

KU LEUVEN

FACULTY OF ECONOMICS
AND BUSINESS

S&L/DA/2017/117

AN EVALUATION OF THE METHODS TO FORECAST TAX
REVENUES
FINAL REPORT

ANDRÉ DECOSTER
TOON VANHEUKELOM
STIJN VAN HOUTVEN

DEPARTMENT OF ECONOMICS – KU LEUVEN

Version of 30 September 2019

EXECUTIVE SUMMARY

ANDRÉ DECOSTER
STIJN VAN HOUTVEN

In juni van elk jaar stelt de Federale Overheidsdienst Financiën (FOD Financiën) ramingen op van de verwachte fiscale ontvangsten in het volgende jaar. Op basis van deze raming wordt in september de initiële begroting opgesteld die wordt gestemd in het parlement voor het einde van december, i.e. de begrotingsopmaak. In februari raamt de FOD Financiën eveneens de belastingontvangsten van het lopende jaar. Deze ramingen vormen de basis voor een eventuele aanpassing van de begroting, de begrotingscontrole. Deze studie bespreekt beide ramingen van de fiscale ontvangsten door de FOD Financiën. Hieronder vallen niet alleen de ontvangsten die voor het federale niveau bestemd zijn, maar ook de ontvangsten die de federale overheid doorstort naar de sociale zekerheid, de gewesten en de gemeenschappen.

In 2015 ontstond er twijfel over de kwaliteit van die ramingen. De ramingen werden vaak grondig herzien tijdens de begrotingscontrole. Daarom werd op 22 oktober 2016 een taskforce opgericht met als taak de ramingen en de ramingsmethoden te analyseren. Naar aanleiding van haar taak binnen die taskforce besloot de FOD Financiën om een studie uit te schrijven omtrent de ramingen van de fiscale ontvangsten. In deze tekst vatten we de belangrijkste bevindingen van deze studie samen.

De uitgevoerde studie valt uiteen in vijf delen. Het eerste deel betrof een internationale vergelijking van de gebruikte ramingsmethoden. Vervolgens werd de gedesaggregeerde methode in detail beschreven en besproken. Deze methode wordt gebruikt bij de raming van ongeveer 85% van de belastingontvangsten, m.n. voor de personenbelasting, de vennootschapsbelasting, de belasting op toegevoegde waarde (BTW) en de accijnzen. Het derde deel was de kwantitatieve evaluatie van de ramingen die worden opgesteld op basis van deze gedesaggregeerde methode. Daarna werden op analoge wijze de ramingen buiten de gedesaggregeerde methode geëvalueerd, zowel op vlak van methode, als op kwantitatieve performantie. In het vijfde en laatste deel van het rapport werden de bevindingen samengevat en werden een aantal mogelijke onderzoekspistes voor verbetering aangereikt.

Deze samenvatting is gestructureerd rond een aantal typische vragen die gesteld worden over de ramingen, en het antwoord dat uit de analyse naar voren komt. We geven een antwoord op de volgende vragen:

- Welke fiscale ontvangsten worden geraamd, en met welke methode?
- Zijn er grote afwijkingen tussen de ramingen van fiscale ontvangsten en de realisaties?
- Houden de ramingen een systematische overschatting in van de werkelijke ontvangsten?
- Kunnen we de afwijkingen in de ramingen verklaren?
- Hoe worden de ramingen opgesteld?
- Is deze methode vergelijkbaar met de methode gehanteerd in de buurlanden?
- Zijn er verbeteringen mogelijk?

