129***	0.141***	0.0757***	0.0757***	0.0689***	0.0689***	0.0632**	0.0632*	0.0665***	0.0665**
	(20.63)	(10.39)	(5.86)	(4.33)	(4.83)	(3.54)	(3.30)	(2.42)	(3.66)	(2.69)
Constant	0.200***	0.913	0.117	0.117	0.126	0.126	0.125	0.125	0.114	0.114
	(4.26)	(1.82)	(0.32)	(0.32)	(0.36)	(0.36)	(0.27)	(0.27)	(0.29)	(0.29)
Observations	2513	2513	2513	2513	2440	2440	1780	1780	1900	1900

t statistics in parentheses

* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

CD = 54.659, p = 0.000,

Table 6.11: Estimation Results (Women's Studies) - OECD

	FE	CCEP	CCEPnw	CCEPwpn	TCCEPnw	TCCEPwpn	MCD-nw	MCD-wpn	MVE-nw	MVE-wpn
Gender Pay Gap	0.187	0.0272	0.0824	0.0824	0.0614	0.0614	0.831***	0.831***	-0.0127	-0.0127
	(1.71)	(0.06)	(0.72)	(0.83)	(0.50)	(0.49)	(4.82)	(5.07)	(-0.10)	(-0.10)
Female Human Capital	0.799**	2.372*	1.066**	1.066**	0.749*	0.749	3.303***	3.303***	0.982*	0.982*
	(2.86)	(2.08)	(2.78)	(3.02)	(2.02)	(1.82)	(8.77)	(7.39)	(2.46)	(2.26)
Male Human Capital	-0.925***	-1.299	-0.374	-0.374	-0.380	-0.380	-2.561***	-2.561***	-0.797	-0.797*
	(-3.34)	(-0.84)	(-1.00)	(-1.14)	(-1.00)	(-0.98)	(-6.73)	(-6.45)	(-1.96)	(-2.00)
Fixed Capital Growth	0.272***	0.219***	0.208***	0.208***	0.215***	0.215***	0.218***	0.218***	0.216***	0.216***
	(22.70)	(8.76)	(12.05)	(15.43)	(12.08)	(15.41)	(10.98)	(18.52)	(10.13)	(13.21)
Constant	-0.0569	0.340	-8.73e-14	-8.73e-14	4.32e-14	4.32e-14	-8.36e-14	-8.36e-14	-1.69e-14	-1.69e-14
	(-0.35)	(0.19)	(-0.00)	(-0.00)	(0.00)	(0.00)	(-0.00)	(-0.00)	(-0.00)	(-0.00)
Observations	660	660	660	660	640	640	420	420	560	560

t statistics in parentheses

* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

CD = 18.427, p = 0.000

Table 6.12: Estimation Results (Women's Studies) - Emerging Markets

	FE	CCEP	CCEPnw	CCEPwpn	TCCEPnw	TCCEPwpn	MCD-nw	MCD-wpn	MVE-nw	MVE-wpn
Gender Pay Gap	-0.0771***	-0.245	-0.249**	-0.249***	-0.291***	-0.291***	-0.208**	-0.208***	-0.266***	-0.266***
	(-5.33)	(-1.01)	(-2.98)	(-4.60)	(-3.68)	(-4.95)	(-2.85)	(-4.56)	(-3.82)	(-6.39)
Female Human Capital	-0.331	1.182	-0.401	-0.401	-0.704	-0.704	-0.0178	-0.0178	-0.348	-0.348
	(-1.29)	(1.10)	(-0.92)	(-1.11)	(-1.52)	(-1.56)	(-0.03)	(-0.03)	(-0.67)	(-0.64)
Male Human Capital	0.157	0.313	1.547***	1.547***	1.498***	1.498***	0.882*	0.882*	1.128**	1.128**
	(0.51)	(0.31)	(3.61)	(4.37)	(3.62)	(4.45)	(2.00)	(2.36)	(2.59)	(3.26)
Fixed Capital Growth	0.151***	0.143***	0.0993***	0.0993***	0.0953***	0.0953***	0.0942***	0.0942***	0.0947***	0.0947***
	(19.77)	(7.94)	(8.48)	(6.53)	(6.80)	(5.04)	(5.95)	(4.47)	(6.15)	(4.70)
Constant	0.242***	0.667	0.131	0.131	0.155	0.155	0.119	0.119	0.127	0.127
	(3.96)	(0.93)	(0.23)	(0.23)	(0.30)	(0.30)	(0.18)	(0.18)	(0.22)	(0.22)
Observations	1453	1453	1453	1453	1360	1360	1040	1040	1100	1100

t statistics in parentheses

* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

CD = 39.743, p = 0.000

Table 6.13: Estimation Results (Women's Studies) - Developing Countries

	FE	CCEP	CCEPnw	CCEPwpn	TCCEPnw	TCCEPwpn	MCD-nw	MCD-wpn	MVE-nw	MVE-wpn
Gender Pay Gap	-0.0935	-0.610	0.0786	0.0786	0.0451	0.0451	0.0538	0.0538	0.226	0.226
	(-0.86)	(-0.74)	(0.62)	(0.53)	(0.37)	(0.32)	(0.43)	(0.37)	(1.38)	(1.05)
Female Human Capital	-0.389	-1.325	1.005	1.005	0.582	0.582	0.898	0.898	1.598	1.598
	(-0.59)	(-0.48)	(0.94)	(0.90)	(0.57)	(0.56)	(0.81)	(0.76)	(1.25)	(1.15)
Male Human Capital	0.447	2.993	0.542	0.542	0.690	0.690	0.818	0.818	1.324	1.324
	(0.57)	(1.08)	(0.49)	(0.45)	(0.65)	(0.65)	(0.70)	(0.60)	(1.05)	(0.96)
Fixed Capital Growth	0.0291	0.106**	0.0238	0.0238	0.0188	0.0188	0.0208	0.0208	0.0125	0.0125
	(1.87)	(3.30)	(0.90)	(0.61)	(0.67)	(0.47)	(0.76)	(0.53)	(0.46)	(0.32)
Constant	0.103	1.508	-1.57e-15	-1.57e-15	7.21e-16	7.21e-16	-1.38e-14	-1.38e-14	-4.49e-15	-4.49e-15
	(0.76)	(1.63)	(-0.00)	(-0.00)	(0.00)	(0.00)	(-0.00)	(-0.00)	(-0.00)	(-0.00)
Observations	540	540	540	540	520	520	460	460	480	480

t statistics in parentheses

* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

CD = 2.084, p = 0.037

Outliers

Table 6.14: Outliers (Women's Studies) - By Income Region

Country	GDP Growth	Human Capital (Female)	MD	MCD	MVE
<u>OECD Countries</u>					
Switzerland	0.036	0.920	0	1	1
Chile	0.144	0.794	0	1	1
Colombia	0.096	0.722	0	1	1
Costa Rica	0.019	0.757	0	1	1
Germany	0.049	0.907	0	1	1
Finland	0.034	0.913	0	1	1
United Kingdom	0.067	0.892	1	1	1
Ireland	0.015	0.893	0	1	1
Republic of Korea	0.070	0.852	0	0	1
Mexico	0.086	0.726	0	1	0
Norway	0.082	0.937	0	1	0
New Zealand	0.043	0.902	0	1	1
Slovenia	0.021	0.892	0	1	0
Sweden	0.052	0.907	1	1	1
<u>Emerging Markets</u>					
Argentina	0.137	0.829	0	1	1
Bulgaria	0.037	0.786	1	1	1
Brazil	0.149	0.718	0	1	0
Cabo Verde	0.076	0.634	0	1	0
Egypt	0.136	0.613	0	1	1
Fiji	0.102	0.687	0	1	1
Ghana	0.112	0.543	1	1	1
Gambia	0.041	0.424	0	1	0
Guatemala	0.053	0.576	0	1	1
Honduras	0.058	0.588	0	1	0
Croatia	-0.011	0.816	0	1	0

Iran	0.069	0.678	0	1	1
Jordan	0.058	0.658	0	1	1
Kenya	0.100	0.526	0	1	0
Kyrgyzstan	0.003	0.652	0	1	0
Morocco	0.065	0.538	0	1	0
Mexico	0.086	0.726	0	1	1
Malaysia	0.084	0.747	0	1	0
Panama	0.091	0.767	1	1	1
Peru	0.129	0.703	0	1	0
Thailand	0.126	0.735	0	1	0
Tajikistan	0.176	0.590	1	1	1
Ukraine	0.033	0.769	0	0	1
Uruguay	0.138	0.787	0	1	0
Zimbabwe	0.091	0.492	1	1	1
Developing Countries					
Bangladesh	0.130	0.494	0	0	1
Comoros	-0.026	0.479	0	1	1
Gabon	0.161	0.632	0	1	1
Cambodia	0.057	0.510	0	1	1
Mozambique	0.022	0.379	0	1	0
Mauritania	0.193	0.465	0	1	1
Nepal	0.105	0.516	0	0	1
Rwanda	0.094	0.472	0	0	1
Sierra Leone	0.043	0.396	0	1	1
Togo	0.062	0.417	0	1	0

1 = flagged as outlier by above diagnostic; 0 otherwise

6.2.5 Summary

We therefore find that after controlling for cross-section dependence, outliers, collinearity and unobserved factors with the CCE and MCD, MVE estimators, we find a negative and significant impact on economic growth when women are being disadvantaged compared to their male counterparts. The results agree with the general literature that improving the wellbeing of women improves economic growth. The results by region also agree with Barro's classical but controversial results as well as Seguino's hypothesis on the wage exploit of women. Therefore, bridging the gap between the genders is a worthwhile pursuit for institutions, especially for emerging markets and developing countries.

6.3 Conclusion

We have shown that controlling for cross-section dependence and outliers produce results that agree with the wider body of literature. We have showcased the performance of the CCE estimators coupled with robust trimming. We identify three areas of future research. Firstly, we have addressed the issue of leverage points (see Rouseeuw and Leroy, 1987 for details) with the MD, MCD and MVE estimators. A future area of research would look at addressing vertical outliers (outliers in the regressand) for the CCE estimators. Potential areas include considering robust alternatives to least squares in the CCE estimators. Estimators like the TCCEMG estimators address only block outliers (see Bramati and Croux, 2007), that is outliers at the cross-section level. Addressing outliers within each unit is another potential research area. Secondly, we have demonstrated that augmenting the CCE estimator with a robust trimming method improves its performance. Developing a robust analogue of the CCE estimator would be another boon to the robust estimation literature, as well as a huge benefit to the empirical field for cross-section dependence. Thirdly, we have identified regulatory quality and gender pay inequalities as determinants of economic growth. With data now widely available, perhaps we can look at the determinants for each country at the micro level. We note that no country is an island (North, 1989), therefore if sector level panel regressions for individual countries are considered, we need to consider foreign contributing factors. Terms of trade is a potential proxy, but perhaps the CCE estimators potentially can partial out these external effects. Overall, we hope that this paper has contributed some insights to the literature and the global economy to a certain degree.

References

Acemoglu, D., 2008. *Introduction to modern economic growth.* Princeton university press.

Rose-Ackerman, S., 2001. Trust, honesty and corruption: Reflection on the state-building process. *European Journal of Sociology/Archives Européennes de Sociologie, 42(3)*, pp.526-570.

Arellano, M. and Bond, S., 1991. Some tests of specification for panel data: Monte Carlo evidence and an application to employment equations. *The review of economic studies*, 58(2), pp.277-297.

Barro, R.J. and Lee, J.W., 1994, June. Sources of economic growth. *In Carnegie-Rochester conference series on public policy* (Vol. 40, pp. 1-46). North-Holland.

Barro, R.J., 1996. Democracy and growth. *Journal of economic growth*, 1, pp.1-27.

Bramati, M. and Croux, C., 2007. Robust estimators for the fixed effects panel data model. *The Econometrics Journal*, 10(3), pp.521-540.

Catrinescu, N., Leon-Ledesma, M., Piracha, M. and Quillin, B., 2009. Remittances, institutions, and economic growth. *World development*, 37(1), pp.81-92.

Charpe, M., Bridji, S. and McAdam, P., 2020. Labor share and growth in the long run. *Macroeconomic Dynamics*, 24(7), pp.1720-1757.

Chudik, A. and Pesaran, M.H., 2013. Large panel data models with cross-sectional dependence: a survey. *CAFE Research Paper*, (13.15).

Davies, L. and Gather, U., 1993. The identification of multiple outliers. *Journal of the American Statistical Association*, 88(423), pp.782-792.

De Maesschalck, R., Jouan-Rimbaud, D. and Massart, D.L., 2000. The mahalanobis distance. Chemometrics and intelligent laboratory systems, 50(1), pp.1-18.

Ditzen, J., 2018. Estimating dynamic common-correlated effects in Stata. *The Stata Journal*, 18(3), pp.585-617.

Feenstra, Robert C., Robert Inklaar and Marcel P. Timmer (2015), The Next Generation of the Penn World Table *American Economic Review*, 105(10), 3150-3182, available for download at www.ggdc.net/pwt

Hubert, M. and Debruyne, M., 2010. Minimum covariance determinant. *Wiley interdisciplinary re-*

views: Computational statistics, 2(1), pp.36-43.

Hubert, M., Debruyne, M. and Rousseeuw, P.J., 2018. Minimum covariance determinant and extensions. *Wiley Interdisciplinary Reviews: Computational Statistics*, 10(3), p.e1421.

Hubert, M., Debruyne, M. and Rousseeuw, P.J., 2018. Minimum covariance determinant and extensions. *Wiley Interdisciplinary Reviews: Computational Statistics*, 10(3), p.e1421.

Innovative instruments for turning data into knowledge (2023) *Global Data Lab.* Available at: https://globaldatalab (Accessed: 04 September 2023).

Kaufmann, D., Kraay, A. and Mastruzzi, M., 2011. The worldwide governance indicators: Methodology and analytical issues1. *Hague journal on the rule of law*, 3(2), pp.220-246.

Lee, Y. and Sul, D., 2020. Depth-Weighted Estimation of Heterogenous Agent Panel Data Models. *Mimeo, Syracuse University.*

Lee, Y. and Sul, D., 2022. Trimmed mean group estimation. In Essays in Honor of M. Hashem Pesaran: Panel Modeling, Micro Applications, and Econometric Methodology (pp. 177-202). *Emerald Publishing Limited.*

Lin, J.Y. and Nugent, J.B., 1995. Institutions and economic development. *Handbook of development economics*, 3, pp.2301-2370.

Mahalanobis, P.C., 1936. Mahalanobis distance. *In Proceedings National Institute of Science of India* (Vol. 49, No. 2, pp. 234-256).

Maronna, R.A., Martin, R.D., Yohai, V.J. and Salibián-Barrera, M., 2019. *Robust statistics: theory and methods (with R).* John Wiley & Sons.

North, D.C., 1989. Institutions and economic growth: An historical introduction. *World development*, 17(9), pp.1319-1332.

Rodrik, D., 2000. Institutions for high-quality growth: what they are and how to acquire them. *Studies in comparative international development*, 35, pp.3-31.

Pesaran, M.H., 2006. Estimation and inference in large heterogeneous panels with a multifactor error structure. *Econometrica*, 74(4), pp.967-1012.

Rousseeuw, P. and Leroy, A., 1987. *Robust regression and outlier detection.* Hoboken, NJ: Wiley-

Interscience.

Rousseeuw, P.J. and Hubert, M., 2011. Robust statistics for outlier detection. *Wiley interdisciplinary reviews: Data mining and knowledge discovery*, 1(1), pp.73-79.

Sachs, J.D. and Warner, A., 1995. Natural resource abundance and economic growth.

Sako, S., 2004. The NEPAD economic and corporate governance initiative: Building institutions for sustainable development. *The new partnership for Africa's development: macroeconomics, institutions, and poverty*, pp.126-141.

Samuelson, P.A. and Solow, R.M., 1956. A complete capital model involving heterogeneous capital goods. *The Quarterly Journal of Economics*, 70(4), pp.537-562.

Seguino, S. (2000) Gender inequality and economic growth: A cross-country analysis, *World Development*, 28(7), pp. 1211–1230. doi:10.1016/s0305-750x(00)00018-8.

Smits, J. and Permanyer, I., 2019. The subnational human development database. *Scientific data*, 6(1), pp.1-15.

Ulubasoglu, M.A. and Doucouliagos, C., 2004, February. Institutions and economic growth: A systems approach. *In Econometric Society 2004 Australasian Meetings Paper* (Vol. 63, pp. 1-33).

Valeriani, E. and Peluso, S., 2011. The impact of institutional quality on economic growth and development: An empirical study. *Journal of Knowledge Management, Economics and Information Technology*, 1(6), pp.1-25.

Van Aelst, S. and Rousseeuw, P., 2009. Minimum volume ellipsoid. *Wiley Interdisciplinary Reviews: Computational Statistics*, 1(1), pp.71-82.

Verardi, V. and Dehon, C., 2010. Multivariate outlier detection in Stata. *The Stata Journal*, 10(2), pp.259-266.

