Influence Analysis with Panel Data

Annalivia Polselli

Institute for Analytics and Data Science (IADS)
University of Essex

ICEEE 2023, University of Cagliari

May 26, 2023

Motivation

- Short panel data sets (small N but $N \gg T$) are common in many fields of Economics, e.g.
  - Macro-level panel data
  - Experimental panel data
- Observational data may contain “anomalous” observations (Rousseeuw and Van Zomeren, 1990; Silva, 2001)
  - Vertical outliers (VO), good leverage (GL) points, bad leverage (BL) points ▸ Example ▸ DGP
- Large influence on the Least Squares (LS) estimates
  $\Longrightarrow$ Biased regression coefficients or standard errors (Donald and Maddala, 1993; Bramati and Croux, 2007; Verardi and Croux, 2009)

Motivation

- Tools for the detection of leveraged and outlying observations
  1. Diagnostic plots (e.g., leverage-vs-residual plots)
  2. Measures of influence (e.g., Cook (1979)'s distance)
- Limitations
  1. For cross-sectional data and less handy for panel data
  2. Fail to flag multiple atypical cases (Atkinson and Mulira, 1993; Chatterjee and Hadi, 1988; Rousseeuw and Van Zomeren, 1990)

This paper

- I develop a method to **detect** and **identify** the type of anomalous units in fixed effects models for short panel data sets
- Following a *unit-wise* approach
  - I formalise measures
    1. Average leverage and residuals
    2. Joint and conditional influence
  - I present graphical tools
    1. Leverage-vs-residual plots
    2. Influence plots

Contribution

- **Individual diagnostic measures**: for cross-sectional models (Martín and Pardo, 2009; Martín, 2015; Pinho et al., 2015), and linear panel data models (Banerjee and Frees, 1997; Belotti and Peracchi, 2020)
  - Fail to detect multiple anomalous cases (Atkinson, 1985; Chatterjee and Hadi, 1988; Rousseeuw and Van Zomeren, 1990; Rousseeuw, 1991)
- **Pair-wise diagnostic measures**: for cross-sectional models (Lawrance, 1995)
  - I extend them to linear panel data models with fixed
  - These resemble a directed and weighted adjacency matrix
    $\Rightarrow$ Plot unit j's influence on i's

Model and estimators

- Model after the *within-group* (WG) transformation

$$\widetilde{y}_{it} = \widetilde{\mathbf{x}}'_{it}\boldsymbol{\beta} + \widetilde{u}_{it}$$

where $\widetilde{y}_{it} = y_{it} - T^{-1}\sum_t y_{it}$, etc.

- WG Estimator

$$\widehat{\boldsymbol{\beta}} = \left(\sum_{i=1}^{N}\sum_{t=1}^{T}\widetilde{\mathbf{x}}_{it}\widetilde{\mathbf{x}}'_{it}\right)^{-1}\sum_{i=1}^{N}\widetilde{\mathbf{x}}_{it}\widetilde{y}_{it}$$

- Leave-One-Out estimator

$$\widehat{\boldsymbol{\beta}}_{(i)} = \widehat{\boldsymbol{\beta}} - (\widetilde{\mathbf{X}}'\widetilde{\mathbf{X}})^{-1}\widetilde{\mathbf{X}}'_i\mathbf{M}_i^{-1}\widehat{\mathbf{u}}_i$$

where $\mathbf{M}_i = \mathbf{I}_i - \mathbf{H}_{ii}$, $\mathbf{H}_{ii} = \widetilde{\mathbf{X}}_i(\widetilde{\mathbf{X}}'\widetilde{\mathbf{X}})^{-1}\widetilde{\mathbf{X}}'_i$, $\widehat{\mathbf{u}}_i = \widetilde{\mathbf{Y}} - \widetilde{\mathbf{X}}\widehat{\boldsymbol{\beta}}$

