

# Estimation and inference in cross-sectionally dependent panel data models

Citation for published version (APA):

Karabiyik, H. (2015). *Estimation and inference in cross-sectionally dependent panel data models*. [Doctoral Thesis, Maastricht University]. Datawyse / Universitaire Pers Maastricht. <https://doi.org/10.26481/dis.20150121hk>

## Document status and date:

Published: 01/01/2015

## DOI:

[10.26481/dis.20150121hk](https://doi.org/10.26481/dis.20150121hk)

## Document Version:

Publisher's PDF, also known as Version of record

## Please check the document version of this publication:

- A submitted manuscript is the version of the article upon submission and before peer-review. There can be important differences between the submitted version and the official published version of record. People interested in the research are advised to contact the author for the final version of the publication, or visit the DOI to the publisher's website.
- The final author version and the galley proof are versions of the publication after peer review.
- The final published version features the final layout of the paper including the volume, issue and page numbers.

[Link to publication](#)

## General rights

Copyright and moral rights for the publications made accessible in the public portal are retained by the authors and/or other copyright owners and it is a condition of accessing publications that users recognise and abide by the legal requirements associated with these rights.

- Users may download and print one copy of any publication from the public portal for the purpose of private study or research.
- You may not further distribute the material or use it for any profit-making activity or commercial gain
- You may freely distribute the URL identifying the publication in the public portal.

If the publication is distributed under the terms of Article 25fa of the Dutch Copyright Act, indicated by the "Taverne" license above, please follow below link for the End User Agreement:

[www.umlib.nl/taverne-license](http://www.umlib.nl/taverne-license)

## Take down policy

If you believe that this document breaches copyright please contact us at:

[repository@maastrichtuniversity.nl](mailto:repository@maastrichtuniversity.nl)

providing details and we will investigate your claim.

# Hande Karabiyik

## Nederlandse samenvatting

Paneldata is een verzameling van gegevens waarbij verschillende eenheden worden gevolgd gedurende een bepaalde tijdsperiode. De data in een panel bevat dus zowel een tijdsdimensie als ook een zogenaamde individuele dimensie. Het woord ‘individuele’ in de voorafgaande zin moet ietwat ruimer worden opgevat dan de naam aanvankelijk doet vermoeden. De individuele dimensie kan inderdaad bestaan uit gegevens betreffende specifieke individuen. Ze kan echter ook refereren naar verschillende bedrijven, industrietakken of zelf landen. Het bruto binnenlands product (BBP) zoals waargenomen voor enkele landen in de Europese Unie over de tijdsperiode 2000-2014 is een voorbeeld van paneldata. Recente analyse heeft aangetoond dat de individuele eenheden vaak afhankelijkheid vertonen. Deze afhankelijkheid kan verregaande gevolgen hebben, als ze wordt genegeerd in de modellering. Er is dan ook theoretisch onderzoek gedaan ter beantwoording van de vraag: hoe kan individuele afhankelijkheid in paneldata (het best) gemodelleerd worden? Gedurende de laatste jaren is de gerelateerde literatuur gestaag gegroeid. Toch is er nog steeds interesse in nieuwe methoden. Methoden die werken onder minder strenge aannamen, of statistische toetsen of te zien of deze aannamen waarschijnlijk zijn. Dit proefschrift levert een bijdrage aan de ontwikkeling van zulke vernieuwende methoden.

De statistische afhankelijkheid in de individuele dimensie kan veroorzaakt worden door het ontbreken van relevante variabelen in het model, de aanwezigheid van globale verstoringen, geografische effecten, etc. Stel dat er een variabele is die niet onafhankelijk is over de individuele eenheden. Het ontbreken van deze variabele in de modellering impliceert ogenblikkelijk dat

de individuele verstoringen in het model deze variabele zullen bevatten. Dit resulteert in de genoemde afhankelijkheid in de individuele dimensie. De aanwezigheid van globale verstoringen is vergelijkbaar. In dit geval is er a priori een verstoring die alle individuele eenheden beïnvloedt. Laten we terugkeren naar ons voorbeeld: het bbp van de landen in de Europese Unie. Volgens empirische studies is het aannemelijk dat het bruto binnenlands product wordt beïnvloed door zulke globale verstoringen. In andere woorden, het bbp van de verschillende landen in de Europese Unie is niet alleen een voorbeeld van paneldata, maar zelfs een voorbeeld van paneldata met afhankelijk in de individuele dimensie.