Appendix

6.A Institutions

6.A.1 Country List for Institutions

Table 6.15: Country List (and Mean Values by Country)

Country	GDP Growth	CRR	GOV	POL	REG	RUL	VOC
Angola	0.082	-1.303	-1.083	-0.784	-1.105	-1.304	-1.183
Albania	0.040	-0.689	-0.337	-0.107	0.000	-0.572	0.024
United Arab Emirates	0.041	0.913	1.024	0.832	0.741	0.574	-0.879
Argentina	0.035	-0.342	-0.035	-0.066	-0.510	-0.514	0.381
Armenia	0.062	-0.633	-0.208	-0.201	0.160	-0.382	-0.553
Australia	0.034	1.902	1.687	0.999	1.711	1.758	1.409
Austria	0.027	1.703	1.687	1.142	1.506	1.848	1.385
Azerbaijan	0.083	-1.094	-0.623	-0.721	-0.518	-0.840	-1.279
Burundi	0.026	-1.122	-1.301	-1.766	-1.107	-1.266	-1.148
Belgium	0.023	1.442	1.605	0.799	1.266	1.366	1.369
Benin	0.066	-0.606	-0.545	0.346	-0.457	-0.480	0.247
Bangladesh	0.058	-1.085	-0.752	-1.215	-0.944	-0.850	-0.448
Bulgaria	0.023	-0.200	0.012	0.280	0.549	-0.120	0.478
Bahrain	0.086	0.234	0.457	-0.449	0.674	0.427	-1.037
Belarus	0.037	-0.441	-0.820	0.205	-1.205	-1.001	-1.452
Belize	0.013	-0.140	-0.154	0.194	-0.252	-0.341	0.675
Bolivia	0.060	-0.649	-0.484	-0.477	-0.645	-0.857	0.002
Brazil	0.030	-0.133	-0.177	-0.196	0.068	-0.221	0.440
Barbados	-0.024	1.439	1.241	1.079	0.794	1.089	1.125
Brunei Darussalam	0.031	0.530	0.906	1.181	0.956	0.551	-0.786
Botswana	0.047	0.894	0.457	1.023	0.681	0.541	0.548
Central African Republic	0.007	-1.150	-1.522	-1.790	-1.206	-1.490	-1.093
Canada	0.024	1.939	1.803	1.073	1.650	1.752	1.468
Switzerland	0.032	2.054	1.950	1.336	1.673	1.870	1.534
Chile	0.039	1.322	1.069	0.505	1.398	1.171	1.025
China	0.065	-0.414	0.093	-0.437	-0.296	-0.468	-1.578

Côte d'Ivoire	0.055	-0.815	-0.936	-1.306	-0.664	-1.069	-0.823
Cameroon	0.044	-1.160	-0.867	-0.726	-0.870	-1.124	-1.042
D.R. of the Congo	0.041	-1.428	-1.635	-2.200	-1.520	-1.688	-1.467
Congo	0.062	-1.161	-1.195	-0.774	-1.242	-1.222	-1.121
Colombia	0.034	-0.330	-0.159	-1.534	0.165	-0.470	-0.153
Costa Rica	0.041	0.572	0.297	0.655	0.587	0.513	1.047
Cyprus	0.026	1.003	1.215	0.496	1.165	0.988	1.031
Czech Republic	0.023	0.431	0.902	0.956	1.106	0.944	0.966
Germany	0.020	1.830	1.596	0.875	1.574	1.664	1.381
Djibouti	0.043	-0.588	-0.911	-0.292	-0.720	-0.860	-1.184
Denmark	0.028	2.316	2.011	1.097	1.748	1.912	1.572
Dominican Republic	0.055	-0.755	-0.492	-0.018	-0.145	-0.593	0.126
Algeria	0.030	-0.645	-0.561	-1.269	-0.926	-0.848	-0.971
Ecuador	0.046	-0.685	-0.614	-0.505	-0.851	-0.848	-0.174
Egypt	0.062	-0.571	-0.428	-0.915	-0.461	-0.268	-1.106
Spain	0.032	1.042	1.222	0.069	1.109	1.134	1.137
Estonia	0.045	1.047	0.963	0.694	1.402	1.048	1.100
Finland	0.023	2.268	2.064	1.377	1.765	1.977	1.550
Fiji	0.037	0.325	-0.295	0.348	-0.313	-0.217	-0.304
France	0.023	1.353	1.488	0.459	1.154	1.417	1.228
Gabon	0.044	-0.833	-0.723	0.210	-0.403	-0.519	-0.805
United Kingdom	0.025	1.823	1.654	0.474	1.756	1.692	1.329
Ghana	0.051	-0.159	-0.157	-0.019	-0.134	-0.010	0.344
Guinea	0.005	-1.038	-1.066	-1.254	-0.981	-1.349	-1.060
Gambia	0.025	-0.560	-0.682	0.143	-0.520	-0.405	-0.938
Greece	0.011	0.155	0.506	0.231	0.637	0.583	0.887
Guatemala	0.043	-0.724	-0.665	-0.731	-0.279	-1.042	-0.301
Guyana	0.051	-0.472	-0.191	-0.373	-0.471	-0.479	0.180
Honduras	0.038	-0.863	-0.665	-0.423	-0.414	-0.964	-0.346
Croatia	0.034	0.029	0.472	0.548	0.324	0.057	0.444
Haiti	0.018	-1.279	-1.609	-1.017	-1.074	-1.351	-0.815
Hungary	0.026	0.408	0.707	0.842	0.951	0.726	0.852

Indonesia	0.052	-0.734	-0.302	-1.051	-0.297	-0.585	-0.110
India	0.072	-0.394	-0.038	-1.115	-0.355	0.060	0.416
Ireland	0.071	1.584	1.503	1.144	1.670	1.613	1.355
Iran	0.049	-0.637	-0.498	-1.026	-1.436	-0.854	-1.364
Iraq	0.085	-1.396	-1.432	-2.272	-1.371	-1.649	-1.307
Iceland	0.030	2.009	1.699	1.370	1.334	1.744	1.439
Israel	0.032	0.943	1.212	-1.170	1.120	0.975	0.673
Italy	0.014	0.302	0.532	0.541	0.849	0.512	1.019
Jamaica	0.023	-0.231	0.238	-0.035	0.177	-0.324	0.574
Jordan	0.078	0.170	0.111	-0.350	0.186	0.291	-0.660
Japan	0.004	1.372	1.432	1.045	1.088	1.361	1.013
Kazakhstan	0.069	-0.924	-0.494	0.112	-0.277	-0.833	-1.117
Kenya	0.052	-0.992	-0.548	-1.184	-0.284	-0.786	-0.334
Kyrgyzstan	0.038	-1.129	-0.688	-0.743	-0.379	-1.036	-0.751
Cambodia	0.065	-1.153	-0.828	-0.359	-0.430	-1.104	-0.969
Republic of Korea	0.031	0.471	0.956	0.381	0.857	0.969	0.704
Kuwait	0.050	0.219	0.015	0.251	0.122	0.411	-0.512
Lao People's DR	0.080	-1.011	-0.769	-0.028	-1.036	-0.989	-1.646
Liberia	0.101	-0.886	-1.377	-1.082	-1.270	-1.181	-0.489
Sri Lanka	0.054	-0.269	-0.171	-0.901	-0.129	0.052	-0.299
Lesotho	0.007	0.015	-0.442	0.001	-0.511	-0.163	-0.063
Lithuania	0.040	0.383	0.744	0.748	1.045	0.723	0.921
Luxembourg	0.044	1.979	1.751	1.413	1.719	1.800	1.551
Latvia	0.035	0.262	0.691	0.522	0.981	0.664	0.797
Morocco	0.030	-0.297	-0.199	-0.349	-0.148	-0.178	-0.623
Republic of Moldova	0.040	-0.725	-0.583	-0.213	-0.176	-0.389	-0.175
Maldives	0.072	-0.505	0.066	0.394	0.039	-0.275	-0.621
Mexico	0.032	-0.516	0.124	-0.599	0.258	-0.519	0.114
Mali	0.068	-0.685	-0.899	-0.607	-0.505	-0.449	-0.008
Myanmar	0.075	-1.247	-1.358	-1.161	-1.737	-1.407	-1.716
Mongolia	0.074	-0.425	-0.392	0.710	-0.248	-0.161	0.257
Mozambique	0.051	-0.599	-0.634	-0.056	-0.538	-0.759	-0.212

Mauritania	0.042	-0.633	-0.680	-0.412	-0.570	-0.742	-0.848
Mauritius	0.024	0.271	0.783	0.906	0.682	0.887	0.852
Malawi	0.027	-0.596	-0.628	-0.075	-0.590	-0.266	-0.204
Malaysia	0.048	0.199	0.989	0.217	0.573	0.407	-0.375
Namibia	0.039	0.351	0.224	0.693	0.183	0.241	0.413
Niger	0.041	-0.748	-0.799	-0.745	-0.658	-0.618	-0.455
Nicaragua	0.029	-0.764	-0.812	-0.320	-0.438	-0.749	-0.372
Netherlands	0.027	2.017	1.858	1.082	1.816	1.791	1.539
Norway	0.037	2.107	1.884	1.286	1.502	1.939	1.614
Nepal	0.048	-0.694	-0.814	-1.274	-0.658	-0.609	-0.506
New Zealand	0.031	2.261	1.753	1.357	1.831	1.874	1.562
Pakistan	0.047	-0.950	-0.633	-2.125	-0.670	-0.832	-0.861
Panama	0.073	-0.373	0.144	0.115	0.463	-0.119	0.523
Peru	0.049	-0.385	-0.333	-0.748	0.337	-0.605	0.044
Philippines	0.039	-0.594	0.022	-1.221	-0.039	-0.428	0.065
Poland	0.041	0.541	0.566	0.696	0.855	0.618	0.962
Portugal	0.022	1.033	1.088	1.015	0.983	1.125	1.236
Paraguay	0.048	-1.052	-0.895	-0.553	-0.436	-0.820	-0.184
Qatar	0.126	0.810	0.668	0.982	0.446	0.633	-0.929
Romania	0.043	-0.305	-0.179	0.211	0.398	0.066	0.435
Russian Federation	0.041	-0.955	-0.407	-0.952	-0.362	-0.871	-0.814
Rwanda	0.082	0.055	-0.399	-0.633	-0.469	-0.512	-1.281
Saudi Arabia	0.066	-0.020	-0.060	-0.346	0.000	0.083	-1.710
Sudan	0.058	-1.314	-1.318	-2.189	-1.401	-1.372	-1.737
Senegal	0.036	-0.177	-0.333	-0.250	-0.238	-0.158	0.090
Singapore	0.075	2.152	2.139	1.248	1.968	1.618	-0.085
Sierra Leone	0.025	-0.824	-1.272	-0.514	-1.031	-1.008	-0.382
El Salvador	0.073	-0.491	-0.336	-0.022	0.044	-0.702	0.061
Serbia	0.033	-0.497	-0.276	-0.464	-0.283	-0.540	-0.021
Sao Tome and Principe	0.056	-0.052	-0.688	0.397	-0.765	-0.465	0.264
Slovakia	0.024	0.210	0.722	0.910	0.910	0.449	0.903
Slovenia	0.022	0.907	0.972	1.025	0.785	0.997	1.060

Sweden	0.026	2.176	1.869	1.196	1.656	1.884	1.572
Eswatini	0.005	-0.120	-0.568	-0.213	-0.339	-0.534	-1.363
Syria	0.085	-1.145	-1.095	-1.303	-1.318	-1.031	-1.748
Togo	0.044	-0.880	-1.240	-0.415	-0.822	-0.840	-0.970
Thailand	0.034	-0.359	0.235	-0.697	0.153	0.028	-0.381
Tajikistan	0.048	-1.222	-1.083	-1.103	-1.126	-1.242	-1.444
Trinidad and Tobago	0.047	-0.054	0.288	0.064	0.412	-0.008	0.568
Tunisia	0.022	-0.099	0.223	-0.273	-0.134	-0.053	-0.572
Turkey	0.049	-0.113	0.144	-1.094	0.232	-0.079	-0.280
Tanzania	0.067	-0.592	-0.602	-0.384	-0.485	-0.401	-0.309
Uganda	0.057	-0.957	-0.569	-1.029	-0.209	-0.450	-0.641
Ukraine	0.027	-0.980	-0.643	-0.674	-0.476	-0.822	-0.191
Uruguay	0.024	1.205	0.495	0.861	0.581	0.628	1.058
United States	0.025	1.447	1.562	0.461	1.491	1.565	1.165
Venezuela	-0.135	-1.165	-1.093	-1.127	-1.360	-1.591	-0.772
Vietnam	0.070	-0.568	-0.228	0.234	-0.579	-0.394	-1.406
Yemen	0.062	-1.179	-1.231	-2.058	-0.911	-1.327	-1.252
South Africa	0.027	0.176	0.362	-0.151	0.447	0.045	0.655
Zambia	0.075	-0.544	-0.799	0.228	-0.540	-0.413	-0.275
Zimbabwe	-0.020	-1.223	-1.160	-0.946	-1.731	-1.518	-1.315

Key	
CRR	Control of Corruption
GOV	Government Effectiveness
POL	Poltical Stability
REG	Regulatory Quality
RUL	Rule of Law
VOC	Voice of Accountability

6.A.2 Full Results for Institutions

Untreated Results

These results have not been treated for outliers.

Table 6.16: Estimation Results - World

	FE	CCEMG	CCEP	CEEPnw	CCEPwpn
Control of Corruption	0.0426	-0.500	-0.500	-0.0206	-0.0206
	(1.31)	(-0.90)	(-0.90)	(-0.63)	(-0.47)
Government Effectiveness	-0.0907**	0.878	0.878	-0.0430	-0.0430
	(-2.72)	(1.25)	(1.25)	(-1.37)	(-0.98)
Political Stability	0.0309	0.115	0.115	0.0325	0.0325
	(1.74)	(0.23)	(0.23)	(1.49)	(1.22)
Regulatory Quality	0.147***	-0.158	-0.158	0.173***	0.173***
	(4.76)	(-0.25)	(-0.25)	(5.03)	(4.03)
Rule of Law	-0.0510	-0.0277	-0.0277	0.105*	0.105
	(-1.38)	(-0.03)	(-0.03)	(2.58)	(1.88)
Voice of Accountability	0.00647	2.930	2.930	-0.0545	-0.0545
	(0.23)	(1.04)	(1.04)	(-1.81)	(-1.45)
Fixed Capital Growth	0.0237***	-0.0970	-0.0970	-0.0255	-0.0255
	(3.33)	(-1.27)	(-1.27)	(-1.89)	(-1.61)
Human Capital	0.316***	0.108	0.108	1.216***	1.216***
	(9.12)	(0.03)	(0.03)	(5.20)	(3.74)
$y_{i,t-1}$	-0.0854***	-0.219	-0.219	-0.497***	-0.497***
	(-13.51)	(-0.52)	(-0.52)	(-15.95)	(-11.85)
Constant	1.109***	-1.729	-1.729	6.25e-14	6.25e-14
	(13.03)	(-0.26)	(-0.26)	(0.00)	(0.00)
Observations	2976	2919	2919	2919	2919

t statistics in parentheses

CD = 79.985, p = 0.000,

* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Table 6.17: Estimation Results - OECD

	FE	CCEMG	CCEP	CEEPnw	CCEPwpn
Control of Corruption	0.150***	2.524	2.524	0.0351	0.0351
	(3.48)	(0.94)	(0.94)	(0.87)	(0.66)
Government Effectiveness	-0.0619	-0.223	-0.223	0.0214	0.0214
	(-1.19)	(-0.19)	(-0.19)	(0.48)	(0.38)
Political Stability	0.0684**	-0.840	-0.840	0.0261	0.0261
	(2.63)	(-1.67)	(-1.67)	(1.23)	(1.20)
Regulatory Quality	0.0545	0.582	0.582	0.196***	0.196***
	(1.07)	(1.25)	(1.25)	(4.33)	(4.21)
Rule of Law	-0.0364	-0.664	-0.664	-0.0546	-0.0546
	(-0.62)	(-0.32)	(-0.32)	(-0.87)	(-0.70)
Voice of Accountability	0.0138	-0.747	-0.747	0.0803	0.0803
	(0.23)	(-0.71)	(-0.71)	(1.18)	(0.99)
Fixed Capital Growth	0.137***	0.127	0.127	0.0214	0.0214
	(8.84)	(1.31)	(1.31)	(0.87)	(0.58)
Human Capital	0.638***	2.687	2.687	1.762***	1.762***
	(6.76)	(0.59)	(0.59)	(5.92)	(6.24)
$y_{i,t-1}$	-0.122***	-1.498***	-1.498***	-0.547***	-0.547***
	(-7.95)	(-5.60)	(-5.60)	(-15.12)	(-16.89)
Constant	1.554***	-0.211	-0.211	-3.21e-14	-3.21e-14
	(7.37)	(-0.03)	(-0.03)	(-0.00)	(-0.00)
Observations	798	798	798	798	798

t statistics in parentheses

CD = 32.729, p = 0.000

* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Table 6.18: Estimation Results - Emerging Markets

	FE	CCEMG	CCEP	CEEPnw	CCEPwpn
Control of Corruption	0.0838	0.147	0.147	-0.0360	-0.0360
	(1.93)	(0.39)	(0.39)	(-0.91)	(-0.72)
Government Effectiveness	-0.122**	-1.510	-1.510	-0.0108	-0.0108
	(-2.88)	(-1.42)	(-1.42)	(-0.30)	(-0.22)
Political Stability	0.0222	-0.0175	-0.0175	0.0815***	0.0815**
	(0.87)	(-0.03)	(-0.03)	(3.30)	(2.78)
Regulatory Quality	0.179***	0.619	0.619	0.114**	0.114*
	(4.57)	(1.39)	(1.39)	(2.96)	(2.25)
Rule of Law	-0.0519	-0.0184	-0.0184	0.0844	0.0844
	(-1.07)	(-0.03)	(-0.03)	(1.73)	(1.22)
Voice of Accountability	0.0197	-0.231	-0.231	-0.0348	-0.0348
	(0.54)	(-0.25)	(-0.25)	(-0.94)	(-0.70)
Fixed Capital Growth	0.0547***	-0.0431	-0.0431	0.0205	0.0205
	(5.82)	(-0.48)	(-0.48)	(1.28)	(0.90)
Human Capital	0.365***	-0.418	-0.418	1.841***	1.841***
	(6.40)	(-0.24)	(-0.24)	(7.07)	(4.40)
$y_{i,t-1}$	-0.0791***	-0.931***	-0.931***	-0.517***	-0.517***
	(-9.58)	(-5.38)	(-5.38)	(-16.34)	(-11.47)
Constant	1.049***	-0.434	-0.434	-3.24e-14	-3.24e-14
	(9.18)	(-0.08)	(-0.08)	(-0.00)	(-0.00)
Observations	1678	1659	1659	1659	1659

t statistics in parentheses

CD = 56.576, p = 0.000

* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Table 6.19: Estimation Results - Developing Countries