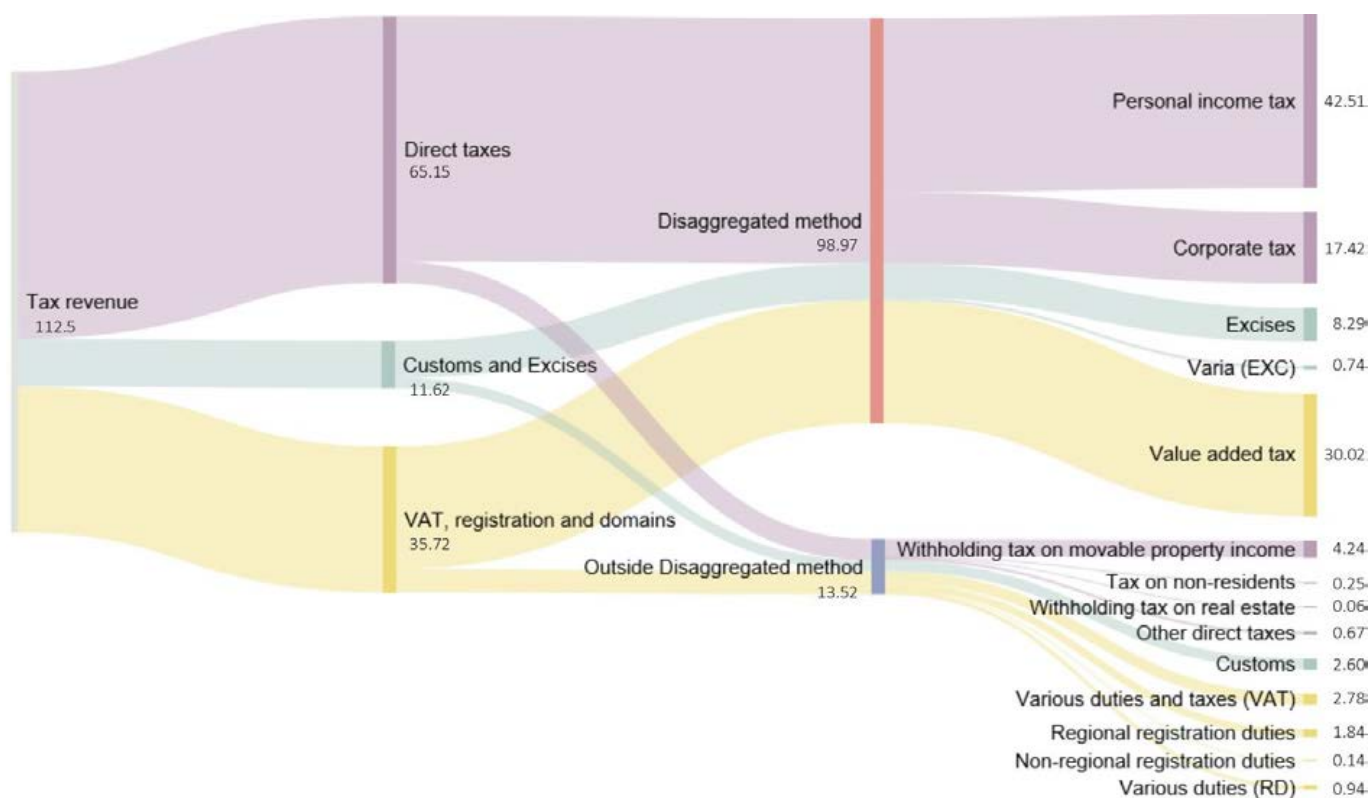
Voor een gedetailleerde beschrijving van de ramingsmethoden en van de analyses die werden uitgevoerd verwijzen we naar het eindrapport van de studie.¹

WELKE FISCALE ONTVANGSTEN WORDEN GERAAMD, EN MET WELKE METHODE?

De eerste vraag betreft het onderwerp van deze studie: de fiscale ontvangsten. Specifieker, welke verschillende categorieën ontvangsten worden beschouwd als deel van de fiscale ontvangsten?

Figuur 1 toont de samenstelling van de totale fiscale ontvangsten die in 2017 samen 112,5 miljard euro bedragen. De eerste opdeling die wordt gemaakt in de begrotingsdocumenten komt overeen met de drie categorieën die we in het middelste deel van de figuur weergeven, nl. de ontvangsten uit de directe belastingen, de ontvangsten uit douane en accijnzen, en de ontvangsten uit BTW, registratie en domeinen. Maar een belangrijkere opdeling van de ontvangsten tonen we aan de rechterzijde van de figuur. Het is de weergave van de gebruikte methode, waarbij we het verschil maken tussen de ramingen met de gedesaggregeerde methode en ramingen buiten de gedesaggregeerde methode. In 2017 werd 99 miljard euro opgehaald uit categorieën geraamd met de gedesaggregeerde methode. Buiten die methode bedroegen de ontvangsten 13,5 miljard euro, wat overeenkomt met slechts 12% van de totale fiscale ontvangsten.