- Leave-Two-Out estimator

$$\begin{aligned}\widehat{\boldsymbol{\beta}}_{(i,j)} = \widehat{\boldsymbol{\beta}}_{(i)} - &(\widetilde{\mathbf{X}}'\widetilde{\mathbf{X}})^{-1}(\widetilde{\mathbf{X}}'_i\mathbf{M}_i^{-1}\mathbf{H}_{ij} + \widetilde{\mathbf{X}}'_j)\\ &(\mathbf{M}_j - \mathbf{H}'_{ij}\mathbf{M}_i^{-1}\mathbf{H}_{ij})^{-1}(\mathbf{H}'_{ij}\mathbf{M}_i^{-1}\widehat{\mathbf{u}}_i + \widehat{\mathbf{u}}_j)\end{aligned}$$

where $\mathbf{M}_j = \mathbf{I}_j - \mathbf{H}_j$, $\mathbf{H}_j = \widetilde{\mathbf{X}}_j(\widetilde{\mathbf{X}}'\widetilde{\mathbf{X}})^{-1}\widetilde{\mathbf{X}}'_j$, $\mathbf{H}_{ij} = \widetilde{\mathbf{X}}_i(\widetilde{\mathbf{X}}'\widetilde{\mathbf{X}})^{-1}\widetilde{\mathbf{X}}'_j$

Model and estimators

- Model after the *within-group* (WG) transformation

$$\widetilde{y}_{it} = \widetilde{\mathbf{x}}'_{it}\boldsymbol{\beta} + \widetilde{u}_{it}$$

where $\widetilde{y}_{it} = y_{it} - T^{-1}\sum_t y_{it}$, etc.

- WG Estimator

$$\widehat{\boldsymbol{\beta}} = \left(\sum_{i=1}^{N}\sum_{t=1}^{T}\widetilde{\mathbf{x}}_{it}\widetilde{\mathbf{x}}'_{it}\right)^{-1}\sum_{i=1}^{N}\widetilde{\mathbf{x}}_{it}\widetilde{y}_{it}$$

- Leave-One-Out estimator

$$\widehat{\boldsymbol{\beta}}_{(i)} = \widehat{\boldsymbol{\beta}} - \left(\widetilde{\mathbf{X}}'\widetilde{\mathbf{X}}\right)^{-1}\widetilde{\mathbf{X}}'_i\mathbf{M}_i^{-1}\widehat{\mathbf{u}}_i$$

where $\mathbf{M}_i = \mathbf{I}_i - \mathbf{H}_{ii}$, $\mathbf{H}_{ii} = \widetilde{\mathbf{X}}_i(\widetilde{\mathbf{X}}'\widetilde{\mathbf{X}})^{-1}\widetilde{\mathbf{X}}'_i$, $\widehat{\mathbf{u}}_i = \widetilde{\mathbf{Y}} - \widetilde{\mathbf{X}}\widehat{\boldsymbol{\beta}}$

- Leave-Two-Out estimator

$$\begin{aligned}\widehat{\boldsymbol{\beta}}_{(i,j)} = \widehat{\boldsymbol{\beta}}_{(i)} - \left(\widetilde{\mathbf{X}}'\widetilde{\mathbf{X}}\right)^{-1}&\left(\widetilde{\mathbf{X}}'_i\mathbf{M}_i^{-1}\mathbf{H}_{ij} + \widetilde{\mathbf{X}}'_j\right)\\ &\left(\mathbf{M}_j - \mathbf{H}'_{ij}\mathbf{M}_i^{-1}\mathbf{H}_{ij}\right)^{-1}\left(\mathbf{H}'_{ij}\mathbf{M}_i^{-1}\widehat{\mathbf{u}}_i + \widehat{\mathbf{u}}_j\right)\end{aligned}$$

where $\mathbf{M}_j = \mathbf{I}_j - \mathbf{H}_j$, $\mathbf{H}_j = \widetilde{\mathbf{X}}_j(\widetilde{\mathbf{X}}'\widetilde{\mathbf{X}})^{-1}\widetilde{\mathbf{X}}'_j$, $\mathbf{H}_{ij} = \widetilde{\mathbf{X}}_i(\widetilde{\mathbf{X}}'\widetilde{\mathbf{X}})^{-1}\widetilde{\mathbf{X}}'_j$

Residuals

The **average normalised residual** squared

$$\widehat{u}_i^* = \frac{1}{T}\sum_{t=1}^{T}\left(\frac{\widehat{u}_{it}}{\sqrt{\sum_i \widehat{u}_{it}^2}}\right)^2$$

where $\widehat{u}_{it} = \widetilde{y}_{it} - \widetilde{\mathbf{x}}_{it}'\widehat{\boldsymbol{\beta}}$ are LS Residuals.