	FE	CCEMG	CCEP	CEEPnw	CCEPwpn
Control of Corruption	-0.0208	-0.354	-0.354	-0.0965	-0.0965
	(-0.30)	(-1.08)	(-1.08)	(-1.37)	(-0.98)
Government Effectiveness	-0.0563	-0.236	-0.236	-0.115	-0.115
	(-0.66)	(-0.64)	(-0.64)	(-1.20)	(-1.14)
Political Stability	0.0871*	0.164	0.164	0.103	0.103
	(2.51)	(0.50)	(0.50)	(1.76)	(1.44)
Regulatory Quality	-0.00212	-0.0366	-0.0366	0.154	0.154
	(-0.03)	(-0.07)	(-0.07)	(1.88)	(1.50)
Rule of Law	-0.104	0.120	0.120	0.0174	0.0174
	(-1.26)	(0.23)	(0.23)	(0.16)	(0.17)
Voice of Accountability	-0.0670	0.629	0.629	0.0862	0.0862
	(-1.13)	(1.32)	(1.32)	(1.45)	(1.07)
Fixed Capital Growth	-0.0691***	0.0264	0.0264	-0.0217	-0.0217
	(-4.84)	(0.28)	(0.28)	(-1.31)	(-0.91)
Human Capital	0.416***	0.785	0.785	0.978***	0.978*
	(6.59)	(0.41)	(0.41)	(3.80)	(2.48)
$y_{i,t-1}$	-0.139***	-0.861**	-0.861**	-0.650***	-0.650***
	(-7.96)	(-3.30)	(-3.30)	(-16.00)	(-13.18)
Constant	1.776***	4.665	4.665	5.50e-14	5.50e-14
	(8.17)	(0.40)	(0.40)	(0.00)	(0.00)
Observations	626	588	588	588	588

t statistics in parentheses

CD = 2.563, p = 0.0103

* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

MD-trimmed results

Table 6.20: TCCEMG(Lee and Sul, 2020) - World

	FE	CCEMG	CCEP	CEEPnw	CCEPwpn
Control of Corruption	0.0505	0.316	0.0573	0.0573	0.0573
	(1.50)	(0.50)	(0.33)	(1.83)	(1.28)
goveff	-0.0487	-0.894	0.0259	0.0259	0.0259
	(-1.39)	(-1.87)	(0.10)	(0.76)	(0.56)
Political Stability	0.0392*	-0.0181	0.0701	0.0701***	0.0701**
	(2.15)	(-0.10)	(0.60)	(3.41)	(2.63)
Regulatory Quality	0.126***	0.772	0.0964	0.0964**	0.0964*
	(3.95)	(1.34)	(0.47)	(2.77)	(2.50)
Rule of Law	-0.105**	0.674	0.0482	0.0482	0.0482
	(-2.66)	(1.19)	(0.20)	(0.93)	(0.61)
Voice of Accountability	0.00766	-2.168	-0.0412	-0.0412	-0.0412
	(0.27)	(-1.42)	(-0.09)	(-1.34)	(-0.97)
gfcfgrowth	0.0359***	0.0297	-0.0220	-0.0220*	-0.0220
	(4.87)	(0.45)	(-0.82)	(-2.07)	(-1.42)
humancap	0.297***	3.763*	0.993	0.993***	0.993***
	(7.96)	(2.20)	(1.03)	(5.82)	(3.93)
$y_{I,t-1}$	-0.0850***	-0.799***	-0.396***	-0.396***	-0.396***
	(-12.00)	(-6.10)	(-3.87)	(-15.41)	(-10.24)
Constant	1.107***	7.190	7.51e-15	7.51e-15	7.51e-15
	(11.52)	(1.44)	(0.00)	(0.00)	(0.00)
NT	2577	2520	2520	2520	2520

t statistics in parentheses

* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Table 6.21: TCCEMG(Lee and Sul, 2020) - OECD

	FE	CCEMG	CCEP	CEEPnw	CCEPwpn
Control of Corruption	0.147***	-0.289	0.0833	0.0833*	0.0833
	(3.34)	(-0.31)	(0.25)	(2.32)	(1.69)
goveff	-0.0586	-1.420	-0.0333	-0.0333	-0.0333
	(-1.02)	(-1.03)	(-0.07)	(-0.71)	(-0.63)
Political Stability	0.0583*	-0.856	0.0409	0.0409	0.0409
	(2.02)	(-1.61)	(0.13)	(1.96)	(1.90)
Regulatory Quality	0.0803	1.109	0.201	0.201***	0.201***
	(1.48)	(1.88)	(0.62)	(3.95)	(3.43)
Rule of Law	-0.0442	0.640	-0.0946	-0.0946	-0.0946
	(-0.71)	(0.28)	(-0.22)	(-1.40)	(-1.19)
Voice of Accountability	-0.0174	-1.026	-0.0537	-0.0537	-0.0537
	(-0.27)	(-0.58)	(-0.06)	(-0.59)	(-0.42)
gfcfgrowth	0.149***	0.0855	0.0580	0.0580**	0.0580*
	(9.35)	(1.27)	(0.76)	(2.83)	(2.32)
humancap	0.513***	3.490*	2.077	2.077***	2.077***
	(5.00)	(2.23)	(0.92)	(4.57)	(6.41)
$y_{I,t-1}$	-0.0975***	-0.878*	-0.561*	-0.561***	-0.561***
	(-5.62)	(-2.45)	(-2.28)	(-8.87)	(-12.53)
Constant	1.219***	-4.473	-1.66e-13	-1.66e-13	-1.66e-13
	(5.21)	(-0.73)	(-0.00)	(-0.00)	(-0.00)
NT	798	609	609	609	609

t statistics in parentheses

* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Table 6.22: TCCEMG(Lee and Sul, 2020) - Emerging Markets

	FE	CCEMG	CCEP	CEEPnw	CCEPwpn
Control of Corruption	0.0891	0.488	0.00144	0.00144	0.00144
	(1.88)	(0.88)	(0.01)	(0.03)	(0.03)
goveff	-0.100*	0.404	0.00463	0.00463	0.00463
	(-2.23)	(0.61)	(0.03)	(0.13)	(0.09)
Political Stability	-0.00514	0.607	0.0612	0.0612*	0.0612
	(-0.18)	(1.24)	(0.48)	(2.12)	(1.78)
Regulatory Quality	0.181***	0.206	0.135	0.135**	0.135*
	(4.26)	(0.60)	(0.79)	(2.87)	(2.53)
Rule of Law	-0.133*	-1.101*	0.0577	0.0577	0.0577
	(-2.48)	(-2.35)	(0.26)	(0.97)	(0.70)
Voice of Accountability	0.0565	-0.131	-0.0760	-0.0760	-0.0760
	(1.38)	(-0.32)	(-0.52)	(-1.48)	(-1.10)
gfcfgrowth	0.0678***	0.0101	0.00651	0.00651	0.00651
	(6.45)	(0.12)	(0.15)	(0.45)	(0.36)
humancap	0.264***	-0.452	1.350	1.350***	1.350*
	(3.98)	(-0.32)	(1.38)	(3.48)	(2.45)
$y_{I,t-1}$	-0.0636***	-0.626***	-0.554***	-0.554***	-0.554***
	(-6.82)	(-6.48)	(-4.16)	(-15.42)	(-11.29)
Constant	0.852***	-9.280	5.31e-14	5.31e-14	5.31e-14
	(6.51)	(-1.53)	(0.00)	(0.00)	(0.00)
NT	1426	1407	1407	1407	1407

t statistics in parentheses

* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Table 6.23: TCCEMG(Lee and Sul, 2020) - Developing Countries

	FE	CCEMG	CCEP	CEEPnw	CCEPwpn
Control of Corruption	-0.000939	-0.786	-0.105	-0.105	-0.105
	(-0.01)	(-1.01)	(-0.16)	(-1.56)	(-1.48)
goveff	-0.0775	0.0586	-0.199	-0.199	-0.199
	(-0.74)	(0.06)	(-0.34)	(-1.96)	(-1.86)
Political Stability	0.0857*	0.538	0.0395	0.0395	0.0395
	(2.20)	(2.07)	(0.11)	(0.57)	(0.48)
Regulatory Quality	0.00397	0.305	0.296	0.296**	0.296**
	(0.05)	(0.46)	(0.53)	(2.82)	(2.96)
Rule of Law	-0.0811	1.190	0.0798	0.0798	0.0798
	(-0.87)	(0.91)	(0.11)	(0.67)	(0.61)
Voice of Accountability	-0.0988	0.892	0.0898	0.0898	0.0898
	(-1.45)	(1.11)	(0.15)	(1.45)	(1.17)
gfcfgrowth	-0.0716***	-0.0202	-0.0286	-0.0286	-0.0286
	(-4.56)	(-0.29)	(-0.47)	(-1.05)	(-0.76)
humancap	0.542***	1.934	0.953	0.953**	0.953
	(7.00)	(0.91)	(0.51)	(2.64)	(1.86)
$y_{I,t-1}$	-0.162***	-0.750**	-0.602***	-0.602***	-0.602***
	(-7.97)	(-3.55)	(-4.84)	(-14.05)	(-14.89)
Constant	2.089***	37.30*	6.73e-14	6.73e-14	6.73e-14
	(8.21)	(2.18)	(0.00)	(0.00)	(0.00)
NT	521	483	483	483	483

t statistics in parentheses

* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

MCD-trimmed results

Table 6.24: TCCEMG(MCD) - World

	FE	CCEMG	CCEP	CEEPnw	CCEPwpn
Control of Corruption	0.0304	0.554	-0.00322	-0.00322	-0.00322
	(0.69)	(0.56)	(-0.02)	(-0.08)	(-0.05)
goveff	-0.0166	-0.0260	0.0499	0.0499	0.0499
	(-0.40)	(-0.06)	(0.30)	(1.21)	(0.86)
Political Stability	0.0171	0.189	0.0314	0.0314	0.0314
	(0.79)	(0.58)	(0.32)	(1.10)	(0.79)
Regulatory Quality	0.137***	0.357	0.102	0.102*	0.102*
	(3.57)	(1.24)	(0.83)	(2.00)	(2.24)
Rule of Law	-0.129**	-0.200	0.0794	0.0794	0.0794
	(-2.69)	(-0.16)	(0.39)	(1.47)	(0.92)
Voice of Accountability	0.0283	1.124	0.0449	0.0449	0.0449
	(0.87)	(1.03)	(0.16)	(1.22)	(0.94)
gfcfgrowth	0.0706***	-0.156*	-0.0423	-0.0423**	-0.0423**
	(7.90)	(-2.04)	(-1.12)	(-2.98)	(-2.74)
humancap	0.195***	-0.0299	0.557	0.557**	0.557*
	(4.42)	(-0.01)	(1.08)	(2.92)	(2.47)
$y_{I,t-1}$	-0.0683***	-0.740**	-0.427***	-0.427***	-0.427***
	(-7.85)	(-2.86)	(-5.77)	(-13.49)	(-10.92)
Constant	0.866***	-11.19*	2.70e-14	2.70e-14	2.70e-14
	(7.47)	(-2.26)	(0.00)	(0.00)	(0.00)
NT	1657	1638	1638	1638	1638

t statistics in parentheses

TCCEMG, MCD

* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Table 6.25: TCCEMG(MCD) - OECD

	FE	CCEMG	CCEP	CEEPnw	CCEPwpn
Control of Corruption	0.165***	0.833	0.0105	0.0105	0.0105
	(3.41)	(0.93)	(0.03)	(0.26)	(0.19)
goveff	-0.0831	0.109	-0.151	-0.151*	-0.151*
	(-1.27)	(0.09)	(-0.75)	(-2.43)	(-1.98)
Political Stability	0.0675*	-0.131	0.0768	0.0768***	0.0768**
	(2.28)	(-0.46)	(0.84)	(3.61)	(2.95)
Regulatory Quality	0.0920	-0.364	0.219	0.219***	0.219**
	(1.56)	(-0.92)	(1.13)	(3.55)	(3.13)
Rule of Law	-0.0296	-0.980	-0.0211	-0.0211	-0.0211
	(-0.44)	(-1.15)	(-0.07)	(-0.29)	(-0.22)
Voice of Accountability	-0.0138	-0.378	0.127	0.127	0.127
	(-0.20)	(-0.41)	(0.43)	(1.30)	(1.01)
gfcfgrowth	0.139***	0.0670	0.0430	0.0430	0.0430
	(7.98)	(0.70)	(0.93)	(1.84)	(1.40)
humancap	0.579***	3.638*	2.512***	2.512***	2.512***
	(5.13)	(2.46)	(3.60)	(7.18)	(5.47)
$y_{I,t-1}$	-0.114***	-0.610*	-0.582***	-0.582***	-0.582***
	(-5.64)	(-2.73)	(-4.39)	(-14.79)	(-11.92)
Constant	1.416***	-0.357	-1.35e-13	-1.35e-13	-1.35e-13
	(5.24)	(-0.06)	(-0.00)	(-0.00)	(-0.00)
NT	504	504	504	504	504

t statistics in parentheses

TCCEMG, MCD

* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Table 6.26: TCCEMG(MCD) - Emerging Markets

	FE	CCEMG	CCEP	CEEPnw	CCEPwpn
Control of Corruption	0.113	0.0742	0.145	0.145**	0.145*
	(1.93)	(0.17)	(0.53)	(2.95)	(2.28)
goveff	-0.114*	0.457	-0.0719	-0.0719	-0.0719
	(-2.01)	(0.60)	(-0.20)	(-1.83)	(-1.46)
Political Stability	-0.0237	-0.691	0.102	0.102**	0.102**
	(-0.67)	(-1.82)	(0.46)	(3.01)	(2.86)
Regulatory Quality	0.141*	0.769	0.0170	0.0170	0.0170
	(2.49)	(1.31)	(0.06)	(0.31)	(0.23)
Rule of Law	-0.135*	-1.005	0.0531	0.0531	0.0531
	(-2.12)	(-0.77)	(0.12)	(0.83)	(0.64)
Voice of Accountability	0.127**	1.582	-0.0662	-0.0662	-0.0662
	(2.59)	(1.52)	(-0.12)	(-1.53)	(-1.47)
gfcfgrowth	0.0864***	-0.126	0.0199	0.0199	0.0199
	(7.49)	(-1.34)	(0.47)	(1.05)	(0.86)
humancap	0.180*	3.247	1.494	1.494***	1.494**
	(2.35)	(1.39)	(1.15)	(3.83)	(2.65)
$y_{I,t-1}$	-0.0491***	-1.089**	-0.485**	-0.485***	-0.485***
	(-4.65)	(-3.12)	(-2.59)	(-13.42)	(-10.34)
Constant	0.662***	-10.66	4.68e-14	4.68e-14	4.68e-14
	(4.42)	(-1.14)	(0.00)	(0.00)	(0.00)
NT	1027	1008	1008	1008	1008

t statistics in parentheses

TCCEMG, MCD

* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Table 6.27: TCCEMG(MCD) - Developing Countries

	FE	CCEMG	CCEP	CEEPnw	CCEPwpn
Control of Corruption	0.00847	-0.602	-0.115	-0.115	-0.115
	(0.09)	(-1.06)	(-0.27)	(-1.19)	(-0.92)
goveff	-0.0275	-0.900	-0.114	-0.114	-0.114
	(-0.22)	(-1.61)	(-0.17)	(-0.99)	(-0.84)
Political Stability	0.0801	0.216	0.113	0.113	0.113
	(1.76)	(0.53)	(0.57)	(1.75)	(1.45)
Regulatory Quality	0.00688	-0.936	0.269	0.269*	0.269*
	(0.07)	(-0.63)	(0.44)	(2.09)	(2.06)
Rule of Law	-0.0670	0.612	0.0806	0.0806	0.0806
	(-0.65)	(1.20)	(0.10)	(0.58)	(0.49)
Voice of Accountability	-0.120	1.991	0.0864	0.0864	0.0864
	(-1.32)	(1.44)	(0.11)	(0.63)	(0.44)
gfcfgrowth	-0.0563**	-0.0204	-0.0451	-0.0451	-0.0451
	(-2.81)	(-0.23)	(-0.77)	(-1.55)	(-1.04)
humancap	0.523***	-4.510	0.947	0.947*	0.947
	(5.64)	(-1.12)	(0.35)	(2.25)	(1.71)
$y_{I,t-1}$	-0.153***	-1.224***	-0.717**	-0.717***	-0.717***
	(-6.55)	(-4.75)	(-3.16)	(-14.30)	(-13.68)
Constant	1.988***	-6.356	1.12e-13	1.12e-13	1.12e-13
	(6.71)	(-0.42)	(0.00)	(0.00)	(0.00)
NT	418	399	399	399	399

t statistics in parentheses

TCCEMG, MCD

* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

MVE-trimmed results

Table 6.28: TCCEMG(MVE) - World

	FE	CCEMG	CCEP	CEEPnw	CCEPwpn
Control of Corruption	0.0395	1.073*	-0.0347	-0.0347	-0.0347
	(1.00)	(2.07)	(-0.16)	(-0.87)	(-0.64)
goveff	-0.0421	-0.687	0.0337	0.0337	0.0337
	(-1.02)	(-1.78)	(0.15)	(0.80)	(0.55)
Political Stability	0.0318	0.260	0.0571	0.0571	0.0571
	(1.47)	(1.69)	(0.49)	(1.78)	(1.27)
Regulatory Quality	0.134***	0.412	0.220	0.220***	0.220***
	(3.59)	(1.65)	(1.36)	(4.01)	(4.54)
Rule of Law	-0.111*	-0.626	0.0560	0.0560	0.0560
	(-2.39)	(-0.97)	(0.22)	(0.95)	(0.63)
Voice of Accountability	0.00702	-0.495	0.0690	0.0690	0.0690
	(0.22)	(-1.35)	(0.44)	(1.59)	(1.08)
gfcfgrowth	0.0383***	-0.0425	-0.0332	-0.0332**	-0.0332*
	(4.74)	(-0.81)	(-1.09)	(-2.74)	(-1.98)
humancap	0.224***	1.965	0.782	0.782**	0.782*
	(5.24)	(1.35)	(1.06)	(3.02)	(2.55)
$y_{I,t-1}$	-0.0752***	-0.838***	-0.480***	-0.480***	-0.480***
	(-9.08)	(-5.88)	(-3.91)	(-13.52)	(-10.36)
Constant	0.964***	-3.606	4.97e-14	4.97e-14	4.97e-14
	(8.67)	(-0.56)	(0.00)	(0.00)	(0.00)
NT	1907	1869	1869	1869	1869

t statistics in parentheses

* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Table 6.29: TCCEMG(MVE) - OECD