Figuur 1: de samenstelling van de fiscale ontvangsten in 2017 (in miljard euro)



¹ Het eindrapport bestaat uit de verzameling van alle deliverables die werden opgeleverd in het kader van de studie. Dat zijn: D1: International comparative study on forecasting tax revenue in the five neighbouring countries, D2: Description of the disaggregated method, D3: Quantitative evaluation of the disaggregated method, D4: Analysis of the forecasts outside the disaggregated method, D5: Possible improvements in the forecast of tax revenues.

Die gedesaggregeerde methode werd ingevoerd in 1989. Voordien werden de totale ontvangsten samen geraamd door de toepassing van één groeivoet die werd berekend op basis van de voorspelde groei van het bruto nationale inkomen. Sinds 1990 worden de ontvangsten op een gedesaggregeerd niveau geraamd: er wordt een afzonderlijke voorspelling opgesteld voor de verschillende categorieën binnen de fiscale ontvangsten. Om die reden werd deze nieuwe methode de ‘gedesaggregeerde methode’ genoemd. Vandaag wordt de naam vooral gebruikt om de gehanteerde methode in de raming van de belangrijkste ontvangstencategorieën te karakteriseren. Ramingen op basis van de gedesaggregeerde methode vertrekken van macro-economische voorspellingen. De ramingen die niet worden gekenmerkt door dergelijke aanpak worden dan ‘buiten de gedesaggregeerde methode’ geplaatst, ook al zijn deze ramingen buiten de gedesaggregeerde methode vaak van een meer gedesaggregeerde aard dan deze binnen die methode.

Voor de directe belastingen wordt de gedesaggregeerde methode gebruikt voor de raming van de opbrengst van de personenbelasting en de vennootschapsbelasting. Een klein deel van de directe belastingen wordt geraamd buiten de gedesaggregeerde methode, zoals bijvoorbeeld de roerende voorheffing (*withholding tax on movable property income*).

Binnen douane en accijnzen worden de accijnzen geraamd met de gedesaggregeerde methode, samen met een aantal kleine ontvangstenstromen die nauw verbonden zijn aan de accijnzen. Ook al is de raming van de douaneontvangsten afhankelijk van een macro-economische projectie, nl. de groeivoet van import, ze maken formeel geen deel uit van de ontvangsten binnen de gedesaggregeerde methode.

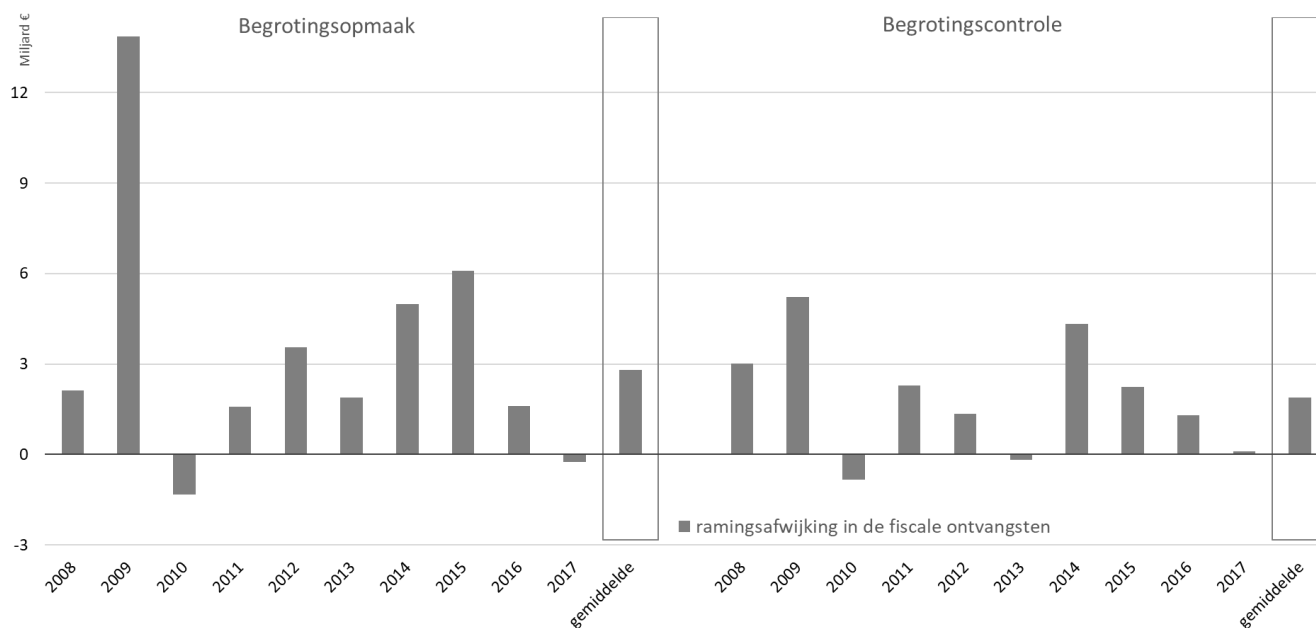
De BTW is een belangrijke bron van ontvangsten die wordt geraamd met de gedesaggregeerde methode. De overige ontvangsten in de categorie registratie en domeinen worden alle voorspeld buiten de gedesaggregeerde methode. Het gaat hierbij voornamelijk om registratie- en successierechten.