Cut-off value: $c_{\widehat{u}_i^*} = \frac{2}{NT}$

Leverage

Individual leverage matrix

$$\underset{(T\times T)}{\mathbf{H}_{ii}} = \widetilde{\mathbf{X}}_i\left(\widetilde{\mathbf{X}}'\widetilde{\mathbf{X}}\right)^{-1}\widetilde{\mathbf{X}}_i' = \begin{pmatrix} h_{ii,11} & h_{ii,12} & \dots & h_{ii,1T} \\ h_{ii,21} & h_{ii,22} & \dots & h_{ii,2T} \\ \vdots & \vdots & \ddots & \vdots \\ h_{ii,T1} & h_{ii,T2} & \dots & h_{ii,TT} \end{pmatrix}$$

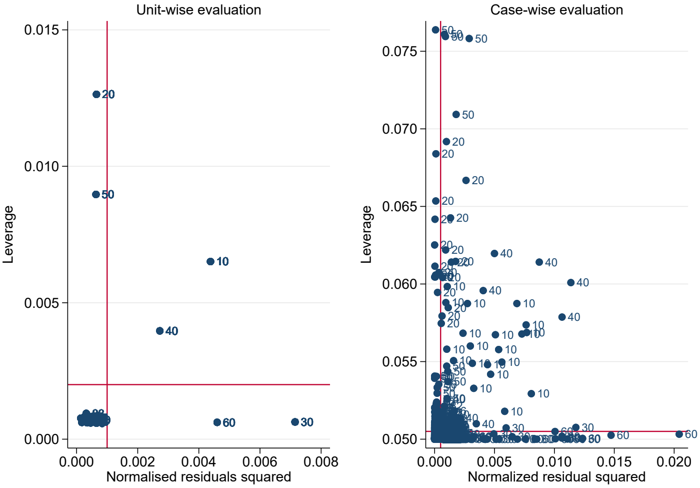
where $h_{ii,tt} = \widetilde{\mathbf{x}}_{it}'(\widetilde{\mathbf{X}}'\widetilde{\mathbf{X}})^{-1}\widetilde{\mathbf{x}}_{it}$, and $h_{ii,ts} = \widetilde{\mathbf{x}}_{it}'(\widetilde{\mathbf{X}}'\widetilde{\mathbf{X}})^{-1}\widetilde{\mathbf{x}}_{is}$ for $t, s = 1, \dots, T$.

The **average individual leverage** of unit i at time t is

$$\overline{h}_i = \frac{1}{T}\sum_{t=1}^{T} h_{ii,tt}$$

Cut-off value: $c_{\overline{h}_i} = \frac{2(K+1)}{NT}$

Leverage-vs-residual plot



Note: Units 10 and 40 are bad leverage units; units 20 and 50 are good leverage units; units 30 and 60 are vertical outliers.

Influence analysis: Overview

- How anomalous units may affect the LS estimates
  1. Joint influence
  2. Joint effects
  3. Conditional influence
  4. Conditional effects

Influence analysis: Joint influence

- For $i \neq j$,

$$\mathrm{C}_{ij}(\widehat{\boldsymbol{\beta}}) = \left(\widehat{\boldsymbol{\beta}} - \widehat{\boldsymbol{\beta}}_{(i,j)}\right)' \left(\widetilde{\mathbf{X}}'\widetilde{\mathbf{X}}\right) \left(\widehat{\boldsymbol{\beta}} - \widehat{\boldsymbol{\beta}}_{(i,j)}\right) (s^2 K)^{-1}$$

  - Influence exerted by a pair (i,j) on LS estimates *jointly*
  - Comparison of LS estimates *with* and *without* the *pair*
  - $\mathrm{C}_{ij}(\widehat{\boldsymbol{\beta}}) = \mathrm{C}_{ji}(\widehat{\boldsymbol{\beta}})$

- For $i = j$,

$$\mathrm{C}_{ii}(\widehat{\boldsymbol{\beta}}) = \left(\widehat{\boldsymbol{\beta}} - \widehat{\boldsymbol{\beta}}_{(i)}\right)' \left(\widetilde{\mathbf{X}}'\widetilde{\mathbf{X}}\right) \left(\widehat{\boldsymbol{\beta}} - \widehat{\boldsymbol{\beta}}_{(i)}\right) (s^2 K)^{-1}$$

