

Commentaar op de WODC-schatting van de export van in Nederland geteelde cannabis

Nicole Maalsté heeft mij gevraagd om in de hoedanigheid van methodoloog en statisticus de gebruikte schattingsmethoden voor de export van de in Nederland geteelde cannabis te evalueren, als ook de daaruit vloeiende schattingen, zoals beschreven in Van der Giessen, Moolenaar & Van Ooyen-Houben (2014), hierna aangeduid als het WODC-rapport. Bij mijn evaluatie heb ik naast mijn eigen expertise gebruik gemaakt van de volgende bronnen; paragraaf 6.1, hoofdstukken 7 en 8, en bijlagen 3 en 4 uit het WODC-rapport, en een pdf-bestand van het Centraal Planbureau (CPB) gedownload van <http://www.cpb.nl/veelgestelde-vragen-ramingen-cpb>.

Mijn evaluatie hieronder zal bestaan uit drie paragrafen. In de eerste paragraaf beschrijf ik hoe het Centraal Planbureau macro-economische grootheden schat. In de tweede paragraaf ga ik in op hoe de export van cannabis wordt geschat. De vergelijking met de schattingsmethoden van het Centraal Planbureau helpen mij duidelijk te maken dat de gebruikte schattingsmethoden en de daaruit vloeiende schattingen in het WODC-rapport niet valide zijn. Ik zal me in de tweede paragraaf in hoofdzaak richten op de validiteit van de gerapporteerde “95% betrouwbaarheidsintervallen” (8.2.1), maar zal ook kort iets zeggen over de gevoeligheidsanalyses (7.2) en de analyses van het effect van het bereik van de exogenen (7.3). In de laatste paragraaf vat ik de belangrijkste conclusies van mijn evaluatie van betrouwbaarheidsintervallen samen.

1. Schattingen van macro-economische grootheden door het Centraal Planbureau

Eén van de vele schattingen die het CPB maakt is de korte-termijnraming van het bruto binnenlands product. Hoe maakt het CPB deze schatting? Allereerst is een model nodig met daarin de exogenen die van invloed zijn op de endogeen. De endogeen is de variabele die voorspeld wordt met het model, hier het binnenlands bruto product. De exogenen zijn de variabelen die in het model worden gebruikt om de endogeen te voorspellen; een voorbeeld van een exogeen is hier de buitenlandse groei. Belangrijk daarbij is dat het model de samenhangen tussen de exogenen en het bbp goed weergeeft. Het model en de waarden van de exogenen zijn gebaseerd op (inter)nationale data. Ook worden, in ieder geval sinds 1971, de schattingen systematisch vergeleken met latere feitelijk gerealiseerde ontwikkelingen (ook data). Natuurlijk treden er voorspelfouten op. Deze hebben verschillende oorzaken, bijvoorbeeld (i) het model beschrijft de samenhangen tussen de exogenen en het binnenlands bruto product niet goed, (ii) de waarden van de exogenen blijken verkeerd ingeschat, en (iii) de ‘situatie’ of context is veranderd. Het CPB bouwt onzekerheidsmarges in haar ramingen, vanwege deze voorspelfouten. Het berekenen van de onzekerheidsmarges is ingewikkeld, waarbij de uitkomsten afhankelijk zijn van de gemaakte veronderstellingen. De berekende onzekerheidsmarges of betrouwbaarheidsintervallen, ook zal zijn ze gebaseerd op veel data en gevalideerde modellen, zijn niet automatisch representatief voor de toekomst. Mede daarom publiceert het CPB alleen de puntschattingen en niet de onzekerheidsmarges.