	FE	CCEMG	CCEP	CEEPnw	CCEPwpn
Control of Corruption	0.138**	0.889	0.0747	0.0747	0.0747
	(2.61)	(0.94)	(0.13)	(1.71)	(1.25)
goveff	-0.0322	-3.297	-0.00526	-0.00526	-0.00526
	(-0.45)	(-1.57)	(-0.00)	(-0.10)	(-0.10)
Political Stability	0.0589	-0.658	0.0529	0.0529*	0.0529*
	(1.70)	(-1.33)	(0.16)	(2.22)	(2.30)
Regulatory Quality	0.0584	0.477	0.208	0.208**	0.208*
	(0.88)	(1.04)	(0.62)	(3.06)	(2.52)
Rule of Law	-0.0262	0.00557	-0.234	-0.234**	-0.234**
	(-0.34)	(0.00)	(-0.34)	(-3.07)	(-2.80)
Voice of Accountability	-0.0289	-0.896	-0.0686	-0.0686	-0.0686
	(-0.38)	(-0.64)	(-0.07)	(-0.70)	(-0.58)
gfcfgrowth	0.119***	-0.153	0.0363	0.0363	0.0363
	(6.26)	(-1.27)	(0.32)	(1.63)	(1.09)
humancap	0.643***	5.118*	2.750	2.750***	2.750***
	(5.36)	(2.38)	(0.90)	(5.04)	(6.18)
$y_{I,t-1}$	-0.124***	-0.985**	-0.683	-0.683***	-0.683***
	(-6.41)	(-2.96)	(-1.97)	(-11.62)	(-13.95)
Constant	1.593***	6.895	-9.03e-14	-9.03e-14	-9.03e-14
	(6.02)	(0.74)	(-0.00)	(-0.00)	(-0.00)
NT	546	546	546	546	546

t statistics in parentheses

CD = 32.728855266992, p = 0

* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Table 6.30: TCCEMG(MVE) - Emerging Markets

	FE	CCEMG	CCEP	CEEPnw	CCEPwpn
Control of Corruption	0.103	-0.0329	0.0490	0.0490	0.0490
	(1.87)	(-0.07)	(0.30)	(1.02)	(0.81)
goveff	-0.0946	0.188	0.0381	0.0381	0.0381
	(-1.79)	(0.29)	(0.16)	(1.00)	(0.81)
Political Stability	-0.00752	-0.242	0.122	0.122***	0.122***
	(-0.23)	(-0.40)	(0.84)	(4.03)	(3.73)
Regulatory Quality	0.169***	-0.374	-0.00970	-0.00970	-0.00970
	(3.48)	(-0.73)	(-0.06)	(-0.17)	(-0.15)
Rule of Law	-0.145*	-0.323	0.0208	0.0208	0.0208
	(-2.41)	(-1.03)	(0.07)	(0.34)	(0.29)
Voice of Accountability	0.0758	-0.983*	-0.0953	-0.0953	-0.0953
	(1.68)	(-2.07)	(-0.40)	(-1.77)	(-1.31)
gfcfgrowth	0.0791***	0.0452	0.0196	0.0196	0.0196
	(6.95)	(0.79)	(0.71)	(1.18)	(0.94)
humancap	0.204**	1.689	1.725	1.725***	1.725**
	(2.71)	(0.90)	(1.92)	(4.80)	(3.28)
$y_{I,t-1}$	-0.0552***	-0.610**	-0.545***	-0.545***	-0.545***
	(-5.40)	(-2.71)	(-6.63)	(-15.71)	(-12.61)
Constant	0.737***	-2.875	-7.60e-14	-7.60e-14	-7.60e-14
	(5.08)	(-0.60)	(-0.00)	(-0.00)	(-0.00)
NT	1153	1134	1134	1134	1134

t statistics in parentheses

CD = 56.57558705407147, p = 0

* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Table 6.31: TCCEMG(MVE) - Developing Countries

	FE	CCEMG	CCEP	CEEPnw	CCEPwpn
Control of Corruption	0.0192	0.342	-0.147	-0.147	-0.147
	(0.21)	(0.56)	(-0.15)	(-1.32)	(-1.01)
goveff	-0.112	-0.517	-0.256	-0.256	-0.256
	(-0.90)	(-1.08)	(-0.60)	(-1.87)	(-1.64)
Political Stability	0.100*	-0.258	0.105	0.105	0.105
	(2.24)	(-0.42)	(0.22)	(1.70)	(1.50)
Regulatory Quality	0.0313	-1.029	0.312	0.312**	0.312***
	(0.33)	(-1.22)	(0.33)	(2.62)	(3.75)
Rule of Law	-0.109	0.0550	0.0778	0.0778	0.0778
	(-1.09)	(0.12)	(0.18)	(0.47)	(0.44)
Voice of Accountability	-0.125	1.871	0.203	0.203	0.203
	(-1.52)	(1.77)	(0.19)	(1.50)	(1.12)
gfcfgrowth	-0.0746***	-0.235	-0.0267	-0.0267	-0.0267
	(-3.86)	(-0.98)	(-0.20)	(-0.80)	(-0.61)
humancap	0.538***	2.200	0.810	0.810*	0.810
	(5.79)	(0.82)	(0.31)	(2.18)	(1.57)
$y_{I,t-1}$	-0.154***	-1.048	-0.635*	-0.635***	-0.635***
	(-6.67)	(-2.09)	(-2.57)	(-14.73)	(-18.19)
Constant	2.029***	41.87	-4.77e-14	-4.77e-14	-4.77e-14
	(6.85)	(1.24)	(-0.00)	(-0.00)	(-0.00)
NT	418	399	399	399	399

t statistics in parentheses

CD = 2.563417535747073, p = .0103647319388731

* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

6.A.3 Global Outliers for Institutions

Table 6.32: Outliers (Institutions) - World

Country	GDP Growth	MD Outlier	MCD Outlier	MVE Outlier
Angola	0.324	0	1	0
Albania	0.113	0	1	0
United Arab Emirates	0.042	0	1	1
Argentina	0.137	1	1	1
Azerbaijan	0.184	1	1	1
Burundi	0.060	0	1	1
Benin	0.032	0	1	0
Bangladesh	0.130	0	1	0
Bahrain	0.098	1	1	1
Belize	0.118	1	1	1
Botswana	0.113	0	1	1
Central African Republic	0.044	0	0	1
Chile	0.144	0	1	1
Côte d'Ivoire	0.132	0	1	0
D.R. of the Congo	0.114	0	1	1
Colombia	0.096	0	1	0
Czech Republic	0.015	0	1	1
Denmark	0.065	0	1	1
Algeria	0.145	0	1	0
Ecuador	0.087	0	1	0
Egypt	0.136	0	1	0
Finland	0.034	1	1	1
Gabon	0.161	0	1	1
United Kingdom	0.067	0	1	1
Gambia	0.041	0	1	0
Guatemala	0.053	0	1	1
India	0.145	0	1	1
Iran (Islamic Republic of)	0.069	0	1	1
Italy	0.012	0	1	0

Jordan	0.058	0	1	0
Japan	0.044	0	1	0
Kenya	0.100	1	1	1
Cambodia	0.057	0	1	1
Kuwait	0.122	0	1	0
Lao People's DR	0.062	0	1	1
Luxembourg	0.135	1	1	1
Latvia	0.032	1	1	1
Morocco	0.065	1	1	1
Republic of Moldova	0.126	0	1	1
Mexico	0.086	0	1	0
Myanmar	0.181	1	1	1
Mongolia	0.177	0	1	1
Mauritania	0.193	0	1	0
Malaysia	0.084	0	1	1
Namibia	0.077	0	1	1
Nepal	0.105	0	1	0
Peru	0.129	1	1	1
Portugal	0.041	0	1	0
Qatar	0.227	1	1	1
Romania	0.018	1	1	1
Russian Federation	0.127	0	1	0
Saudi Arabia	0.204	1	1	1
Sudan	0.143	0	1	1
El Salvador	0.151	0	1	0
Sao Tome and Principe	0.101	0	1	0
Eswatini	0.050	1	1	1
Togo	0.062	1	1	1
Tajikistan	0.176	0	1	1
Trinidad and Tobago	0.144	1	1	1
Uganda	0.093	0	1	1
Ukraine	0.033	0	1	0

Viet Nam	0.128	1	1	1
South Africa	0.041	0	1	0
Zambia	0.125	0	0	1
Zimbabwe	0.091	1	1	1

1 = flagged as outlier; 0 otherwise

6.B Women's Studies

6.B.1 Country List for Women's Studies

Table 6.33: Country List (and Mean Values by Country)

Country	Growth Rate	Fixed Capital Growth	Female HC	Male HC	Pay Gap Ratio
Albania	0.039	0.074	0.738	0.765	1.124
Argentina	0.030	0.001	0.818	0.823	1.123
Armenia	0.067	0.058	0.720	0.741	1.207
Australia	0.032	0.018	0.906	0.937	1.091
Austria	0.025	0.028	0.881	0.915	1.095
Azerbaijan	0.103	0.071	0.694	0.722	1.141
Burundi	0.031	0.031	0.366	0.404	1.092
Belgium	0.023	0.039	0.898	0.925	1.083
Benin	0.065	0.063	0.444	0.533	1.102
Bangladesh	0.070	0.084	0.507	0.598	1.491
Bulgaria	0.036	0.052	0.777	0.786	1.092
Belarus	0.041	0.050	0.786	0.773	1.091
Belize	0.022	0.036	0.667	0.702	1.206
Brazil	0.032	0.029	0.724	0.732	1.124
Barbados	-0.031	0.002	0.786	0.773	1.078
Botswana	0.040	0.061	0.644	0.667	1.081
Central African Republic	0.012	0.035	0.320	0.413	1.201
Canada	0.022	0.021	0.906	0.921	1.070
Switzerland	0.033	0.034	0.910	0.947	1.057
Chile	0.047	0.054	0.796	0.832	1.147
China	0.070	0.096	0.670	0.703	1.104
Côte d'Ivoire	0.061	0.121	0.441	0.527	1.249
Cameroon	0.043	0.060	0.478	0.558	1.117
Congo	0.069	0.083	0.524	0.575	1.073
Colombia	0.043	0.050	0.718	0.726	1.104
Comoros	0.025	-0.006	0.478	0.552	1.256

Cabo Verde	0.047	0.016	0.625	0.644	1.167
Costa Rica	0.043	0.035	0.758	0.780	1.126
Cyprus	0.026	0.030	0.840	0.870	1.083
Czech Republic	0.029	0.036	0.853	0.874	1.098
Germany	0.018	0.015	0.907	0.942	1.076
Denmark	0.025	0.037	0.913	0.928	1.064
Algeria	0.031	0.035	0.629	0.750	1.531
Egypt	0.061	0.046	0.618	0.713	1.411
Spain	0.028	0.025	0.858	0.877	1.096
Estonia	0.050	0.075	0.868	0.843	1.084
Ethiopia	0.102	0.169	0.373	0.440	1.190
Finland	0.017	0.017	0.914	0.918	1.061
Fiji	0.042	0.051	0.686	0.735	1.222
France	0.019	0.032	0.871	0.883	1.069
Gabon	0.053	0.081	0.635	0.703	1.142
United Kingdom	0.021	0.027	0.892	0.916	1.092
Georgia	0.070	0.084	0.747	0.768	1.203
Ghana	0.053	0.052	0.542	0.602	1.123
Guinea	0.009	0.013	0.371	0.457	1.052
Gambia	0.024	0.024	0.423	0.490	1.215
Greece	0.003	-0.020	0.846	0.876	1.101
Guatemala	0.041	0.044	0.574	0.628	1.222
Guyana	0.055	0.057	0.634	0.665	1.236
Honduras	0.045	0.040	0.581	0.604	1.169
Croatia	0.033	0.052	0.812	0.829	1.081
Hungary	0.029	0.050	0.818	0.827	1.087
Indonesia	0.067	0.109	0.631	0.687	1.190
India	0.072	0.086	0.507	0.623	1.494
Ireland	0.066	0.105	0.891	0.909	1.084
Iran	0.040	0.046	0.673	0.778	1.485
Iraq	0.049	-0.007	0.541	0.685	1.717
Italy	0.011	0.011	0.859	0.889	1.117

Jamaica	0.024	0.034	0.693	0.703	1.107
Jordan	0.090	0.081	0.651	0.747	1.436
Japan	0.003	-0.008	0.883	0.917	1.121
Kazakhstan	0.093	0.103	0.767	0.764	1.111
Kenya	0.058	0.078	0.516	0.565	1.087
Kyrgyzstan	0.062	0.072	0.652	0.673	1.205
Cambodia	0.077	0.092	0.506	0.564	1.087
Republic of Korea	0.034	0.037	0.849	0.910	1.136
Kuwait	0.058	0.069	0.806	0.807	1.126
Laos	0.083	0.078	0.526	0.578	1.068
Liberia	0.062	0.036	0.414	0.503	1.165
Lesotho	0.008	0.023	0.476	0.482	1.135
Morocco	0.033	0.049	0.544	0.657	1.395
Republic of Moldova	0.066	0.066	0.724	0.723	1.081
Madagascar	0.038	0.063	0.468	0.503	1.112
Maldives	0.075	0.126	0.664	0.717	1.219
Mexico	0.028	0.027	0.726	0.765	1.169
Mali	0.060	0.079	0.356	0.423	1.231
Mongolia	0.086	0.110	0.703	0.683	1.073
Mozambique	0.046	0.059	0.373	0.418	1.074
Mauritania	0.046	0.077	0.471	0.547	1.346
Mauritius	0.021	0.004	0.731	0.774	1.186
Malawi	0.035	0.039	0.432	0.470	1.158
Malaysia	0.056	0.048	0.749	0.781	1.141
Namibia	0.039	0.016	0.585	0.596	1.088
Niger	0.042	0.090	0.298	0.379	1.248
Nicaragua	0.025	-0.021	0.603	0.634	1.259
Netherlands	0.021	0.028	0.899	0.935	1.074
Norway	0.033	0.044	0.935	0.943	1.046
Nepal	0.055	0.088	0.513	0.572	1.079
New Zealand	0.028	0.030	0.904	0.930	1.081
Pakistan	0.050	0.044	0.410	0.563	1.745

Panama	0.074	0.092	0.769	0.775	1.118
Peru	0.054	0.051	0.708	0.749	1.110
Philippines	0.049	0.057	0.661	0.688	1.146
Poland	0.038	0.032	0.842	0.840	1.100
Portugal	0.015	0.004	0.825	0.835	1.076
Paraguay	0.055	0.047	0.676	0.702	1.163
Romania	0.053	0.085	0.781	0.796	1.088
Rwanda	0.070	0.113	0.451	0.488	1.110
Saudi Arabia	0.074	0.073	0.749	0.848	1.314
Senegal	0.032	0.051	0.424	0.499	1.291
Sierra Leone	0.066	0.073	0.386	0.453	1.105
El Salvador	0.072	0.097	0.638	0.665	1.149
Serbia	0.036	0.061	0.747	0.778	1.118
Sao Tome and Principe	0.055	0.048	0.525	0.592	1.296
Slovakia	0.028	0.045	0.822	0.830	1.098
Slovenia	0.022	0.017	0.883	0.886	1.067
Sweden	0.023	0.030	0.914	0.923	1.053
Eswatini	0.005	-0.006	0.507	0.535	1.089
Syrian Arab Republic	0.084	0.072	0.535	0.643	1.621
Togo	0.042	0.051	0.433	0.538	1.125
Thailand	0.046	0.040	0.738	0.742	1.082
Tajikistan	0.078	0.110	0.589	0.667	1.296
Trinidad and Tobago	0.048	0.022	0.772	0.789	1.116
Tunisia	0.020	-0.005	0.665	0.739	1.332
Turkey	0.054	0.071	0.711	0.793	1.221
Tanzania	0.065	0.081	0.465	0.506	1.112
Uganda	0.059	0.053	0.459	0.508	1.142
Ukraine	0.044	0.041	0.762	0.751	1.100
Uruguay	0.028	0.020	0.790	0.779	1.117
United States	0.021	0.013	0.905	0.915	1.078
Uzbekistan	0.065	0.096	0.650	0.691	1.176
Venezuela	-0.159	-0.385	0.731	0.729	1.158

Vietnam	0.072	0.090	0.654	0.663	1.079
Yemen	0.046	0.142	0.336	0.553	2.967
South Africa	0.028	0.033	0.660	0.689	1.123
Zimbabwe	-0.012	-0.040	0.498	0.541	1.108

Mean Values by Country

6.B.2 Full Results for Women's Studies

Untreated Results

Table 6.34: Estimation Results - World

	FE	CCEMG	CCEP	CEEPnw	CCEPwpn
Gender Pay Gap	-0.0741***	-0.493	-0.493	-0.192**	-0.192***
	(-5.55)	(-1.55)	(-1.55)	(-3.00)	(-4.70)
Female Human Capital	-0.262	-0.0000328	-0.0000328	-0.692	-0.692*
	(-1.28)	(-0.00)	(-0.00)	(-1.78)	(-2.04)
Male Human Capital	0.138	0.929	0.929	1.839***	1.839***
	(0.56)	(0.82)	(0.82)	(4.55)	(5.40)
Fixed Capital Growth	0.129***	0.141***	0.141***	0.0757***	0.0757***
	(20.63)	(10.39)	(10.39)	(5.86)	(4.33)
Constant	0.200***	0.913	0.913	0.117	0.117
	(4.26)	(1.82)	(1.82)	(0.32)	(0.32)
Observations	2513	2513	2513	2513	2513

t statistics in parentheses

CD = 54.659, p = 0.000,

* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Table 6.35: Estimation Results - OECD