- $\mathrm{C}_{ij}(\widehat{\boldsymbol{\beta}}), \mathrm{C}_{ii}(\widehat{\boldsymbol{\beta}}) \sim \mathrm{F}(\nu_1, \nu_2)$

  where $\nu_1 = k + 1$ and $\nu_2 = NT - N - (k + 1)$

Influence analysis: Joint influence

- For $i \neq j$,

$$\mathrm{C}_{ij}(\widehat{\boldsymbol{\beta}}) = \left(\widehat{\boldsymbol{\beta}} - \widehat{\boldsymbol{\beta}}_{(i,j)}\right)' \left(\widetilde{\mathbf{X}}'\widetilde{\mathbf{X}}\right) \left(\widehat{\boldsymbol{\beta}} - \widehat{\boldsymbol{\beta}}_{(i,j)}\right) (s^2 K)^{-1}$$

  - Influence exerted by a pair (i,j) on LS estimates *jointly*
  - Comparison of LS estimates *with* and *without* the *pair*
  - $\mathrm{C}_{ij}(\widehat{\boldsymbol{\beta}}) = \mathrm{C}_{ji}(\widehat{\boldsymbol{\beta}})$

- For $i = j$,

$$\mathrm{C}_{ii}(\widehat{\boldsymbol{\beta}}) = \left(\widehat{\boldsymbol{\beta}} - \widehat{\boldsymbol{\beta}}_{(i)}\right)' \left(\widetilde{\mathbf{X}}'\widetilde{\mathbf{X}}\right) \left(\widehat{\boldsymbol{\beta}} - \widehat{\boldsymbol{\beta}}_{(i)}\right) (s^2 K)^{-1}$$

- $\mathrm{C}_{ij}(\widehat{\boldsymbol{\beta}}), \mathrm{C}_{ii}(\widehat{\boldsymbol{\beta}}) \sim \mathrm{F}(\nu_1, \nu_2)$

  where $\nu_1 = k + 1$ and $\nu_2 = NT - N - (k + 1)$

Influence analysis: Joint influence

- For $i \neq j$,

$$\mathrm{C}_{ij}(\widehat{\boldsymbol{\beta}}) = \left(\widehat{\boldsymbol{\beta}} - \widehat{\boldsymbol{\beta}}_{(i,j)}\right)' \left(\widetilde{\mathbf{X}}'\widetilde{\mathbf{X}}\right) \left(\widehat{\boldsymbol{\beta}} - \widehat{\boldsymbol{\beta}}_{(i,j)}\right) (s^2 K)^{-1}$$

  - Influence exerted by a pair (i,j) on LS estimates *jointly*
  - Comparison of LS estimates *with* and *without* the *pair*
  - $\mathrm{C}_{ij}(\widehat{\boldsymbol{\beta}}) = \mathrm{C}_{ji}(\widehat{\boldsymbol{\beta}})$

- For $i = j$,

$$\mathrm{C}_{ii}(\widehat{\boldsymbol{\beta}}) = \left(\widehat{\boldsymbol{\beta}} - \widehat{\boldsymbol{\beta}}_{(i)}\right)' \left(\widetilde{\mathbf{X}}'\widetilde{\mathbf{X}}\right) \left(\widehat{\boldsymbol{\beta}} - \widehat{\boldsymbol{\beta}}_{(i)}\right) (s^2 K)^{-1}$$

- $\mathrm{C}_{ij}(\widehat{\boldsymbol{\beta}}), \mathrm{C}_{ii}(\widehat{\boldsymbol{\beta}}) \sim \mathrm{F}(\nu_1, \nu_2)$

  where $\nu_1 = k + 1$ and $\nu_2 = NT - N - (k + 1)$

Influence analysis: Joint effects

- For $i \neq j$,

$$\mathrm{K}_{j|i} = \frac{\mathrm{C}_{ij}(\widehat{\boldsymbol{\beta}})}{\mathrm{C}_{ii}(\widehat{\boldsymbol{\beta}})}$$

  - How much the pair is influential wrt i
  - For $i = j$, $\mathrm{K}_{j|i} = 1$

- For large values of $\mathrm{K}_{j|i}$
  - j *swamps* i
  - the most influential unit *swamps* the least
  - j drives the LS estimates *swamping* the effect of i

Influence analysis: Joint effects

- For $i \neq j$,

$$\mathrm{K}_{j|i} = \frac{\mathrm{C}_{ij}(\widehat{\boldsymbol{\beta}})}{\mathrm{C}_{ii}(\widehat{\boldsymbol{\beta}})}$$