2. Schattingen van export cannabis

Bijlage 3 en Bijlage 4 van het WODC-rapport geven respectievelijk een overzicht van de exogenen en de modellen die de samenhang weergeven tussen de exogenen en de endogenen. Een voorbeeld van een

endogeen in het WODC-rapport is het percentage export van in Nederland geteelde cannabis. Andere voorbeelden van endogenen zijn de totale productie van in Nederland geteelde cannabis en de totale consumptie van in Nederland geteelde cannabis. Voorbeelden van de exogenen, waarvan er meer dan 40 zijn, zijn het aantal hennepplanten per m², de gemiddelde opbrengst per plant en het marktaandeel van coffeeshops. Aangezien ik geen expertise heb op het terrein van cannabisproductie en export, neem ik in het vervolg voor het gemak aan dat alle exogenen en hun samenhangen met de endogenen juist zijn bepaald. Ik richt me hier op de waarden of verdelingen van de exogenen en de onzekerheden hierin, en de gevolgen die dat heeft voor de validiteit van de schattingen van de betrouwbaarheidsintervallen.

De kern van het probleem ligt in het feit dat voor het bepalen van de waarden van een belangrijk aantal van de exogenen geen of nauwelijks data voorhanden zijn, data die wel voorhanden zijn als het CPB haar ramingen maakt. De in paragraaf 8.2.1 gerapporteerde betrouwbaarheidsintervallen van het WODC-rapport kunnen alleen valide zijn als de mogelijke waarden van de exogenen goed zijn geschat of gespecificeerd. Meer precies wordt hier bedoeld de meest waarschijnlijke waarde, en de onzekerheid hiervan. Maar ook de samenhang tussen de exogenen. Het ligt voor de hand dat een mogelijke onderschatting van een exogeen, bijvoorbeeld het marktaandeel van coffeeshops, samengaat met een onderschatting van andere exogenen, zoals de totale omzet van de coffeeshops in Nederland. Om al deze mogelijke waarden goed te bepalen zijn data nodig. Door de afwezigheid van goede data en de aannames in de gedane Monte Carlo simulaties treden de volgende vertekeningen op in de schattingen van de betrouwbaarheidsintervallen:

- (i) Onjuiste schattingen van de gemiddelde waarde van exogenen leiden tot een systematische vertekening in de schatting van het 'centrum' van het betrouwbaarheidsinterval.
- (ii) Onjuiste schattingen van de spreiding van de exogenen leiden tot een systematische vertekening in de schatting van de 'breedte' van het betrouwbaarheidsinterval.
- (iii) De Monte Carlo simulaties gaan uit van onafhankelijk verdeelde exogenen, dat wil zeggen dat voor elke exogeen apart een willekeurige waarde wordt getrokken. Het is echter waarschijnlijk dat de exogenen positief met elkaar samenhangen, zoals eerder genoemd het voorbeeld van de totale omzet van coffeeshops en het marktaandeel van coffeeshops. Gevolg hiervan is dat het betrouwbaarheidsinterval dat werd verkregen door Monte Carlo simulaties aanzienlijk smaller zal zijn dan op basis van deze positieve samenhangen verwacht mag worden.

Door deze vertekeningen is het onmogelijk een gefundeerde uitspraak te doen over de kans dat de gerapporteerde betrouwbaarheidsintervallen (zie figuren 11-14 van paragraaf 8.2) de werkelijke waarden bevatten. Het is zeer wel mogelijk dat geen van de vier gerapporteerde betrouwbaarheidsintervallen de werkelijke waarde bevat, ook omdat deze intervallen waarschijnlijk te smal zijn (zie (iii) hierboven). De afwezigheid van voldoende data maken goed schatten niet alleen ondoenlijk, het maakt ook het toetsen van de voorspelling onmogelijk. Merk op dat de situatie geheel anders is bij ramingen van het CPB, waar voorspellingen, data, en toetsing hand in hand gaan. Maar zelfs in die gunstige situatie maakt het CPB geen onzekerheidsmarges openbaar.