	FE	CCEMG	CCEP	CEEPnw	CCEPwpn
Gender Pay Gap	0.187	0.0272	0.0272	0.0824	0.0824
	(1.71)	(0.06)	(0.06)	(0.72)	(0.83)
Female Human Capital	0.799**	2.372*	2.372*	1.066**	1.066**
	(2.86)	(2.08)	(2.08)	(2.78)	(3.02)
Male Human Capital	-0.925***	-1.299	-1.299	-0.374	-0.374
	(-3.34)	(-0.84)	(-0.84)	(-1.00)	(-1.14)
Fixed Capital Growth	0.272***	0.219***	0.219***	0.208***	0.208***
	(22.70)	(8.76)	(8.76)	(12.05)	(15.43)
Constant	-0.0569	0.340	0.340	-8.73e-14	-8.73e-14
	(-0.35)	(0.19)	(0.19)	(-0.00)	(-0.00)
Observations	660	660	660	660	660

t statistics in parentheses

CD = 18.427, p = 0.000

* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Table 6.36: Estimation Results - Emerging Markets

	FE	CCEMG	CCEP	CEEPnw	CCEPwpn
Gender Pay Gap	-0.0771***	-0.245	-0.245	-0.249**	-0.249***
	(-5.33)	(-1.01)	(-1.01)	(-2.98)	(-4.60)
Female Human Capital	-0.331	1.182	1.182	-0.401	-0.401
	(-1.29)	(1.10)	(1.10)	(-0.92)	(-1.11)
Male Human Capital	0.157	0.313	0.313	1.547***	1.547***
	(0.51)	(0.31)	(0.31)	(3.61)	(4.37)
Fixed Capital Growth	0.151***	0.143***	0.143***	0.0993***	0.0993***
	(19.77)	(7.94)	(7.94)	(8.48)	(6.53)
Constant	0.242***	0.667	0.667	0.131	0.131
	(3.96)	(0.93)	(0.93)	(0.23)	(0.23)
Observations	1453	1453	1453	1453	1453

t statistics in parentheses

CD = 39.743, p = 0.000

* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Table 6.37: Estimation Results - Developing Countries

	FE	CCEMG	CCEP	CEEPnw	CCEPwpn
Gender Pay Gap	-0.0935	-0.610	-0.610	0.0786	0.0786
	(-0.86)	(-0.74)	(-0.74)	(0.62)	(0.53)
Female Human Capital	-0.389	-1.325	-1.325	1.005	1.005
	(-0.59)	(-0.48)	(-0.48)	(0.94)	(0.90)
Male Human Capital	0.447	2.993	2.993	0.542	0.542
	(0.57)	(1.08)	(1.08)	(0.49)	(0.45)
Fixed Capital Growth	0.0291	0.106**	0.106**	0.0238	0.0238
	(1.87)	(3.30)	(3.30)	(0.90)	(0.61)
Constant	0.103	1.508	1.508	-1.57e-15	-1.57e-15
	(0.76)	(1.63)	(1.63)	(-0.00)	(-0.00)
Observations	540	540	540	540	540

t statistics in parentheses

CD = 2.084, p = 0.037

* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

MD-Trimmed Results

Table 6.38: TCCEMG(Lee and Sul, 2020) - World

	FE	CCEMG	CCEP	CEEPnw	CCEPwpn
Gender Pay Gap	-0.0692***	-0.530	-0.203	-0.203**	-0.203***
	(-5.26)	(-1.44)	(-0.79)	(-2.89)	(-4.90)
Female Human Capital	-0.0840	0.00417	-0.583	-0.583	-0.583
	(-0.39)	(0.00)	(-0.46)	(-1.44)	(-1.56)
Male Human Capital	-0.106	0.920	2.045	2.045***	2.045***
	(-0.41)	(0.84)	(1.63)	(4.58)	(5.51)
Fixed Capital Growth	0.132***	0.139***	0.0689*	0.0689***	0.0689***
	(20.29)	(9.64)	(2.07)	(4.83)	(3.54)
Constant	0.249***	1.081*	0.126	0.126	0.126
	(5.12)	(2.15)	(0.36)	(0.36)	(0.36)
N	2433	2440	2440	2440	2440

t statistics in parentheses

* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Table 6.39: TCCEMG(Lee and Sul, 2020) - OECD

	FE	CCEMG	CCEP	CEEPnw	CCEPwpn
Gender Pay Gap	0.198	-0.0150	0.0614	0.0614	0.0614
	(1.76)	(-0.03)	(0.29)	(0.50)	(0.49)
Female Human Capital	0.825**	2.446*	0.749	0.749*	0.749
	(2.83)	(1.99)	(1.16)	(2.02)	(1.82)
Male Human Capital	-0.950**	-1.296	-0.380	-0.380	-0.380
	(-3.28)	(-0.78)	(-0.45)	(-1.00)	(-0.98)
Fixed Capital Growth	0.271***	0.214***	0.215***	0.215***	0.215***
	(22.11)	(8.33)	(13.88)	(12.08)	(15.41)
Constant	-0.0694	0.420	4.32e-14	4.32e-14	4.32e-14
	(-0.42)	(0.25)	(0.00)	(0.00)	(0.00)
N	640	640	640	640	640

t statistics in parentheses

* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Table 6.40: TCCEMG(Lee and Sul, 2020) - Emerging Markets

	FE	CCEMG	CCEP	CEEPnw	CCEPwpn
Gender Pay Gap	-0.0785***	-0.200	-0.291**	-0.291***	-0.291***
	(-5.54)	(-0.78)	(-2.62)	(-3.68)	(-4.95)
Female Human Capital	-0.368	1.172	-0.704	-0.704	-0.704
	(-1.45)	(0.99)	(-0.99)	(-1.52)	(-1.56)
Male Human Capital	0.155	0.318	1.498*	1.498***	1.498***
	(0.51)	(0.29)	(2.47)	(3.62)	(4.45)
Fixed Capital Growth	0.163***	0.136***	0.0953**	0.0953***	0.0953***
	(20.04)	(7.05)	(2.93)	(6.80)	(5.04)
Constant	0.269***	0.803	0.155	0.155	0.155
	(4.34)	(1.14)	(0.30)	(0.30)	(0.30)
N	1353	1360	1360	1360	1360

t statistics in parentheses

* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Table 6.41: TCCEMG(Lee and Sul, 2020) - Developing Countries

	FE	CCEMG	CCEP	CEEPnw	CCEPwpn
Gender Pay Gap	-0.0848	-0.771	0.0451	0.0451	0.0451
	(-0.76)	(-0.86)	(0.11)	(0.37)	(0.32)
Female Human Capital	-0.324	-2.263	0.582	0.582	0.582
	(-0.47)	(-0.76)	(0.36)	(0.57)	(0.56)
Male Human Capital	0.366	3.461	0.690	0.690	0.690
	(0.45)	(1.23)	(0.41)	(0.65)	(0.65)
Fixed Capital Growth	0.0281	0.104**	0.0188	0.0188	0.0188
	(1.74)	(3.03)	(0.27)	(0.67)	(0.47)
Constant	0.105	1.918	7.21e-16	7.21e-16	7.21e-16
	(0.74)	(1.78)	(0.00)	(0.00)	(0.00)
N	520	520	520	520	520

t statistics in parentheses

* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

MCD-Trimmed Results

Table 6.42: TCCEMG(MCD) - World

	FE	CCEMG	CCEP	CEEPnw	CCEPwpn
Gender Pay Gap	-0.0669***	-0.344	-0.350***	-0.350***	-0.350***
	(-4.78)	(-0.83)	(-7.46)	(-4.77)	(-7.25)
Female Human Capital	0.0948	0.416	-0.920	-0.920	-0.920*
	(0.36)	(0.30)	(-1.62)	(-1.92)	(-2.04)
Male Human Capital	-0.326	0.569	2.146***	2.146***	2.146***
	(-1.05)	(0.43)	(3.42)	(4.64)	(6.83)
Fixed Capital Growth	0.133***	0.142***	0.0632	0.0632**	0.0632*
	(17.44)	(8.87)	(1.60)	(3.30)	(2.42)
Constant	0.286***	0.836	0.125	0.125	0.125
	(4.85)	(1.31)	(0.27)	(0.27)	(0.27)
N	1775	1780	1780	1780	1780

t statistics in parentheses

TCCEMG, MCD

* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Table 6.43: TCCEMG(MCD) - OECD

	FE	CCEMG	CCEP	CEEPnw	CCEPwpn
Gender Pay Gap	0.213	0.709	0.831*	0.831***	0.831***
	(1.77)	(1.31)	(2.13)	(4.82)	(5.07)
Female Human Capital	1.147***	2.973	3.303**	3.303***	3.303***
	(3.72)	(1.80)	(2.67)	(8.77)	(7.39)
Male Human Capital	-1.338***	-1.948	-2.561	-2.561***	-2.561***
	(-4.29)	(-1.05)	(-1.82)	(-6.73)	(-6.45)
Fixed Capital Growth	0.260***	0.214***	0.218***	0.218***	0.218***
	(20.81)	(6.86)	(12.92)	(10.98)	(18.52)
Constant	-0.0202	1.186	-8.36e-14	-8.36e-14	-8.36e-14
	(-0.11)	(0.53)	(-0.00)	(-0.00)	(-0.00)
N	420	420	420	420	420

t statistics in parentheses
TCCEMG, MCD
* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Table 6.44: TCCEMG(MCD) - Emerging Markets

	FE	CCEMG	CCEP	CEEPnw	CCEPwpn
Gender Pay Gap	-0.0767***	-0.384	-0.208	-0.208**	-0.208***
	(-4.90)	(-1.51)	(-1.34)	(-2.85)	(-4.56)
Female Human Capital	-0.223	0.728	-0.0178	-0.0178	-0.0178
	(-0.72)	(0.59)	(-0.02)	(-0.03)	(-0.03)
Male Human Capital	-0.0304	1.020	0.882	0.882*	0.882*
	(-0.08)	(0.83)	(0.91)	(2.00)	(2.36)
Fixed Capital Growth	0.163***	0.133***	0.0942**	0.0942***	0.0942***
	(17.67)	(6.31)	(2.90)	(5.95)	(4.47)
Constant	0.301***	0.567	0.119	0.119	0.119
	(4.21)	(0.62)	(0.18)	(0.18)	(0.18)
N	1033	1040	1040	1040	1040

t statistics in parentheses
TCCEMG, MCD
* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Table 6.45: TCCEMG(MCD) - Developing Countries

	FE	CCEMG	CCEP	CEEPnw	CCEPwpn
Gender Pay Gap	-0.155	-0.958	0.0538	0.0538	0.0538
	(-1.28)	(-0.94)	(0.11)	(0.43)	(0.37)
Female Human Capital	-1.167	-3.485	0.898	0.898	0.898
	(-1.43)	(-1.02)	(0.43)	(0.81)	(0.76)
Male Human Capital	1.377	4.593	0.818	0.818	0.818
	(1.43)	(1.40)	(0.36)	(0.70)	(0.60)
Fixed Capital Growth	0.0271	0.108**	0.0208	0.0208	0.0208
	(1.56)	(2.64)	(0.30)	(0.76)	(0.53)
Constant	0.0454	1.943	-1.38e-14	-1.38e-14	-1.38e-14
	(0.29)	(1.48)	(-0.00)	(-0.00)	(-0.00)
N	460	460	460	460	460

t statistics in parentheses

TCCEMG, MCD

* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

MVE-Trimmed Results

Table 6.46: TCCEMG(MVE) - World

	FE	CCEMG	CCEP	CEEPnw	CCEPwpn
Gender Pay Gap	-0.0662***	-0.459	-0.292**	-0.292***	-0.292***
	(-4.84)	(-1.20)	(-2.64)	(-3.71)	(-6.95)
Female Human Capital	0.0911	0.452	-0.587	-0.587	-0.587
	(0.37)	(0.35)	(-0.85)	(-1.33)	(-1.46)
Male Human Capital	-0.310	0.582	1.857**	1.857***	1.857***
	(-1.04)	(0.45)	(2.59)	(4.20)	(5.95)
Fixed Capital Growth	0.135***	0.141***	0.0665	0.0665***	0.0665**
	(18.37)	(8.81)	(1.74)	(3.66)	(2.69)
Constant	0.276***	0.737	0.114	0.114	0.114
	(4.96)	(1.24)	(0.29)	(0.29)	(0.29)
N	1895	1900	1900	1900	1900

t statistics in parentheses

* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Table 6.47: TCCEMG(MVE) - OECD

	FE	CCEMG	CCEP	CEEPnw	CCEPwpn
Gender Pay Gap	0.207	-0.410	-0.0127	-0.0127	-0.0127
	(1.70)	(-0.50)	(-0.05)	(-0.10)	(-0.10)
Female Human Capital	1.062**	1.686	0.982	0.982*	0.982*
	(3.20)	(1.32)	(1.36)	(2.46)	(2.26)
Male Human Capital	-1.225***	-0.641	-0.797	-0.797	-0.797*
	(-3.64)	(-0.35)	(-0.81)	(-1.96)	(-2.00)
Fixed Capital Growth	0.271***	0.207***	0.216***	0.216***	0.216***
	(20.38)	(6.59)	(12.09)	(10.13)	(13.21)
Constant	-0.0412	0.614	-1.69e-14	-1.69e-14	-1.69e-14
	(-0.23)	(0.30)	(-0.00)	(-0.00)	(-0.00)
N	560	560	560	560	560

t statistics in parentheses

CD = 18.427, p = 0.000

* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Table 6.48: TCCEMG(MVE) - Emerging Markets

	FE	CCEMG	CCEP	CEEPnw	CCEPwpn
Gender Pay Gap	-0.0768***	-0.378	-0.266**	-0.266***	-0.266***
	(-4.97)	(-1.52)	(-2.69)	(-3.82)	(-6.39)
Female Human Capital	-0.250	0.396	-0.348	-0.348	-0.348
	(-0.82)	(0.33)	(-0.39)	(-0.67)	(-0.64)
Male Human Capital	-0.00813	1.275	1.128	1.128**	1.128**
	(-0.02)	(1.09)	(1.41)	(2.59)	(3.26)
Fixed Capital Growth	0.162***	0.132***	0.0947**	0.0947***	0.0947***
	(17.96)	(5.97)	(3.15)	(6.15)	(4.70)
Constant	0.303***	0.692	0.127	0.127	0.127
	(4.36)	(0.90)	(0.22)	(0.22)	(0.22)
N	1093	1100	1100	1100	1100

t statistics in parentheses

* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

CD = 39.743, p = 0.000

Table 6.49: TCCEMG(MVE) - Developing Countries

	FE	CCEMG	CCEP	CEEPnw	CCEPwpn
Gender Pay Gap	-0.0781	-0.604	0.226	0.226	0.226
	(-0.60)	(-0.74)	(0.60)	(1.38)	(1.05)
Female Human Capital	-0.418	-1.634	1.598	1.598	1.598
	(-0.57)	(-0.63)	(0.74)	(1.25)	(1.15)
Male Human Capital	0.490	3.681	1.324	1.324	1.324
	(0.57)	(1.28)	(0.69)	(1.05)	(0.96)
Fixed Capital Growth	0.0267	0.105**	0.0125	0.0125	0.0125
	(1.60)	(2.86)	(0.19)	(0.46)	(0.32)
Constant	0.0766	1.117	-4.49e-15	-4.49e-15	-4.49e-15
	(0.46)	(1.35)	(-0.00)	(-0.00)	(-0.00)
N	480	480	480	480	480

t statistics in parentheses

* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

CD = 2.084, p = 0.037

6.B.3 Global Outliers for Women's Studies

Table 6.50: Outliers (Women's Studies) - World

Country	GDP Growth	Human Capital (Female)	MD	MCD	MVE
Argentina	0.137	0.829	0	1	1
Burundi	0.060	0.385	1	1	1
Belgium	0.080	0.898	0	1	1
Belarus	0.100	0.801	0	1	0
Belize	0.118	0.686	0	1	1
Brazil	0.149	0.718	0	1	0
Central African Republic	0.044	0.323	0	1	1
Switzerland	0.036	0.920	0	1	1
Chile	0.144	0.794	0	1	0
Côte d'Ivoire	0.132	0.426	0	1	1
Congo	0.460	0.536	0	1	0
Cabo Verde	0.076	0.634	0	1	1
Denmark	0.065	0.907	0	1	1
Egypt	0.136	0.613	0	1	1
Finland	0.034	0.913	1	1	1
Fiji	0.102	0.687	0	1	0

Georgia	0.045	0.740	0	1	1
Ghana	0.112	0.543	0	1	1
Gambia	0.041	0.424	0	1	0
Honduras	0.058	0.588	0	1	1
Indonesia	0.171	0.632	0	1	1
Iran (Islamic Republic of)	0.069	0.678	0	1	1
Iraq	0.175	0.544	0	1	1
Japan	0.044	0.875	0	1	1
Kazakhstan	0.174	0.764	0	1	1
Kyrgyzstan	0.003	0.652	0	1	0
Republic of Korea	0.070	0.852	0	1	0
Kuwait	0.122	0.803	0	1	1
Morocco	0.065	0.538	0	1	1
Republic of Moldova	0.126	0.731	0	1	0
Mali	0.053	0.375	0	1	1
Mongolia	0.177	0.711	0	1	1
Mozambique	0.022	0.379	0	1	1
Mauritius	0.025	0.728	0	1	1
Nicaragua	0.036	0.594	1	1	1
Norway	0.082	0.937	1	1	1
Nepal	0.105	0.516	0	1	0
Romania	0.018	0.801	1	1	1
Saudi Arabia	0.204	0.742	0	1	1
Sierra Leone	0.043	0.396	0	1	1
Syrian Arab Republic	0.088	0.573	0	1	0
Togo	0.062	0.417	1	1	1
Tajikistan	0.176	0.590	0	1	0
Tunisia	0.060	0.674	0	1	1
Turkey	0.139	0.698	0	1	1
U.R. of Tanzania: Mainland	0.139	0.473	0	1	1
Uganda	0.093	0.479	0	1	0
Ukraine	0.033	0.769	0	1	0

Zimbabwe	0.091	0.492	1	1	1

1 = flagged as outlier by above diagnostic; 0 otherwise

Chapter 7

A Robust CCE Estimator

7.1 Introduction

In the previous chapters, we have looked at addressing outliers in cross-section units, such as countries that do not fit a model. Often times, outlying observations occur within non-outlying units. In practice, outliers in country macro- and micro-data are the norm and not the exception. It could be that these outlying observations are randomly interspersed among the bulk of the data in the panel. While large outliers within an individual unit can potentially cause that unit to be flagged as a block outlier, it may be the case that the majority of units (especially if more than half the units) have at least one such outlier. In this case, simply detecting outliers at the unit level is insufficient. In this paper, we are considering a robust approach to the CCE estimator, by augmenting the CCE estimator with methods in the robust literature. We seek to replace least squares in the CCE estimators with robust analogues in the literature. In addition, we will also consider a robust alternative to the cross-section averages.