  - How much the pair is influential wrt i
  - For $i = j$, $\mathrm{K}_{j|i} = 1$

- For large values of $\mathrm{K}_{j|i}$
  - j *swamps* i
  - the most influential unit *swamps* the least
  - j drives the LS estimates *swamping* the effect of i

Influence analysis: Conditional measure

For $i \neq j$,

$$\mathrm{C}_{i(j)}(\widehat{\boldsymbol{\beta}}) = \left(\widehat{\boldsymbol{\beta}}_{(i,j)} - \widehat{\boldsymbol{\beta}}_{(j)}\right)' \left(\sum_{\substack{i=1 \\ i \neq j}}^{N} \widetilde{\mathbf{X}}'_{i(j)} \widetilde{\mathbf{X}}_{i(j)}\right) \left(\widehat{\boldsymbol{\beta}}_{(i,j)} - \widehat{\boldsymbol{\beta}}_{(j)}\right) (s^2 K)^{-1}$$

- Influence exerted by i on LS estimates *conditional* on removing j from the sample
- How the absence of j affects the influence i on LS estimates
- $\mathrm{C}_{i(j)}(\widehat{\boldsymbol{\beta}}) = 0$ for $i = j$
- $\mathrm{C}_{i(j)}(\widehat{\boldsymbol{\beta}}) \neq \mathrm{C}_{j(i)}(\widehat{\boldsymbol{\beta}})$
- $\mathrm{C}_{i(j)}(\widehat{\boldsymbol{\beta}}) \sim F(\nu_1, \nu_2)$

Influence analysis: Conditional effects

- For $i \neq j$

$$\mathrm{M}_{i(j)} = \frac{\mathrm{C}_{i(j)}(\widehat{\boldsymbol{\beta}})}{\mathrm{C}_{ii}(\widehat{\boldsymbol{\beta}})}$$

  - How influence of i changes before and after the deletion of j

- If $\mathrm{M}_{i(j)} \geq 1$
  - j *masks* i
  - influence of i increases without j in the sample
- If $\mathrm{M}_{i(j)} < 1$
  - j *boosts* i
  - influence of i decreases without j in the sample

Influence analysis: Conditional effects

- For $i \neq j$

$$\mathrm{M}_{i(j)} = \frac{\mathrm{C}_{i(j)}(\widehat{\boldsymbol{\beta}})}{\mathrm{C}_{ii}(\widehat{\boldsymbol{\beta}})}$$

  - How influence of i changes before and after the deletion of j

- If $\mathrm{M}_{i(j)} \geq 1$
  - j *masks* i
  - influence of i increases without j in the sample

- If $\mathrm{M}_{i(j)} < 1$
  - j *boosts* i
  - influence of i decreases without j in the sample

Influence analysis: Conditional effects

- For $i \neq j$

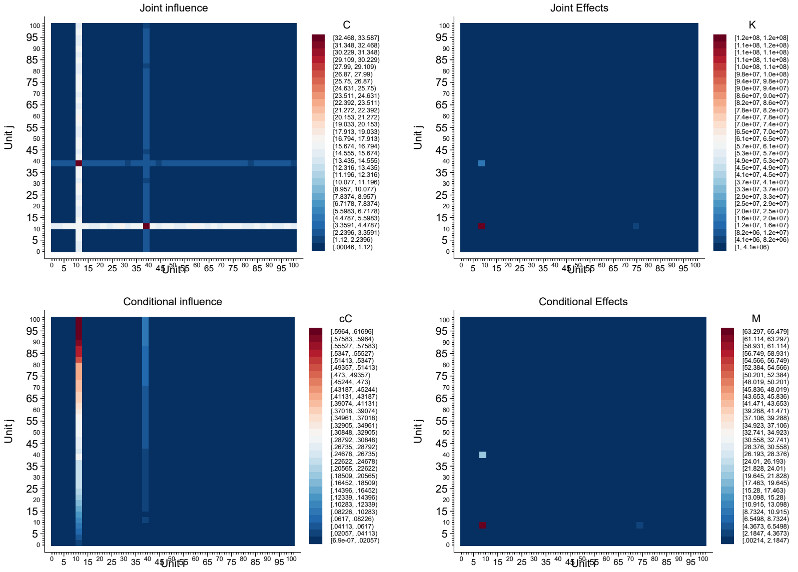
$$\mathrm{M}_{i(j)} = \frac{\mathrm{C}_{i(j)}(\widehat{\boldsymbol{\beta}})}{\mathrm{C}_{ii}(\widehat{\boldsymbol{\beta}})}$$