Hoe moeten de gerapporteerde 95% betrouwbaarheidsintervallen dan wel worden geïnterpreteerd? Zij geven aan dat 95% van de *door de modellen voorspelde waarden* worden omvat door deze intervallen. De intervallen zijn alleen valide met betrekking tot *mogelijke werkelijke waarden* als zowel het model als de mogelijke waarden van de exogenen goed is gespecificeerd. Dat laatste is niet het geval, of onmogelijk, door het ontbreken van goede data. Dus zijn de betrouwbaarheidsintervallen niet valide voor mogelijke werkelijke waarden.

In het WODC-rapport worden ook gevoeligheidsanalyses uitgevoerd (paragraaf 7.2) en het effect geanalyseerd van het bereik van de exogenen (paragraaf 7.3). In zowel de gevoeligheidsanalyses als de

bereikanalyses wordt gekeken wat het effect is van een exogeen op een endogeen, door de waarde van één exogeen te veranderen terwijl alle andere exogenen constant worden gehouden. Als kanttekening bij deze werkwijze wil ik opmerken dat de modellen rekenkundige producten (in de statistiek worden deze producten 'interacties' genoemd) van exogenen bevatten (zie Bijlage 4). Bijvoorbeeld, productie cannabis per m² is een product van de twee exogenen 'planten per m²' en 'gemiddelde opbrengst per plant'. Het gevolg daarvan is dat het effect van een exogeen afhangt van de waarde van andere exogenen, en dus niet zonder meer los van de deze andere exogenen kan worden bepaald. Bijvoorbeeld: Stel dat 'planten per m²' wordt ingeschat op 10, met minimum 7 en maximum 13, en 'gemiddelde opbrengst per plant' op 100 met minimum 50 en maximum 150. Volgens de werkwijze in het WODC-rapport zou het bereik van 'gemiddelde opbrengst per plant' dan worden vastgesteld op $(150-50) \times 10 = 1.000$, ofwel het domein van gemiddelde opbrengst per plant keer ingeschatte aantal planten per m². Echter, het bereik van 'gemiddelde opbrengst per plant' kan ook $(150-50) \times 8 = 800$ zijn, als het aantal planten per m² lager is. Kortom, in modellen met producten is het vaststellen van 'het' bereik van een exogeen strikt genomen niet mogelijk: het bereik hangt af van de waarde van andere exogenen. Dit geldt in het bijzonder voor de modellen in Bijlage 4, die producten kunnen bevatten van een tiental exogenen (zie vergelijking 36' in Bijlage B4.1 voor het voorspellen van de productie van cannabis).

3. Conclusie

In het onderhavige commentaar evalueer ik de gebruikte methoden om de export van de in Nederland geteelde cannabis te schatten, als ook de daaruit vloeiende schattingen. De kwaliteit van schattingen staat of valt met de kwaliteit van de gebruikte data. Een 95% betrouwbaarheidsinterval is alleen dan valide als zowel model als de mogelijke waarden van de exogenen juist zijn gespecificeerd op basis van inschattingen en data. Zelfs in een relatieve gunstige situatie waar veel data voorhanden zijn en het model veelvuldig is getoetst, zoals bij ramingen door het CPB, zal het '95% betrouwbaarheidsinterval' niet, zoals het hoort gemiddeld 19 van de 20 keer de werkelijke waarde bevatten. Dit omdat de modellen zijn gebaseerd op data en inschattingen uit het verleden, die niet automatisch representatief zijn voor de toekomst. In een ongunstige situatie waar weinig data voorhanden zijn, zoals bij de schatting van de export van de in Nederland geteelde cannabis, is de kans onbekend dat het betrouwbaarheidsinterval de werkelijke waarde bevat. Het betrouwbaarheidsinterval is dan louter een *voorspellingsinterval*, zonder duidelijke relatie met de realiteit. Na bestudering van de schattingsmethode (modellen, specificatie exogenen en hun gezamenlijke verdeling, aannames Monte Carlo simulatie) concludeer ik dan ook dat de gerapporteerde betrouwbaarheidsintervallen niet valide zijn.