Motivation

The factor model of Pesaran (2006) is given by

$$\begin{aligned} y_{it} &= \boldsymbol{\alpha}_i'\mathbf{d}_t + \boldsymbol{\beta}_i'\mathbf{x}_{it} + \boldsymbol{\gamma}_i'\mathbf{f}_t + \epsilon_{it} \\ \mathbf{x}_{it} &= \mathbf{A}_i'\mathbf{d}_t + \boldsymbol{\Gamma}_i'\mathbf{f}_t + \mathbf{v}_{it} \\ e_{it} &= \boldsymbol{\gamma}_i'\mathbf{f}_t + \epsilon_{it} \end{aligned} \tag{7.1}$$

of which the CCE estimates are least squares (LS) estimates of

$$y_{it} = \boldsymbol{\alpha}_i'\mathbf{d}_t + \boldsymbol{\beta}_i'\mathbf{x}_{it} + \boldsymbol{\delta}_i'\bar{\mathbf{z}}_{wt} + e_{it} \tag{7.2}$$

where $\bar{\mathbf{z}}_{wt} = (\bar{y}_{wt}, \bar{\mathbf{x}}'_{wt})'$ are the cross-section averages with weights w_i satisfying assumption 5 of Pesaran (2006) (see also Lee and Sul, 2022; Ditzen, 2018). In a private email exchange with Markus Eberhardt of the University of Nottingham. Regarding CCE estimation, Eberhardt remarked that estimating the cross-sections with a robust estimator instead of least squares causes the Mean Group Estimator to become robust to outliers (for reference, see Eberhardt, 2012). Eberhardt's works include, among others, the *xtmg* software package in Stata (Eberhardt, 2012). The *xtmg* package estimates the mean group (MG) estimator (Pesaran and Smith, 1995) and the CCE estimator of Pesaran (2006). The *xttmg* command allows for robust estimation via the *rreg* command in Stata. The *rreg* command uses the Cook's distance and iterations of the Huber-weighted M-estimators (see Huber, 1964; Rousseeuw and Leroy, 1987 for details) to produce weights for the CCE and MG estimators for outlier robustness. To our knowledge, we have not found anything else in the literature examining (outlier-) robust CCE estimators, apart from Lee and Sul (2022) and Eberhardt (2012). However, the robust literature is large, among them influential works by Peter Rousseeuw (Rousseeeuw, 1984; Rousseeuw and Leroy, 1987; Hubert, Debruyne and Rousseeuw, 2018, to name a few) and Maronna et. al. (2019). The Stata literature for robust estimation includes software packages for: S-estimators (Verardi and McCathie, 2012); Minimum Cavariance Determinant (MCD) and Minimum Volume Ellipsoid (MVE) estimators (Verardi and Dehon, 2010), robust regression estimators (Jann, 2010), Verardi and Croux (2009), and many more. However, the *rreg* command does not have the desired robustness properties for two main reasons (Verardi and Croux, 2009). Firstly, Cook distances are only effective at identifying isolated outliers, but suffer from masking and swamping when clusters of outliers exist (Rousseeuw and van Zomeren, 1990). Therefore, it does not necessarily identify all leverage points effectively. Secondly, the initial values for the iteratively reweighted OLS algorithm are monotone M-estimators that are not robust to bad leverage points, which may lead the algorithm to converge to a local instead of a global minimum (Verardi and Croux, 2009). In this paper, we will examine the effects of outliers on the CCE augmented with robust estimators and develop robust estimators for the CCE. Our work closely resembles that of Maria Caterina Bramati and Christopher Croux (Bramati and Croux, 2007). In their paper, Bramati and Croux developed a robust Fixed Effects estimator by replacing the time series averages within each unit with the medians, and estimating the transformed data with robust estimators. Our approach is similar: we will address the issue of outliers within the cross-section averages, and estimate the augmented models with a robust estimator.

7.2 Models and Estimation

We consider the effect of outliers in the estimation of (7.2). We look at two areas: in $\bar{\mathbf{z}}_{wt}$ and in the overall least squares estimation. The CCE Mean Group (CCEMG) estimator is a mean group estimation of the individual cross-section units. The Pooled CCE (CCEP) estimator is a pooled LS estimation of (7.2) (see also Ditzen, 2018). In the next section, we consider robust alternatives to least squares.

7.2.1 Least Squares and Robust Estimators

It is well-known in the literature that least squares is non-robust. A single arbitrarily large outlier can distort slope estimates (see Rousseeuw and Leroy, 1987 for examples). A measure of robustness in the literature is the breakdown point (see Rousseeuw and Leroy, 1987). The breakdown point (BP) is the maximum proportion of outlying observations an estimator can take before the estimator fails. It also implies that the maximum breakdown point of any estimator is 50%. Two 50% (high-) breakdown robust analogues to LS are the Least Median of Squares (LMS) and the Least-Trimmed Squares (LTS) estimators (Rousseeuw, 1984; Rousseeuw and Leroy, 1987). The LMS minimises the median of the squared residuals, while the LTS minimises the sum of the squared residuals over a subsample of fixed proportion h of the entire sample (if $h = n/2$, where n is the sample size, then the LTS has a breakdown point of 50%). However, the LMS and LTS suffer from low efficiencies. The MM estimator (Yohai, 1987; Maronna et.al., 2019) is a reweighted M-estimator using a robust scale estimate (or S-estimator). In our investigation, we will consider replacing LS in the CCEP estimator with the MM-estimator. The resulting standard errors for the slope estimates will be inherited from the MM-estimator.

7.2.2 Cross-section Averages

The cross-section averages $\bar{\mathbf{z}}_{wt}$ are simple weighted averages of the observation vectors $\mathbf{z}_{it} = (y_{it}, \mathbf{x}'_{it})'$. Again, it is well-known in the literature that the mean is non-robust (see Rousseeuw and Hubert , 2011 for an illustration). The robust estimate of the sample mean with the highest BP is the sample median. We will replace the cross-section averages in the CCE estimator with cross-section medians. In addition to medians, we will consider a second approach: to estimate the center of the regressors with the MCD estimator (Rousseeuw, 1984; Hubert, Debruyne, and Rousseeuw, 2018). The MCD estimator robustly estimates the center of the data cloud by minimising the determinant of the covariance matrix estimate over all subsamples of size $h \leq n$ (where n is the sample size). By estimating a subsample of the cross-section units, we obtain robust estimates for the cross-section means.

7.2.3 The Robust CCE estimators

The methods in current robust literature assume normality or symmetry in the distributions. There is some literature on robust statistics for asymmetric distributions, such as the modified box-plot for skewed data (Hubert and Van der Veeken, 2008) and the asymmetric M-estimator (Allende et. al., 2006). However, the application of these will be considered for future research. To this end, we will restrict our scope to elliptically symmetric, unimodal distributions:

Definition 7.2.1. A multivariate distribution is elliptically symmetric and unimodal if there exists a strictly decreasing real function g such that the density function f of the distribution is of the form

$$f(\mathbf{x}) = \frac{1}{\sqrt{|\mathbf{\Sigma}|}} g(d^2(\mathbf{x}, \boldsymbol{\mu}, \mathbf{\Sigma})) \tag{7.3}$$

The statistical distance d is given by

$$d(\mathbf{x}, \boldsymbol{\mu}, \mathbf{\Sigma}) = \sqrt{(\mathbf{x} - \boldsymbol{\mu})' \mathbf{\Sigma}^{-1} (\mathbf{x} - \boldsymbol{\mu})} \tag{7.4}$$

The definition extends to the univariate case. The multivariate normal distribution is an example of an elliptically symmetric and unimodal distribution. Regarding the scope of our discussion,

Assumption 7.2.1. *The data $\mathbf{y}_i$ and $\mathbf{x}_i$ are elliptically symmetric and unimodal across all $i = 1, 2, \ldots, N$.*

We note that asymmetry is a problem present in small samples, as when the dimensionality increases, the data are more and more concentrated in an outer shell of the distribution (Hubert and Van der Veeken, 2008). We will consider asymmetric distributions in future work. Henceforth, we propose two new estimators: the CCE Robust Mean Group (CCERG) and the CCE Robust Pooled (CCERP). The CCERG is set up as follows. First, we compute robust cross-section location estimates as an analogue to cross-section averages. We consider two multivariate robust estimators: the marginal median and the MCD location estimate. The marginal median of a sample of i.i.d. sample of n $p \times 1$ vectors $\mathbf{x}_i$ is the pointwise median of the components of the vectors:

$$\operatorname*{med}_{1 \le i \le n} (\mathbf{x}) = (\operatorname*{med}_{1 \le i \le n} x_1, \operatorname*{med}_{1 \le i \le n} x_2, \ldots, \operatorname*{med}_{1 \le i \le n} x_p)' \tag{7.5}$$

The MCD can estimate a robust center (along with the corresponding covariance matrix) $\hat{\boldsymbol{\mu}}_i^{MCD}(\mathbf{x}) = \hat{\boldsymbol{\mu}}_i^{MCD}(\mathbf{x}_1, \mathbf{x}_2, \ldots, \mathbf{x}_n)$ across all observations $\mathbf{x}_i, i = 1, 2, \ldots, n$ in a sample. Again, this requires $\mathbf{x}_{it}$ to be elliptically symmetric across all cross-sections. For each $t = 1, 2, \ldots, T$, for every $i = 1, 2, \ldots, N$, we have the (reweighted[1]) MCD-estimate

$$\hat{\boldsymbol{\mu}}_t^{MCD}(\mathbf{x}_t) = \hat{\boldsymbol{\mu}}_t^{MCD}(\mathbf{x}_{1t}, \mathbf{x}_{2t}, \ldots, \mathbf{x}_{Nt}) \tag{7.6}$$

In place of the cross-section averages $\mathbf{z}_{wt}$, we have the cross-section location estimates $\tilde{\mathbf{z}}_t$, where $\tilde{\mathbf{z}}_t =$. We will consider $\tilde{\mathbf{z}}_t = \underset{1 \leq i \leq N}{\text{med}}(\mathbf{z}_{it}\text{or}(\hat{\boldsymbol{\mu}}_t^{MCD}(y_t), \hat{\boldsymbol{\mu}}_t^{MCD}(\mathbf{x}_t))$. The factor model (7.2) now becomes

$$y_{it} = \boldsymbol{\alpha}_i' \mathbf{d}_t + \boldsymbol{\beta}_i' \mathbf{x}_{it} + \boldsymbol{\theta}_i' \tilde{\mathbf{z}}_t + e_{it} \tag{7.7}$$

7.2.4 The Robust CCEMG estimator

The CCERG estimator applies a robust estimate $T(\mathbf{Z})$ of (7.7) to each cross-section unit:

$$\hat{\mathbf{b}}_i^{RG} = T(\mathbf{y}_i, \mathbf{X}_i, \tilde{\mathbf{z}}_t, \mathbf{d}_t) \tag{7.8}$$

Where T is a robust estimator, such as the LMS or LTS for example. The robust estimator T addresses the presence of outliers in each individual cross-section unit. $\mathbf{y}_i = (y_{i1}, y_{i2}, \ldots, y_{iT})'$ and $\mathbf{X}_i = (\mathbf{x}_{i1}', \mathbf{x}_{i2}', \ldots, \mathbf{x}_{iT}')'$ are the stacked forms of the data y_{it} and $\mathbf{x}_{it}$ by cross-section. We note that the breakdown point of T determines the proportion of outliers a single unit can take before the individual slope estimate fails, i.e. becomes outlying itself. The variability of each individual slope estimate depends on the efficiency of T. The CCERG estimator is then a robust location estimate of the individual slope estimates $\hat{\mathbf{b}}_i^{RG}$:

$$\hat{\boldsymbol{\beta}}_{CCERG} = R(\hat{\mathbf{b}}_1^{RG}, \hat{\mathbf{b}}_2^{RG}, \ldots, \hat{\mathbf{b}}_N^{RG}) \tag{7.9}$$

where R is a robust location estimate, such as the MCD or the (marginal) median. R is chosen to avoid outlying slope estimates due to sufficiently aberrant cross-section units. It then follows that the breakdown point of R determines the proportion of outlying cross-section units the CCERG estimator can take before it fails. This gives the CCERG estimator a (finite sample) breakdown point of

$$\epsilon_{CCERG} = \epsilon_T \times \epsilon_R \tag{7.10}$$

[1]The reweighted MCD can produce a more efficient estimate of the centre of a data cloud. Both the original and rewieghted MCD are also normally distributed, but the variance depends on the distribution of the data. For more information on the reweighted MCD, see Hubert, Debryune and Rousseeuw(2018).

where ϵ_T and ϵ_R are the breakdown points of the estimators T and R respectively. The maximal finite sample BP of the CCERG is therefore 25%, with both T and R both 50% BP estimators. However, when considering outliers in panel data, the situation is a little less straightforward. Unlike cross-section or time-series estimators, the two-dimensional structure of panel data allows for more outliers than is demarcated by the breakdown point. It is possible, for example, for each cross-section unit to be contaminated with 40% outliers and the estimates still turn up robust (if $\epsilon_T > 0.4$). If we consider outliers at the block level and the cell level separately, we note that each dimension can possibly attain the maximum finite BP of 50%. Thus, it is possible for the estimator to produce robust estimates even if up to 50% of the data are contaminated, provided the structure of the contamination permits. In practice, we often times do not know which observations are outlying, save for performing preliminary robust outlier diagnostics, but that is a separate field of study and beyond the scope of the paper. We therefore take the finite sample breakdown point of $\epsilon^{max}_{CCERG} = 0.25$ as the maximal breakdown point when faced with an unknown sample for a one-step robust estimation procedure.

7.2.5 The Robust CCEP estimator

The CCERP estimator is simply a pooled robust estimate of (7.7):

$$\hat{\boldsymbol{\beta}}_{CCERP} = T(\mathbf{y}, \mathbf{X}, J_N \otimes \tilde{\mathbf{z}}_t, J_N \otimes \mathbf{d}_t) \tag{7.11}$$

where T is a robust estimator as in (7.8), $\otimes$ is the Kronecker product and J_N is a $N \times 1$ vector of ones. $\mathbf{y}$ and $\mathbf{X}$ are the overall stacked forms of the data. We can improve the efficiency of the slope estimates by using high-efficiency estimators, such as an MM-estimator. The CCERP estimator therefore has a BP of ϵ_T, and hence a maximum BP of 50% (For example, the MM-estimator can be tuned with a breakdown point of 50% and a relative efficiency of 85%). In the next section, we perform simulations for the original CCE estimators as well as the robust CCE estimators.

7.3 Monte Carlo Studies

We performed a Monte Carlo (MC) study of the CCE under contamination. We used the same data-generating process (DGP) as in Pesaran's (2006) paper. We first evaluate the performance of the CCE estimators under outliers. We then estimate the same datasets with the CCERG and CCERP estimators. We then present and evaluate the results of the simulations.

7.3.1 Model Specifications

The DGP of Pesaran (2006) is given by

$$y_{it} = \alpha_{1i}d_{1t} + \beta_{1i}x_{1it} + \beta_{2i}x_{2it} + \gamma_{1i}f_{1t} + \gamma_{2i}f_{2t} + \epsilon_{it} \tag{7.12}$$

$$x_{jit} = a_{1ji}d_{1t} + a_{2ji}d_{2t} + \gamma_{1ji}f_{1t} + \gamma_{3ji}f_{3t} + v_{jit} \qquad j = 1, 2 \tag{7.13}$$

For each $i = 1, 2, \ldots, N$, the common effects are modelled as

$$\begin{aligned} d_{1t} &= 1 \\ d_{2t} &= 0.5d_{2,t-1} + \epsilon_{dt} \qquad \epsilon_{dt} \sim \text{iid}.N(0, (1-0.5)^2) \\ f_{jt} &= 0.5f_{j,t-1} + \epsilon_{fjt} \qquad \epsilon_{fjt} \sim \text{iid}.N(0, (1-0.5)^2) \qquad j = 1, 2, 3 \end{aligned} \tag{7.14}$$

The individual specific errors v_{jit} of x_{jit} are given by

$$\begin{aligned} v_{jit} &= \rho_{vj}v_{ji,t-1} + \epsilon_{vjt} \\ \epsilon_{vjt} &\sim \text{iid}.N(0, 1-\rho_{vj}^2) \\ \rho_{vj} &\sim \text{iid}.U(0.05, 0.95) \end{aligned} \tag{7.15}$$

The factor loadings of the common observed effects $\mathbf{d}_t$ are generated as

$$\begin{aligned} \alpha_{1i} &\sim \text{iid}.N(1, 1) \\ a_{jli} &\sim \text{iid}.N(0.5, 0.5) \qquad \text{for all combinations } j, l = 1, 2 \end{aligned} \tag{7.16}$$

The factor loadings of the unobserved effects $\mathbf{f}_t$ are generated as

$$\begin{gathered} \gamma_{ji} \sim \text{iid}.N(1, 0.2) \qquad j = 1, 2 \\ \begin{pmatrix} \gamma_{11i} & \gamma_{12i} & \gamma_{13i} \\ \gamma_{21i} & \gamma_{22i} & \gamma_{23i} \end{pmatrix} \sim \text{iid}. \begin{pmatrix} N(0.5, 0.5) & 0 & N(0, 0.5) \\ N(0, 0.5) & 0 & N(0.5, 0.5) \end{pmatrix} \end{gathered} \tag{7.17}$$

The individual-specific errors of y_{it} are generated as stationary AR(1) processes for $i = 1, 2, \ldots, N/2$

and as MA(1) processes for the rest (lines 62-70):

$$
\begin{aligned}
&\varepsilon_{it} = \rho_{i\varepsilon}\varepsilon_{i,t-1} + \sigma_i(1-\rho_{i\varepsilon}^2)^{1/2}\varsigma_{it} \qquad i = 1, 2, \ldots, N/2 \\
&\varepsilon_{it} = \sigma_i(1+\theta_{i\varepsilon}^2)^{-1/2}(\varsigma_{it} + \theta_{i\varepsilon}\varsigma_{i,t-1}) \qquad i = N/2+1, N/2+2, \ldots, N \\
&\sigma_i^2 \sim \text{iid}U(0.5, 1.5) \\
&\rho_{i\varepsilon} \sim \text{iid}U(0.05, 0.95) \\
&\theta_{i\varepsilon} \sim \text{iid}U(0, 1)
\end{aligned} \tag{7.18}
$$

For the CCERP, we select the MM-estimator with a 50% BP and and 85% efficiency for $T(Z)$ in (7.11). For the CCERG, we select the LTS estimator for $T(Z)$ in (7.8) and the MCD estimator for $R(\mathbf{b})$ in (7.9). Both the MM-estimator and the LTS can be computed in Stata using the *robreg* package (Verardi and Croux, 2009).