  - How influence of i changes before and after the deletion of j

- If $\mathrm{M}_{i(j)} \geq 1$
  - j *masks* i
  - influence of i increases without j in the sample

- If $\mathrm{M}_{i(j)} < 1$
  - j *boosts* i
  - influence of i decreases without j in the sample

Influence analysis: Network-like plots ▸ Summary



Example: Berka et al. (2018)

- They study relationship between real exchange rate and sectoral productivity in nine Eurozone countries

$$RER_{it} = \beta TFP_{it} + \mathbf{x}_{it}'\boldsymbol{\gamma} + \alpha_i + u_{it}$$

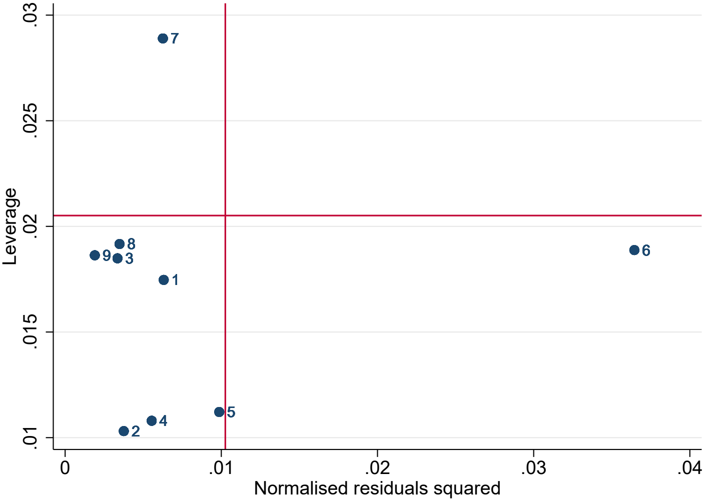
RER_{it}: real exchange rate in log
TFP_{it}: total factor productivity in log
$\mathbf{x}_{it}$: other controls
α_i: country fixed effects

- Finding strong correlation between TFP and RER among high-income countries with floating nominal exchange rates

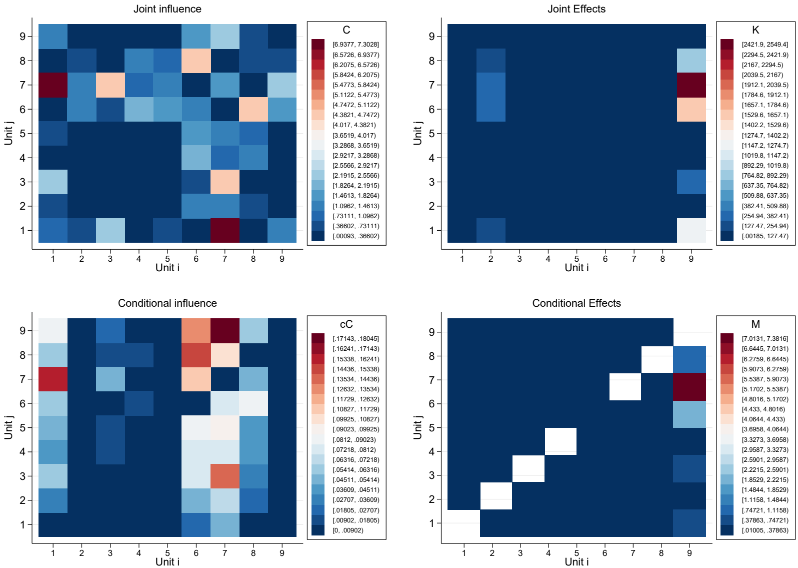
- Balanced panel with 9 countries between 1995–2007

- We use the sample for specification (2a) in Table 4

Example: Leverage-vs-residual plot



Example: Network-like plots ▸ Summary



Conclusion

- This paper develops a method to
  1. Identify anomalous units and their type (unit-wise leverage-vs-residual plot)
  2. Investigate how anomalous units may affect the LS estimates (joint and conditional influence and effects)

- Once identified the type of anomaly in the sample
  1. Methods for measurement error if error in the data entry
  2. Robust estimation techniques if VO and BL (Bramati and Croux, 2007; Verardi and Croux, 2009; Aquaro and Čížek, 2013, 2014; Jiao, 2022)
  3. Jackknife-type standard errors if GL (MacKinnon and White, 1985; Davidson et al., 1993; MacKinnon, 2013; Belotti and Peracchi, 2020; Polselli, 2022)

Thank you for your attention!