ZIJN ER GROTE AFWIJINGEN TUSSEN DE RAMINGEN VAN DE FISCALE ONTVANGSTEN EN DE REALISATIES?

Er zijn inderdaad afwijkingen tussen de ramingen van de fiscale ontvangsten en de realisaties. De afwezigheid van dergelijke afwijkingen zou eerder verbazing wekken. Figuur 2 toont deze afwijkingen. Hierbij werden de ramingsafwijkingen gedefinieerd als het verschil tussen de raming en de realisatie. Dit betekent dat positieve afwijkingen duiden op een raming die de ontvangsten overschatte, en vice versa. Het linkse paneel in de figuur toont deze afwijkingen voor de initiële begrotingen (begrotingsopmaak), en het rechtse paneel toont de afwijkingen die volgen uit de ramingen gemaakt tijdens de begrotingscontrole. De laatste omkaderde balk in elk paneel geeft de gemiddelde afwijking over de getoonde periode, in dit geval 2008-2017.²

In het linkse paneel van de figuur zien we dat de gemiddelde ramingsafwijking in de begrotingsopmaak 2,8 miljard bedraagt. De afwijking in 2009, het jaar van de financiële crisis, is duidelijk een outlier: in dit jaar werden de ontvangsten overschat met 13,8 miljard euro in de begrotingsopmaak.

² In de studie werden verschillende analyses uitgevoerd op de periode 1990-2017. We hebben ervoor gekozen om in de executive summary vooral te kijken naar de periode 2008-2017.

Figuur 2: de afwijkingen in de ramingen van de fiscale ontvangsten (in miljard €)