7.3.2 Outliers

We use the above specification for each experiment. We break our Experiment into 4 sub-experiments 1A, 1B, 1C, and 1D. We estimate the above specifications (7.12) to (7.18) with the with the CCERG and CCERP estimators. For each experiment, we introduce 4 levels of contamination, one for each sub-experiment. The levels of contamination are:

Experiments 1A: "Clean" (uncontaminated) data

Esperiments 1B: 5% of observations replaced with vertical outliers and leverage points

Esperiments 1C: 10% of observations replaced

Esperiments 1B: 20% of observations replaced

For Experiments B, C and D, the observations are randomly selected across the entire panel for replacement. The vertical outlier is $\tilde{y} \sim \text{iid}.N(0, 2)$ and the leverage points are distributed as $(\tilde{x}_1, \tilde{x}_2) \sim \text{iid}.(N(2, 2), N(2, 2))$. Fr each sub-experiment, we perform 4 regressions for the Pooled CCE and 4 regressions for the Mean Group CCE per repetition. For the Pooled estimation, we perform the original CCEP estimator; the CCEP estimator augmented with the usual cross-section averages $\bar{\mathbf{z}}_{wt}$ and the MM-estimator as $T(\mathbf{Z})$ (labelled MM in the tables); the CCEP estimator augmented with cross-section medians $\tilde{\mathbf{z}}_t$ and the MM-estimator as $T(\mathbf{Z})$ (labelled MM+Med); and the CCEP estimator augmented with MCD cross-section means $\hat{\boldsymbol{\mu}}_t^{MCD}(\mathbf{x}_t)$ and the MM-estimator as $T(\mathbf{Z})$ (labelled MCD). For the Mean Group Estimation, we perform the original CCEMG estimation and the MM, MM+Med and MCD estimators, with the MCD estimator as the overall mean group estimator $R(\mathbf{Z}$ for the CCERG.

7.4 Results

We report the biases and root mean-squared errors (RMSEs) for each simulation. Tables 7.1 to 7.4 record the results for the CCERP estimator for β_1. Tables 7.5 to 7.8 report the results for the CCERG estimator for β_1. We report the results for β_2 in the appendix; they are similar. For the CCERP and the CCEMG, we can see that the MM, MM+Med and MCD versions are resistant to outliers, maintaining $O(n^{-2}$ or even $O(n^{-3}$ biases and RMSEs even in the face of heavy contamination, whereas for the original estimators, the biases and RMSEs have ballooned to $O(1)$ in the face of contamination. The MM version of both the CCEP and CCERG estimators support Eberhardt's (2012) findings that robust regression estimators $T(\mathbf{Z}$ make the CCE estimators outlier proof. Curiously, the MM estimator peforms as well as the MM+Med and MCD versions, implying that robust cross-section averages are not necessary. Perhaps the outliers are thinly spread out across the cross-section units; future investigations into (extreme) outlying units (or even half of units with extreme outliers) would be of interest. All 3 robust estimators show slight increases in bias and RMSEs as N and T increase, but appear to be capped. This is due to the bounded influence functions of the robust estimators involved, which puts bounds on the biases that the estimates can deviate to. With bounded influence functions, robust estimators still have asymptotic biases, but they are capped[2]. Nevertheless, the MM, MM+Median and MCD estimators have demonstrated that the CCERP and CCERG are robust estimators over the originals.

Table 7.1: Experiment 1A - Small Sample Properties of the CCERP Estimators (β_1)

		CCEP		MM		MM+Med		MCD	
N	T	Bias	RMSE	Bias	RMSE	Bias	RMSE	Bias	RMSE
20	20	0.034	0.112	0.027	0.066	0.019	0.065	0.019	0.066
30	20	0.030	0.094	0.006	0.051	0.000	0.052	0.000	0.053
30	30	0.030	0.077	0.005	0.042	-0.002	0.043	-0.002	0.043
50	20	0.033	0.079	0.002	0.042	-0.003	0.043	-0.003	0.044
50	30	0.032	0.067	0.002	0.035	-0.003	0.036	-0.003	0.036
50	50	0.032	0.058	0.002	0.028	-0.004	0.029	-0.003	0.029
100	20	0.016	0.053	-0.015	0.032	-0.017	0.033	-0.017	0.033
100	30	0.015	0.044	-0.014	0.027	-0.017	0.029	-0.017	0.028
100	50	0.016	0.037	-0.014	0.023	-0.017	0.025	-0.017	0.024
100	100	0.016	0.031	-0.014	0.019	-0.018	0.022	-0.017	0.021
200	20	0.002	0.038	-0.011	0.022	-0.012	0.023	-0.011	0.023
200	30	0.001	0.031	-0.011	0.019	-0.012	0.020	-0.012	0.020
200	50	0.002	0.026	-0.011	0.016	-0.013	0.017	-0.012	0.017
200	100	0.002	0.021	-0.011	0.014	-0.013	0.015	-0.012	0.014

Uncontaminated data

[2]see Adrover and Yohai, 2002 derived asymptotic biases for various robust estimators (quoted in Maronna et. al., 2019).

Table 7.2: Experiment 1B - Small Sample Properties of the CCERP Estimators (β_1)

		CCEP		MM		MM+Med		MCD	
N	T	Bias	RMSE	Bias	RMSE	Bias	RMSE	Bias	RMSE
20	20	-0.336	0.371	0.014	0.064	0.010	0.066	0.010	0.067
30	20	-0.426	0.449	-0.004	0.052	-0.007	0.054	-0.006	0.054
30	30	-0.407	0.423	-0.003	0.043	-0.007	0.044	-0.006	0.045
50	20	-0.402	0.417	-0.007	0.044	-0.010	0.045	-0.009	0.045
50	30	-0.385	0.395	-0.006	0.036	-0.010	0.037	-0.009	0.037
50	50	-0.369	0.375	-0.006	0.029	-0.010	0.030	-0.009	0.030
100	20	-0.364	0.371	-0.023	0.035	-0.025	0.037	-0.024	0.037
100	30	-0.348	0.353	-0.023	0.032	-0.025	0.033	-0.024	0.033
100	50	-0.335	0.337	-0.023	0.028	-0.026	0.030	-0.025	0.030
100	100	-0.322	0.324	-0.023	0.026	-0.026	0.028	-0.025	0.027
200	20	-0.454	0.457	-0.015	0.024	-0.016	0.025	-0.015	0.025
200	30	-0.432	0.435	-0.015	0.022	-0.017	0.023	-0.016	0.022
200	50	-0.415	0.416	-0.016	0.020	-0.017	0.021	-0.017	0.020
200	100	-0.400	0.400	-0.015	0.017	-0.017	0.018	-0.016	0.018

5% of cross-section units are contaminated

Table 7.3: Experiment 1C - Small Sample Properties of the CCERP Estimators (β_1)

		CCEP		MM		MM+Med		MCD	
N	T	Bias	RMSE	Bias	RMSE	Bias	RMSE	Bias	RMSE
20	20	-0.510	0.530	-0.001	0.067	-0.003	0.069	-0.003	0.071
30	20	-0.624	0.638	-0.012	0.054	-0.014	0.056	-0.013	0.057
30	30	-0.612	0.622	-0.014	0.046	-0.016	0.048	-0.015	0.048
50	20	-0.593	0.602	-0.016	0.047	-0.018	0.048	-0.018	0.049
50	30	-0.576	0.582	-0.017	0.040	-0.020	0.042	-0.019	0.042
50	50	-0.556	0.560	-0.017	0.034	-0.020	0.035	-0.019	0.035
100	20	-0.527	0.532	-0.035	0.044	-0.037	0.046	-0.036	0.045
100	30	-0.514	0.516	-0.034	0.041	-0.037	0.042	-0.036	0.042
100	50	-0.498	0.500	-0.035	0.038	-0.037	0.040	-0.036	0.039
100	100	-0.483	0.484	-0.035	0.036	-0.038	0.039	-0.037	0.038
200	20	-0.637	0.640	-0.021	0.028	-0.023	0.029	-0.022	0.029
200	30	-0.619	0.621	-0.022	0.027	-0.024	0.028	-0.023	0.027
200	50	-0.601	0.602	-0.022	0.025	-0.024	0.026	-0.023	0.025
200	100	-0.586	0.586	-0.022	0.023	-0.024	0.025	-0.023	0.024

10% of cross-section units are contaminated

Table 7.4: Experiment 1D - Small Sample Properties of the CCERP Estimators (β_1)

N	T	CCEP Bias	CCEP RMSE	MM Bias	MM RMSE	MM+Median Bias	MM+Median RMSE	MCD Bias	MCD RMSE
20	20	-0.667	0.679	-0.047	0.092	-0.051	0.097	-0.051	0.099
30	20	-0.812	0.820	-0.048	0.077	-0.050	0.080	-0.049	0.080
30	30	-0.800	0.805	-0.050	0.070	-0.053	0.073	-0.052	0.073
50	20	-0.772	0.777	-0.055	0.072	-0.058	0.075	-0.057	0.075
50	30	-0.758	0.761	-0.055	0.067	-0.058	0.070	-0.058	0.070
50	50	-0.742	0.744	-0.057	0.063	-0.061	0.067	-0.061	0.067
100	20	-0.682	0.685	-0.077	0.082	-0.080	0.085	-0.080	0.084
100	30	-0.668	0.669	-0.078	0.080	-0.081	0.084	-0.081	0.084
100	50	-0.658	0.659	-0.078	0.080	-0.083	0.084	-0.082	0.084
100	100	-0.645	0.646	-0.077	0.078	-0.082	0.083	-0.082	0.083
200	20	-0.805	0.807	-0.047	0.051	-0.049	0.053	-0.049	0.053
200	30	-0.792	0.793	-0.047	0.049	-0.050	0.052	-0.050	0.052
200	50	-0.777	0.778	-0.048	0.049	-0.051	0.052	-0.051	0.052
200	100	-0.764	0.765	-0.048	0.049	-0.051	0.052	-0.051	0.052

20% of cross-section units are contaminated

Table 7.5: Experiment 1A - Small Sample Properties of the CCERG Estimators (β_1)

N	T	CCEMG Bias	CCEMG RMSE	MM Bias	MM RMSE	MM+Median Bias	MM+Median RMSE	MCD Bias	MCD RMSE
20	20	0.034	0.115	0.032	0.631	0.015	0.658	0.009	0.686
30	20	0.028	0.094	0.025	0.629	0.011	0.636	0.008	0.662
30	30	0.028	0.075	0.029	0.521	0.010	0.523	0.008	0.533
50	20	0.031	0.080	0.032	0.638	0.021	0.636	0.019	0.664
50	30	0.032	0.065	0.035	0.528	0.020	0.529	0.022	0.538
50	50	0.032	0.056	0.034	0.443	0.019	0.446	0.021	0.447
100	20	0.016	0.055	0.019	0.631	0.008	0.627	0.010	0.663
100	30	0.016	0.044	0.019	0.521	0.007	0.519	0.011	0.530
100	50	0.016	0.036	0.017	0.430	0.007	0.430	0.009	0.435
100	100	0.016	0.030	0.015	0.359	0.006	0.357	0.009	0.360
200	20	-0.954	0.954	-0.918	1.280	-0.922	1.280	-0.919	1.315
200	30	-0.952	0.952	-0.895	1.170	-0.906	1.177	-0.904	1.188
200	50	0.002	0.024	0.002	0.441	-0.004	0.439	-0.001	0.445
200	100	0.002	0.019	0.001	0.368	-0.004	0.367	-0.001	0.370

Uncontaminated data

Table 7.6: Experiment 1B - Small Sample Properties of the CCERG Estimators (β_1)

		CCEMG		MM		MM+Median		MCD	
N	T	Bias	RMSE	Bias	RMSE	Bias	RMSE	Bias	RMSE
20	20	-0.227	0.268	-0.017	0.683	-0.023	0.702	-0.032	0.730
30	20	-0.280	0.308	-0.014	0.672	-0.017	0.673	-0.025	0.712
30	30	-0.309	0.325	0.003	0.544	-0.005	0.550	-0.005	0.563
50	20	-0.258	0.274	-0.004	0.672	-0.012	0.672	-0.013	0.706
50	30	-0.284	0.294	0.005	0.551	-0.002	0.551	-0.001	0.561
50	50	-0.310	0.316	-0.004	0.422	-0.013	0.424	-0.010	0.428
100	20	-0.227	0.235	-0.023	0.661	-0.029	0.656	-0.029	0.692
100	30	-0.253	0.258	-0.010	0.523	-0.014	0.520	-0.013	0.537
100	50	-0.276	0.279	-0.003	0.442	-0.009	0.440	-0.007	0.445
100	100	-0.292	0.294	0.004	0.367	-0.004	0.366	0.000	0.369
200	20	-0.285	0.289	-0.027	0.668	-0.031	0.662	-0.034	0.711
200	30	-0.315	0.317	-0.016	0.534	-0.020	0.529	-0.018	0.550
200	50	-0.342	0.343	-0.011	0.453	-0.014	0.451	-0.011	0.459
200	100	-0.361	0.362	-0.007	0.378	-0.011	0.377	-0.009	0.378

5% of cross-section units are contaminated

Table 7.7: Experiment 1C - Small Sample Properties of the CCERG Estimators (β_1)

		CCEMG		MM		MM+Median		MCD	
N	T	Bias	RMSE	Bias	RMSE	Bias	RMSE	Bias	RMSE
20	20	-0.405	0.430	-0.079	0.738	-0.083	0.757	-0.092	0.790
30	20	-0.486	0.502	-0.068	0.732	-0.068	0.740	-0.077	0.778
30	30	-0.524	0.534	-0.034	0.585	-0.036	0.588	-0.036	0.603
50	20	-0.453	0.463	-0.055	0.724	-0.058	0.722	-0.066	0.762
50	30	-0.483	0.489	-0.023	0.584	-0.022	0.580	-0.022	0.594
50	50	-0.506	0.510	-0.004	0.473	-0.007	0.474	-0.006	0.475
100	20	-0.395	0.400	-0.072	0.700	-0.077	0.700	-0.083	0.740
100	30	-0.424	0.427	-0.041	0.544	-0.046	0.545	-0.046	0.565
100	50	-0.448	0.449	-0.025	0.456	-0.026	0.456	-0.027	0.465
100	100	-0.460	0.461	-0.013	0.377	-0.017	0.377	-0.015	0.380
200	20	-0.483	0.486	-0.070	0.718	-0.072	0.713	-0.081	0.766
200	30	-0.514	0.516	-0.038	0.559	-0.039	0.554	-0.040	0.580
200	50	-0.540	0.541	-0.024	0.469	-0.027	0.469	-0.026	0.476
200	100	-0.553	0.554	-0.016	0.389	-0.019	0.388	-0.017	0.390

10% of cross-section units are contaminated

Table 7.8: Experiment 1D - Small Sample Properties of the CCERG Estimators (β_1)

		CCEMG		MM		MM+Median		MCD	
N	T	Bias	RMSE	Bias	RMSE	Bias	RMSE	Bias	RMSE
20	20	-0.613	0.627	-0.243	0.879	-0.249	0.905	-0.271	0.944
30	20	-0.736	0.746	-0.237	0.906	-0.237	0.928	-0.256	0.966
30	30	-0.751	0.757	-0.130	0.699	-0.129	0.714	-0.137	0.731
50	20	-0.694	0.700	-0.212	0.869	-0.217	0.878	-0.238	0.932
50	30	-0.706	0.710	-0.112	0.679	-0.116	0.683	-0.124	0.706
50	50	-0.718	0.721	-0.054	0.527	-0.057	0.538	-0.056	0.541
100	20	-0.605	0.608	-0.220	0.804	-0.223	0.806	-0.239	0.851
100	30	-0.619	0.621	-0.132	0.616	-0.138	0.622	-0.145	0.644
100	50	-0.632	0.633	-0.085	0.508	-0.088	0.512	-0.090	0.519
100	100	-0.636	0.637	-0.056	0.411	-0.058	0.416	-0.059	0.417
200	20	-0.715	0.717	-0.212	0.853	-0.216	0.854	-0.245	0.918
200	30	-0.736	0.737	-0.114	0.647	-0.117	0.647	-0.130	0.678
200	50	-0.747	0.748	-0.067	0.525	-0.070	0.528	-0.072	0.537
200	100	-0.752	0.752	-0.045	0.426	-0.047	0.428	-0.048	0.431

20% of cross-section units are contaminated

7.5 Conclusion

We have demonstrated the robustness of the CCERG and CCERP in the face of outliers. Immediate future research includes estimation of robust variances and confidence intervals, extreme outlying cross-section units (extreme block outliers) and asymmetric robust statistics for the general factor model. Finally, there is much work to do regarding the aymptotics of the CCERG and CCERP estimators. We hope to cover some of these topics in future.