✉ annalivia.polselli[at]essex.ac.uk

https://github.com/POLSEAN/Influence-Analysis

@AnnalivPolselli

References I

Aquaro, M. and Čížek, P. (2013). One-step robust estimation of fixed-effects panel data models. *Computational Statistics & Data Analysis*, 57(1):536–548.

Aquaro, M. and Čížek, P. (2014). Robust estimation of dynamic fixed-effects panel data models. *Statistical Papers*, 55(1):169–186.

Atkinson, A. and Mulira, H.-M. (1993). The stalactite plot for the detection of multivariate outliers. *Statistics and Computing*, 3(1):27–35.

Atkinson, A. C. (1985). Plots, transformations and regression; an introduction to graphical methods of diagnostic regression analysis. No. 04; $\mathrm{QA278.\ 2,\ A8}$.

Banerjee, M. and Frees, E. W. (1997). Influence diagnostics for linear longitudinal models. *Journal of the American Statistical Association*, 92(439):999–1005.

Belotti, F. and Peracchi, F. (2020). Fast leave-one-out methods for inference, model selection, and diagnostic checking. *The Stata Journal*, 20(4):785–804.

Berka, M., Devereux, M. B., and Engel, C. (2018). Real exchange rates and sectoral productivity in the eurozone. *American Economic Review*, 108(6):1543–81.

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

Chatterjee, S. and Hadi, A. S. (1988). Impact of simultaneous omission of a variable and an observation on a linear regression equation. *Computational Statistics & Data Analysis*, 6(2):129–144.

Cook, R. D. (1979). Influential observations in linear regression. *Journal of the American Statistical Association*, 74(365):169–174.

References II

Davidson, R., MacKinnon, J. G., et al. (1993). Estimation and inference in econometrics. *OUP Catalogue*.

Donald, S. G. and Maddala, G. (1993). 24 identifying outliers and influential observations in econometric models. In *Econometrics*, volume 11 of *Handbook of Statistics*, pages 663 – 701. Elsevier.

Jiao, X. (2022). A simple robust procedure in instrumental variables regression. Unpublished, Last accessed: 07/02/2023.

Lawrance, A. (1995). Deletion influence and masking in regression. *Journal of the Royal Statistical Society: Series B (Methodological)*, 57(1):181–189.

MacKinnon, J. G. (2013). Thirty years of heteroskedasticity-robust inference. In *Recent advances and future directions in causality, prediction, and specification analysis*, pages 437–461. Springer.

MacKinnon, J. G. and White, H. (1985). Some heteroskedasticity-consistent covariance matrix estimators with improved finite sample properties. *Journal of econometrics*, 29(3):305–325.

Martín, N. (2015). Diagnostics in a simple correspondence analysis model: An approach based on cook's distance for log-linear models. *Journal of Multivariate Analysis*, 136:175–189.

Martín, N. and Pardo, L. (2009). On the asymptotic distribution of cook's distance in logistic regression models. *Journal of Applied Statistics*, 36(10):1119–1146.

References III

Pinho, L. G. B., Nobre, J. S., and Singer, J. M. (2015). Cook's distance for generalized linear mixed models. *Computational Statistics & Data Analysis*, 82:126–136.

Polselli, A. (2022). *Essays on Econometric Methods*. PhD thesis, University of Essex.

Rousseeuw, P. J. (1991). A diagnostic plot for regression outliers and leverage points. *Computational Statistics & Data Analysis*, 11(1):127–129.