De globale verstoringen kunnen gemodelleerd worden door middel van factor modellen. De achterliggende gedachte is dat een klein aantal niet waargenomen variabelen (de factoren) alle eenheden in de individuele dimensie gezamenlijk beïnvloed. Het feit dat deze factoren niet worden waargenomen impliceert dat deze alleen benaderd kunnen worden. Twee populaire benaderingen zijn de *Principale Componenten (PC)* benadering en de *Common Correlated Effects (CCE)* benadering. De *CCE* approximatie is voorgesteld in Pesaran (2006) voor stationaire factoren en stationaire variabelen. De niet geobserveerde factoren worden benaderd door gemiddelden over de individuele eenheden. Deze methode is toepasbaar onder verschillende aannamen. Een voordeel van de *CCE* methode is de eenvoud: gemiddelden volstaan.

De *CCE* approximatie staat centraal in dit proefschrift. Hoofdstuk 1, de introductie, geeft meer informatie over deze benadering en behandelt de concepten die herhaaldelijk gebruikt zullen worden. Daarnaast informeert deze introductie de lezer over de doelstelling en de motivatie onderliggende dit proefschrift. De wetenschappelijke bijdragen die dit proefschrift levert worden ook vermeld.

Hoofdstuk 2 van dit proefschrift bevat de eerste discussie van paneldata met afhankelijkheid in de individuele dimensie. Het vertrekpunt van deze discussie is n van de aannamen die cruciaal is in Pesaran (2006). Deze aanname vereist dat het aantal niet waargenomen factoren niet groter mag zijn dan het aantal geobserveerde variabelen. Deze veronderstelling is noodzakelijk voor de consistente en efficiënte schatting van het model. We stellen voor niet alleen gemiddelden te nemen in de individuele dimensie, maar ook andere combinaties van de waargenomen variabelen toe te staan. Het zal blijken dat de genoemde aanname door deze modificatie overbodig is.

In Hoofdstuk 3 beperken we ons tot de gemiddelden over de individuele eenheden. We beschouwen niet langer andere combinaties van de variabelen. De uitbreiding en uitdaging is nu de toepassing op een specifiek model, namelijk het *Error Correction Model (ECM)*. Een extra aanname, *weak exogeneity*, maakt het mogelijk ons zonder verlies van efficiëntie te concentreren op het *Conditional Error Correction Model (CECM)*. In dit *CECM* laten we vervolgens afhankelijkheid toe in de individuele dimensie. Om met deze afhankelijkheid te kunnen omgaan verrijken we het model met de huidige en vertraagde eerste verschillen van de individuele dimensie gemiddelden. Zowel de lange als ook de korte termijn parameters kunnen nu consistent geschat worden met de (indirecte) kleinste kwadraten methode. We zullen ook aantonen dat een andere consistente schatter van de lange termijn parameters verkregen kan worden door de lange termijn parameters voor iedere individuele eenheid afzonderlijk te schatten, en de resultaten te middelen.

Het gebruik van het *CECM* was gebaseerd op de *weak exogeneity* aanname. In Hoofdstuk 4 beschouwen we Lagrange-Multiplier toetsen om te testen voor *weak exogeneity* van sommige variabelen voor de lange termijn parameters. Allereerst leiden we een test proced af dat geldig is voor de

individuele eenheden afzonderlijk. Hierna construeren we uit de toetsingsgrootheden voor de individuele eenheden nieuwe toetsingsgrootheden die de testen voor de *weak exogeneity* van alle individuele eenheden gezamenlijk. De toegestane afhankelijkheid van de individuele eenheden leidt echter ook tot afhankelijkheid van de individuele toetsingsgrootheden. De standaard methoden die gebruikt worden om individuele toetsen te combineren tot een gezamenlijke toets zijn niet langer direct toepasbaar. We beredeneren dat als gevolg hiervan bootstrap methoden noodzakelijk zijn.

Dit proefschrift vormt hopelijk een waardevolle bijdrage tot de literatuur over paneldata met afhankelijkheid in de individuele dimensie. Een literatuur waarin vele vragen nog onbeantwoord zijn. We hopen dat zowel de theoretische als ook de toegepaste onderzoeker zal profiteren van de inhoud van dit proefschrift. De theoreticus ontdekt wellicht vernieuwende inzichten, of zelfs een opstapje voor een geheel nieuw eigen onderzoek. De toegepaste onderzoeker zou zich kunnen verheugen op een nieuw stukje gereedschap om paneldata te onderzoeken. Echter geen gewoon stukje gereedschap, maar een stukje gereedschap dat expliciet rekening houdt met mogelijke afhankelijkheid in de individuele dimensie.