References

Allende, H., Frery, A.C., Galbiati, J. and Pizarro, L., 2006. M-estimators with asymmetric influence functions: the distribution case. *Journal of Statistical Computation and Simulation*, 76(11), pp.941-956. Bramati, M.C. and Croux, C., 2007. Robust estimators for the fixed effects panel data model. *The econometrics journal*, 10(3), pp.521-540.

De Maesschalck, R., Jouan-Rimbaud, D. and Massart, D.L., 2000. The mahalanobis distance. Chemometrics and intelligent laboratory systems, 50(1), pp.1-18.

Ditzen, J., 2018. Estimating dynamic common-correlated effects in Stata. *The Stata Journal*, 18(3), pp.585-617.

Eberhardt, M., 2012. Estimating panel time-series models with heterogeneous slopes. *The Stata Journal*, 12(1), pp.61-71.

Hardin, J. and Rocke, D.M., 2005. The distribution of robust distances. *Journal of Computational and Graphical Statistics*, 14(4), pp.928-946.

Hubert, M., Debruyne, M. and Rousseeuw, P.J., 2018. Minimum covariance determinant and extensions. *Wiley Interdisciplinary Reviews: Computational Statistics*, 10(3), p.e1421.

Hubert, M. and Van der Veeken, S., 2008. Outlier detection for skewed data. *Journal of Chemometrics: A Journal of the Chemometrics Society*, 22(3-4), pp.235-246.

Ben Jann, 2010. ROBREG10: Stata module providing robust regression estimators, *Statistical Software Components S457114*, Boston College Department of Economics, revised 11 Apr 2021.

Lee, Y. and Sul, D., 2020. Depth-Weighted Estimation of Heterogenous Agent Panel Data Models. *Mimeo, Syracuse University.*

Lee, Y. and Sul, D., 2022. Trimmed mean group estimation. In Essays in Honor of M. Hashem Pesaran: Panel Modeling, Micro Applications, and Econometric Methodology (pp. 177-202). *Emerald Publishing Limited.*

Mahalanobis, P.C., 1936. Mahalanobis distance. *In Proceedings National Institute of Science of India* (Vol. 49, No. 2, pp. 234-256).

Maronna, R.A., Martin, R.D., Yohai, V.J. and Salibián-Barrera, M., 2019. *Robust statistics: theory* *and methods (with R)*. John Wiley & Sons.

Pesaran, M., 2006. Estimation and Inference in Large Heterogeneous Panels with a Multifactor Error Structure. *Econometrica*, 74(4), pp.967-1012.

Pesaran, M.H. and Smith, R., 1995. Estimating long-run relationships from dynamic heterogeneous panels. *Journal of econometrics*, 68(1), pp.79-113.

Rousseeuw, P.J., 1984. Least median of squares regression.

Journal of the American statistical association, 79(388), pp.871-880.

Rousseeuw, P. and Leroy, A., 1987. *Robust regression and outlier detection.* Hoboken, NJ: Wiley-Interscience.

Rousseeuw, P.J. and Hubert, M., 2011. Robust statistics for outlier detection. *Wiley interdisciplinary reviews: Data mining and knowledge discovery*, 1(1), pp.73-79.

Rousseeuw, P.J. and Van Zomeren, B.C., 1990. Unmasking multivariate outliers and leverage points. *Journal of the American Statistical association*, 85(411), pp.633-639.

Van Aelst, S. and Rousseeuw, P., 2009. Minimum volume ellipsoid. *Wiley Interdisciplinary Reviews: Computational Statistics*, 1(1), pp.71-82.

Verardi, V. and Dehon, C., 2010. Multivariate outlier detection in Stata. *The Stata Journal*, 10(2), pp.259-266.

Verardi, V. and McCathie, A., 2012. The S-estimator of multivariate location and scatter in Stata. *The Stata Journal*, 12(2), pp.299-307.

Westerlund, J., Petrova, Y. and Norkute, M., 2019. CCE in fixed-T panels. *Journal of Applied Econometrics*, 34(5), pp.746-761.

Yohai, V.J., 1987. High breakdown-point and high efficiency robust estimates for regression. *The Annals of statistics*, pp.642-656.

Appendix

7.A CCERP Results for β_1

Table 7.9: Experiment 1A - Small Sample Properties of the CCERP Estimators (β_2)

		CCEP		MM		MM+Med		MCD	
N	T	Bias	RMSE	Bias	RMSE	Bias	RMSE	Bias	RMSE
20	20	-0.042	0.113	-0.047	0.074	-0.058	0.081	-0.059	0.083
30	20	-0.050	0.101	-0.037	0.059	-0.045	0.064	-0.045	0.065
30	30	-0.050	0.086	-0.035	0.051	-0.044	0.057	-0.044	0.058
50	20	-0.054	0.089	-0.055	0.064	-0.060	0.068	-0.060	0.069
50	30	-0.055	0.079	-0.055	0.061	-0.061	0.066	-0.060	0.066
50	50	-0.054	0.071	-0.055	0.058	-0.061	0.064	-0.061	0.064
100	20	-0.050	0.070	-0.044	0.049	-0.047	0.052	-0.046	0.051
100	30	-0.049	0.063	-0.043	0.047	-0.047	0.050	-0.046	0.049
100	50	-0.049	0.059	-0.044	0.046	-0.048	0.050	-0.047	0.049
100	100	-0.049	0.056	-0.044	0.045	-0.048	0.049	-0.047	0.047
200	20	-0.012	0.040	-0.004	0.020	-0.005	0.020	-0.004	0.021
200	30	-0.012	0.033	-0.004	0.017	-0.006	0.017	-0.005	0.018
200	50	-0.012	0.028	-0.003	0.014	-0.005	0.014	-0.004	0.014
200	100	-0.013	0.024	-0.004	0.010	-0.006	0.011	-0.005	0.011

Uncontaminated data

Table 7.10: Experiment 1B - Small Sample Properties of the CCERP Estimators (β_2)

		CCEP		MM		MM+Med		MCD	
N	T	Bias	RMSE	Bias	RMSE	Bias	RMSE	Bias	RMSE
20	20	-0.406	0.432	-0.061	0.083	-0.069	0.090	-0.070	0.091
30	20	-0.465	0.487	-0.043	0.063	-0.050	0.068	-0.049	0.068
30	30	-0.449	0.463	-0.044	0.057	-0.051	0.062	-0.050	0.063
50	20	-0.469	0.482	-0.061	0.070	-0.066	0.074	-0.065	0.073
50	30	-0.456	0.464	-0.061	0.067	-0.066	0.071	-0.065	0.070
50	50	-0.439	0.444	-0.061	0.064	-0.067	0.069	-0.065	0.068
100	20	-0.416	0.422	-0.051	0.055	-0.053	0.058	-0.052	0.057
100	30	-0.403	0.407	-0.052	0.055	-0.055	0.058	-0.054	0.057
100	50	-0.390	0.392	-0.052	0.053	-0.055	0.057	-0.054	0.056
100	100	-0.377	0.378	-0.051	0.052	-0.055	0.056	-0.054	0.054
200	20	-0.474	0.477	-0.008	0.021	-0.009	0.022	-0.008	0.022
200	30	-0.453	0.455	-0.007	0.018	-0.009	0.019	-0.008	0.018
200	50	-0.436	0.437	-0.007	0.015	-0.009	0.016	-0.008	0.015
200	100	-0.422	0.423	-0.008	0.012	-0.010	0.013	-0.009	0.013

5% of cross-section units are contaminated

Table 7.11: Experiment 1C - Small Sample Properties of the CCERP Estimators (β_2)

		CCEP		MM		MM+Med		MCD	
N	T	Bias	RMSE	Bias	RMSE	Bias	RMSE	Bias	RMSE
20	20	-0.574	0.592	-0.077	0.096	-0.084	0.102	-0.085	0.104
30	20	-0.642	0.656	-0.050	0.069	-0.056	0.073	-0.055	0.074
30	30	-0.629	0.639	-0.052	0.064	-0.059	0.069	-0.058	0.069
50	20	-0.650	0.658	-0.071	0.078	-0.075	0.082	-0.074	0.081
50	30	-0.635	0.641	-0.069	0.074	-0.074	0.079	-0.073	0.078
50	50	-0.617	0.621	-0.070	0.073	-0.076	0.078	-0.074	0.077
100	20	-0.573	0.577	-0.062	0.066	-0.065	0.068	-0.064	0.067
100	30	-0.562	0.565	-0.062	0.064	-0.065	0.067	-0.064	0.066
100	50	-0.545	0.547	-0.062	0.064	-0.066	0.067	-0.065	0.066
100	100	-0.532	0.533	-0.062	0.062	-0.066	0.066	-0.064	0.065
200	20	-0.661	0.663	-0.013	0.023	-0.014	0.024	-0.013	0.024
200	30	-0.644	0.646	-0.013	0.021	-0.015	0.022	-0.014	0.021
200	50	-0.627	0.628	-0.013	0.018	-0.015	0.020	-0.014	0.019
200	100	-0.611	0.611	-0.014	0.016	-0.016	0.018	-0.015	0.017

10% of cross-section units are contaminated

Table 7.12: Experiment 1D - Small Sample Properties of the CCERP Estimators (β_2)

		CCEP		MM		MM+Med		MCD	
N	T	Bias	RMSE	Bias	RMSE	Bias	RMSE	Bias	RMSE
20	20	-0.740	0.750	-0.125	0.141	-0.134	0.150	-0.135	0.152
30	20	-0.822	0.831	-0.079	0.094	-0.086	0.100	-0.085	0.100
30	30	-0.810	0.816	-0.083	0.092	-0.091	0.099	-0.090	0.099
50	20	-0.809	0.814	-0.101	0.107	-0.107	0.113	-0.106	0.112
50	30	-0.800	0.803	-0.103	0.106	-0.109	0.112	-0.108	0.112
50	50	-0.787	0.789	-0.104	0.106	-0.111	0.113	-0.110	0.112
100	20	-0.717	0.720	-0.097	0.100	-0.102	0.105	-0.101	0.104
100	30	-0.706	0.707	-0.100	0.102	-0.106	0.108	-0.104	0.106
100	50	-0.696	0.697	-0.100	0.102	-0.106	0.107	-0.105	0.106
100	100	-0.687	0.687	-0.100	0.101	-0.107	0.107	-0.106	0.106
200	20	-0.830	0.831	-0.036	0.041	-0.038	0.043	-0.038	0.043
200	30	-0.817	0.818	-0.036	0.039	-0.039	0.042	-0.039	0.042
200	50	-0.804	0.805	-0.037	0.038	-0.040	0.042	-0.040	0.041
200	100	-0.792	0.793	-0.037	0.037	-0.041	0.041	-0.040	0.041

20% of cross-section units are contaminated

7.B CCERG results for β_2

Table 7.13: Experiment 1A - Small Sample Properties of the CCERG Estimators (β_2)

		CCEP		MM		MM+Med		MCD	
N	T	Bias	RMSE	Bias	RMSE	Bias	RMSE	Bias	RMSE
20	20	-0.045	0.117	-0.047	0.638	-0.072	0.653	-0.081	0.677
30	20	-0.052	0.103	-0.056	0.625	-0.083	0.630	-0.086	0.654
30	30	-0.049	0.084	-0.053	0.518	-0.081	0.531	-0.081	0.534
50	20	-0.052	0.089	-0.054	0.634	-0.075	0.637	-0.071	0.667
50	30	-0.055	0.078	-0.056	0.532	-0.076	0.533	-0.074	0.545
50	50	-0.053	0.069	-0.052	0.450	-0.076	0.453	-0.072	0.454
100	20	-0.049	0.071	-0.053	0.643	-0.063	0.637	-0.060	0.670
100	30	-0.049	0.062	-0.051	0.533	-0.063	0.530	-0.060	0.545
100	50	-0.049	0.058	-0.050	0.447	-0.063	0.446	-0.058	0.450
100	100	-0.049	0.055	-0.053	0.377	-0.064	0.379	-0.061	0.379
200	20	-0.976	0.977	-0.936	1.290	-0.944	1.294	-0.943	1.327
200	30	-0.974	0.974	-0.912	1.185	-0.930	1.192	-0.929	1.204
200	50	-0.012	0.026	-0.012	0.444	-0.018	0.443	-0.016	0.448
200	100	-0.012	0.022	-0.011	0.373	-0.018	0.372	-0.015	0.374

Uncontaminated data

Table 7.14: Experiment 1B - Small Sample Properties of the CCERG Estimators (β_2)

		CCEMG		MM		MM+Med		MCD	
N	T	Bias	RMSE	Bias	RMSE	Bias	RMSE	Bias	RMSE
20	20	-0.304	0.333	-0.097	0.677	-0.119	0.701	-0.118	0.731
30	20	-0.336	0.358	-0.099	0.671	-0.113	0.678	-0.112	0.717
30	30	-0.362	0.376	-0.084	0.550	-0.098	0.556	-0.095	0.563
50	20	-0.332	0.344	-0.095	0.684	-0.109	0.679	-0.111	0.714
50	30	-0.358	0.366	-0.084	0.563	-0.097	0.558	-0.092	0.573
50	50	-0.314	0.319	-0.023	0.430	-0.039	0.433	-0.035	0.436
100	20	-0.286	0.292	-0.091	0.675	-0.100	0.674	-0.104	0.715
100	30	-0.309	0.313	-0.078	0.537	-0.087	0.535	-0.085	0.555
100	50	-0.330	0.333	-0.072	0.460	-0.081	0.461	-0.075	0.466
100	100	-0.348	0.349	-0.066	0.389	-0.074	0.389	-0.072	0.388
200	20	-0.303	0.307	-0.044	0.673	-0.049	0.666	-0.051	0.715
200	30	-0.334	0.336	-0.033	0.537	-0.036	0.534	-0.034	0.555
200	50	-0.361	0.363	-0.023	0.457	-0.029	0.455	-0.027	0.461
200	100	-0.381	0.382	-0.020	0.380	-0.025	0.380	-0.023	0.384

5% of cross-section units are contaminated

Table 7.15: Experiment 1C - Small Sample Properties of the CCERG Estimators (β_2)

		CCEMG		MM		MM+Med		MCD	
N	T	Bias	RMSE	Bias	RMSE	Bias	RMSE	Bias	RMSE
20	20	-0.478	0.498	-0.158	0.737	-0.174	0.756	-0.181	0.798
30	20	-0.529	0.544	-0.148	0.745	-0.150	0.741	-0.158	0.781
30	30	-0.557	0.567	-0.116	0.595	-0.124	0.593	-0.121	0.597
50	20	-0.518	0.527	-0.145	0.743	-0.153	0.736	-0.157	0.778
50	30	-0.547	0.553	-0.115	0.593	-0.120	0.591	-0.122	0.608
50	50	-0.571	0.575	-0.094	0.487	-0.106	0.488	-0.103	0.492
100	20	-0.446	0.451	-0.142	0.715	-0.150	0.713	-0.156	0.760
100	30	-0.476	0.478	-0.110	0.565	-0.118	0.565	-0.116	0.584
100	50	-0.497	0.498	-0.092	0.480	-0.099	0.481	-0.097	0.485
100	100	-0.510	0.511	-0.082	0.401	-0.090	0.404	-0.086	0.402
200	20	-0.505	0.507	-0.085	0.722	-0.089	0.718	-0.098	0.772
200	30	-0.536	0.538	-0.054	0.566	-0.059	0.563	-0.059	0.584
200	50	-0.562	0.563	-0.038	0.475	-0.042	0.472	-0.039	0.479
200	100	-0.577	0.578	-0.028	0.392	-0.034	0.393	-0.031	0.395

10% of cross-section units are contaminated

Table 7.16: Experiment 1D - Small Sample Properties of the CCERG Estimators (β_2)

		CCEMG		MM		MM+Med		MCD	
N	T	Bias	RMSE	Bias	RMSE	Bias	RMSE	Bias	RMSE
20	20	-0.682	0.694	-0.319	0.877	-0.336	0.895	-0.357	0.954
30	20	-0.759	0.769	-0.305	0.924	-0.318	0.926	-0.339	0.986
30	30	-0.775	0.781	-0.206	0.711	-0.220	0.726	-0.227	0.744
50	20	-0.742	0.748	-0.304	0.901	-0.311	0.895	-0.325	0.946
50	30	-0.756	0.760	-0.205	0.698	-0.218	0.707	-0.221	0.726
50	50	-0.769	0.771	-0.147	0.551	-0.158	0.558	-0.157	0.563
100	20	-0.648	0.651	-0.283	0.827	-0.293	0.828	-0.306	0.878
100	30	-0.663	0.665	-0.201	0.647	-0.213	0.649	-0.217	0.673
100	50	-0.675	0.676	-0.155	0.534	-0.163	0.538	-0.164	0.545
100	100	-0.679	0.679	-0.126	0.441	-0.133	0.445	-0.133	0.447
200	20	-0.738	0.739	-0.230	0.864	-0.238	0.862	-0.267	0.931
200	30	-0.763	0.764	-0.131	0.655	-0.139	0.657	-0.149	0.688
200	50	-0.776	0.776	-0.082	0.532	-0.089	0.532	-0.090	0.542
200	100	-0.779	0.779	-0.057	0.430	-0.063	0.432	-0.063	0.435

20% of cross-section units are contaminated

Chapter 8

Conclusion

We have showed the sensitivity of the CCE estimators to outliers. We have reviewed and robustified the existing TCCE estimators in the literature and applied them to a few studies. We have explored new robust CCE estimators with potential for one-step outlier-robust estimation. For future research, we would like to explore the robust CCE estimators further. We would like to solve the asymmetry problem in the robust CCE estimators and explore the asymptotic theory behind the robust CCE estimators. All these are open for future research.