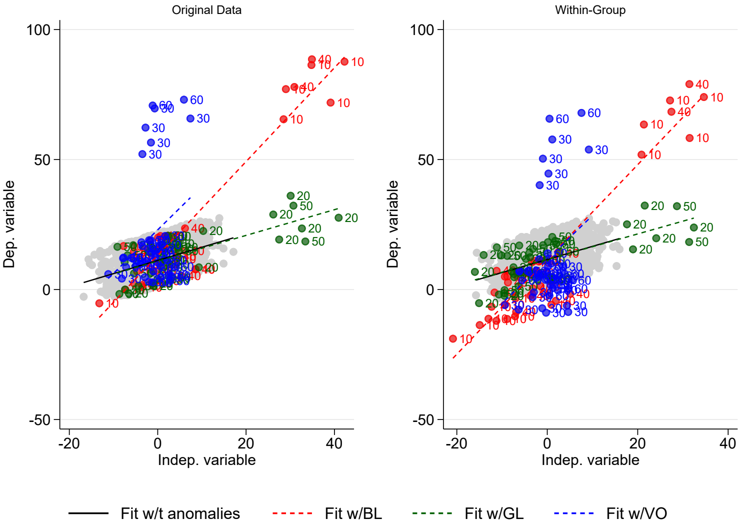
Rousseeuw, P. J. and Van Zomeren, B. C. (1990). Unmasking multivariate outliers and leverage points. *Journal of the American Statistical Association*, 85(411):633–639.

Silva, J. S. (2001). Influence diagnostics and estimation algorithms for powell's scls. *Journal of Business & Economic Statistics*, 19(1):55–62.

Verardi, V. and Croux, C. (2009). Robust regression in stata. *The Stata Journal*, 9(3):439–453.

Scatter Plot DGP

Back



DGP

Back

```
loc numobs 100
set obs 100
gen id = _n
expand 20

bys id: generate t = _n
bys id: gen z = rnormal(0,5)
**GL
bys id: replace z = z + rnormal(30,1) if id==20 & t<=5
bys id: replace z = z + rnormal(30,1) if id==50 & t<=2
**for BL
bys id: replace z = z + rnormal(30,1) if id==10 & t<=5
bys id: replace z = z + rnormal(30,1) if id==40 & t<=2
**line
bys id: gen a = runiform(0,20)
bys id: gen y = 1 + .5*z + a + runiform()
**BL
bys id: replace y = y + rnormal(50,1) if id==10 & t<=5
bys id: replace y = y + rnormal(50,1) if id==40 & t<=2
*VO
bys id: replace y = y + rnormal(50,1) if id==30 & t<=5
bys id: replace y = y + rnormal(50,1) if id==60 & t<=2
```

Influence Analysis: Summary ▸ Back

Variable	Obs	Mean	Std. dev.	Min	Max
C	10,000	.3811386	2.200585	2.35e-11	33.58732
K	10,000	16156.08	1242556	4.42e-08	1.23e+08
cC	10,000	.0038312	.0353837	0	.6169614
M	9,900	.0305928	.6922132	4.39e-06	65.47916

```
-------------------------------------------------
              Influence analysis
-------------------------------------------------
v1 = k+1 =  2
v2 = NT-N-k-1 = 1898
c1 = 4/N = .04
c2 = F(v1,v2,.5) =   0.6934
-------------------------------------------------
Cii >= c1
 - Count : 8
 - List  : 8 10 20 34 40 43 50 65
Cii >= c2
 - Count : 2
 - List  : 10 40
i with K >= p99
 - Count : 30
 - List  : 3 4 6 9 11 13 14 19 24 27 47 49 55 57 62 64 67 68 69 71 72 74 76 77 79 84 86 89 93 95
j with K >= p99
 - Count :
 - List  :
i with M >= 1
 - Count : 2
 - List  : 9 74
j with M >= 1
 - Count : 2
 - List  : 10 40
-------------------------------------------------
```

Example: Summary ▸ Back

Variable	Obs	Mean	Std. dev.	Min	Max
C	81	1.0233	1.472976	.0009253	7.30281
K	81	97.87085	368.2484	.0018538	2549.404
cC	81	.032125	.0439157	0	.1804506
M	72	.2303033	.8915019	.0046645	7.381636

```
---------------------------------------------------
                Influence analysis
---------------------------------------------------
v1 = k+1 =  2
v2 = NT-N-k-1 = 184
c1 = 4/N = .4444444444444444
c2 = F(v1,v2,.5) =   0.6958
---------------------------------------------------
Cii >= c1
 - Count : 4
 - List  : 1 6 7 8
Cii >= c2
 - Count : 3
 - List  : 1 6 7
i with K >= p99
 - Count : 1
 - List  : 9
j with K >= p99
 - Count :
 - List  :
i with M >= 1
 - Count : 1
 - List  : 9
j with M >= 1
 - Count : 2
 - List  : 6 7
---------------------------------------------------
```