



UvA-DARE (Digital Academic Repository)

Responses to the incidental parameter problem

Pua, A.A.Y.

Publication date

2016

Document Version

Final published version

[Link to publication](#)

Citation for published version (APA):

Pua, A. A. Y. (2016). *Responses to the incidental parameter problem*.

General rights

It is not permitted to download or to forward/distribute the text or part of it without the consent of the author(s) and/or copyright holder(s), other than for strictly personal, individual use, unless the work is under an open content license (like Creative Commons).

Disclaimer/Complaints regulations

If you believe that digital publication of certain material infringes any of your rights or (privacy) interests, please let the Library know, stating your reasons. In case of a legitimate complaint, the Library will make the material inaccessible and/or remove it from the website. Please Ask the Library: <https://uba.uva.nl/en/contact>, or a letter to: Library of the University of Amsterdam, Secretariat, Singel 425, 1012 WP Amsterdam, The Netherlands. You will be contacted as soon as possible.

Nederlandse Samenvatting

(Summary in Dutch)

Dit proefschrift bestaat uit een aantal artikelen rond een gezamenlijk thema: het incidentele-parameterprobleem in paneldata modellen. De laatste decennia is er een grote toename in de beschikbaarheid van datasets waarbij een groep individuen, bedrijven of landen over een korte of lange periode worden waargenomen. Dit type data stelt ons in de gelegenheid om dynamisch gedrag te analyseren, waarbij rekening wordt gehouden met tijdsinvariante niet-waargenomen heterogeniteit. Doorgaans heeft deze heterogeniteit de vorm van individu-specifieke effecten, welke een probleem opleveren voor identificatie, schatting en toetsing van parameters, in het bijzonder wanneer standaardmethoden zonder aanpassing worden toegepast op panel data.

In Hoofdstuk 1 wordt een inleiding gegeven op de belangrijkste ontwikkelingen in de paneldata literatuur over de afgelopen decennia, voor zover die relevant zijn voor de motivatie voor de overige hoofdstukken van het proefschrift. Hoofdstukken 2 tot en met 5 bevatten mijn bijdragen aan de paneldata literatuur.

Hoofdstuk 2 laat zien dat het benaderen van gemiddelde marginale effecten in discrete-keuzemodellen op basis van het lineaire kansmodel – een gebruikelijke empirische aanpak in toptijdschriften – zeer onverstandig is in paneldata situaties. Afgezien van het feit dat het gemiddelde marginale effect soms niet uniek geïdentificeerd is, blijkt de instrumentele-variabelenmethode toegepast op een dynamisch lineair kansmodel te leiden tot een inconsistente schatter van het gemiddelde marginale effect, ongeacht welk type asymptotische benadering wordt gebruikt (grote cross-sectie of tijdreeksdimensie).

In Hoofdstuk 3 wordt een paneldata schattingsmethode ontwikkeld voor een simultane-vergelijkingenmodel met discrete endogene variabelen, niet-waargenomen heterogeniteit en dynamiek. Dit type model komt geregeld voor in empirische toepassingen, maar tot nu toe is er nog geen alomvattende econometrische aanpak voor ontwikkeld. De nieuw ontwikkelde methode wordt toegepast op een model voor de interactie van liquiditeits- en hoeveelheidseffecten op het arbeidsaanbod van man-

nelijke hoofden van de huishouding in de *Panel Study of Income Dynamics*. Daarbij worden de resultaten op basis van de nieuwe methode vergeleken met bestaande empirische resultaten.

In Hoofdstuk 4 wordt een methode ontwikkeld om de vertekening in de schatting van parameters in paneldata modellen te corrigeren op basis van orthogonale projecties. De voorgestelde methode is gebaseerd op een gecorrigeerde scorevector, verkregen door de scorevector voor de structurele parameters te projecteren op het orthogonale complement van een ruimte opgespannen door functies die de fluctuatie in incidentele parameters karakteriseren. Onder de aanname dat de individuspecifieke effecten vrijwel elke eindige waarde kunnen aannemen, en dat de kansdichtheden van de data correct gespecificeerd zijn, wordt afgeleid dat de asymptotische verdeling van de parameterschatters normaal is, gecentreerd rond de werkelijke waarde. Dit resultaat correspondeert met dat van bestaande biascorrectiemethoden in de literatuur. De constructie van de gecorrigeerde scorevector kan worden uitgebreid naar situaties waarin de niet-waargenomen heterogeniteit meerdimensionaal is. Monte Carlo simulaties laten zien dat de eindige-steekprofeigenschappen van de nieuwe methode minstens even goed, en in sommige gevallen beter zijn dan die van bestaande methoden, in het bijzonder als er slechts 3 of 4 tijdswaarnemingen zijn.

In Hoofdstuk 5 verken ik de parallellen tussen paneldata methoden die individuspecifieke niet-waargenomen heterogeniteit toelaten, en *big data* methoden gebaseerd op grote hoeveelheden gegevens met een relatief lage informatiewaarde, voor het verkrijgen van informatie over laag-dimensionale parameters. De vraag is hoe *machine learning* methoden zoals Lasso consistente schatters van de structurele parameters kunnen opleveren in lineaire dynamische paneldata modellen met een vast (en klein) aantal tijdswaarnemingen, als we bereid zijn om aan te nemen dat de individuele effecten *sparse* zijn. De resultaten van dit hoofdstuk geven aan dat sterke aannames nodig zijn wat betreft de omvang en het aantal individuele effecten ongelijk aan nul, om consistente schatting en betrouwbare inferentie over de structurele parameters mogelijk te maken.

Elk van de hoofdstukken passen in de onderzoeksagenda zoals hiervoor omschreven. Mijn vervolgonderzoek zal zich richten op de volgende ideeën. In aansluiting op Hoofdstuk 2 is verder onderzoek noodzakelijk in welke situaties het lineaire kansmodel wel of niet werkt. Het ontwikkelen van een identificatieaanpak en van schattings- en toetstingsprocedures aansluitend op de methode zoals ontwikkeld in Hoofdstuk 3 zal van pas komen bij toekomstig empirisch werk dat niet gebaseerd is op parametrische restricties. De tweede-orde orthogonale projectie van Hoofdstuk 4 kan worden uitgebreid naar hogere ordes; het is van belang te onderzoeken of dit uiteindelijk convergeert naar de scorevector van een conditionele of marginale likelihood (als deze bestaan), of naar een andere functie van slechts de structurele parameters. De aanpak in Hoofdstuk 5, tenslotte, kan worden uitgebreid naar niet-lineaire panel data modellen.