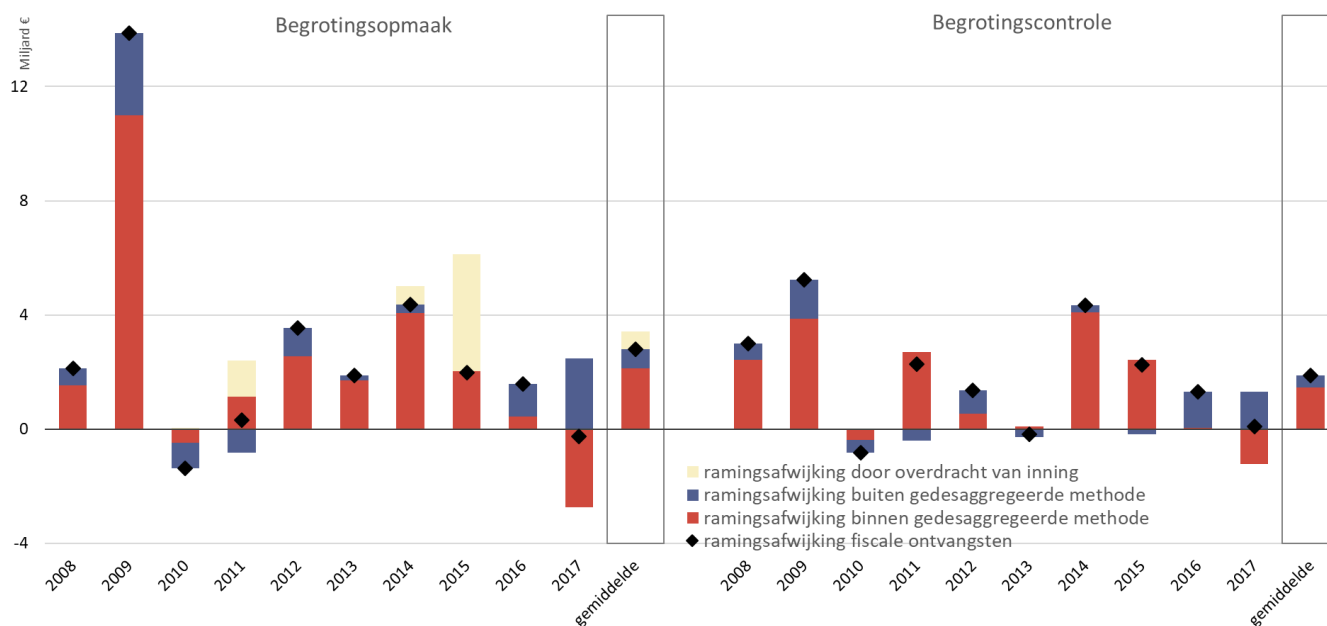
De afwijkingen in de begrotingscontrole zijn meestal kleiner dan de afwijkingen in de ramingen van de begrotingsopmaak. Inderdaad, in de begrotingscontrole is de gemiddelde afwijking slechts 1,8 miljard. Dit is geen verrassing: tijdens de begrotingscontrole is er recentere informatie beschikbaar. Zelfs de grote overschatting van 13,8 miljard in 2009 werd bij de begrotingscontrole teruggebracht tot 5 miljard.

Aangezien over- en onderschattingen elkaar compenseren bij het berekenen van gemiddelden over meerdere jaren, is het gemiddelde van de absolute waarde van de afwijkingen een betere indicator van de *accuraatheid* van een ramingsmethode. De gemiddelde absolute afwijking bedraagt 3,1 miljard euro in de initiële begroting, en 2 miljard euro bij de begrotingscontrole. Ook accuraatheid is dus hoger bij de begrotingscontrole.

De fiscale ontvangsten worden niet als één geheel geraamd. De totale afwijking in de raming van de fiscale ontvangsten is het gevolg van de afwijking in de ramingen van de verschillende categorieën van belastingontvangsten. Aangezien het grootste deel van de ramingen gebeurt binnen het kader van de gedesaggregeerde methode, verwachten we dat de ramingsafwijking grotendeels zal bepaald worden door de afwijking binnen de gedesaggregeerde methode. Figuur 3 bevestigt dit. Daarin tonen we de decompositie van de ramingsafwijking in de totale fiscale ontvangsten naar de gedesaggregeerde methode (rood), ontvangsten geraamd buiten de gedesaggregeerde methode (blauw), en ontvangsten waarvan de inning werd overgedragen (geel)³. De totale ramingsafwijking werd gecorrigeerd voor deze overdrachten, en wordt weergegeven door de zwarte ruiten.

³ Na de toegenomen fiscale autonomie, bijv. in de vorm van gewestelijke belastingen, bleef de federale overheid in eerste instantie verantwoordelijk voor de inning van alle belastingontvangsten. De regionale overheid neemt vaak pas later de inning over. Zo werd bijvoorbeeld in 2011 de inning van de verkeersbelasting overgenomen door het Vlaams Gewest. Tijdens de initiële opmaak van de begroting was echter niet altijd zeker wanneer de inning effectief door de regio's zou overgenomen worden. Daardoor maken die de ontvangsten toch deel uit van de initiële raming. Als de transfer dan toch plaatsvond, is een deel van de ramingsafwijking te wijten aan dergelijke overdracht van inning. In wat volgt corrigeren we systematisch voor deze overdrachten.

Figuur 3: de afwijkingen in de ramingen opgesplitst naar de afwijking binnen en buiten de gedesaggregeerde methode

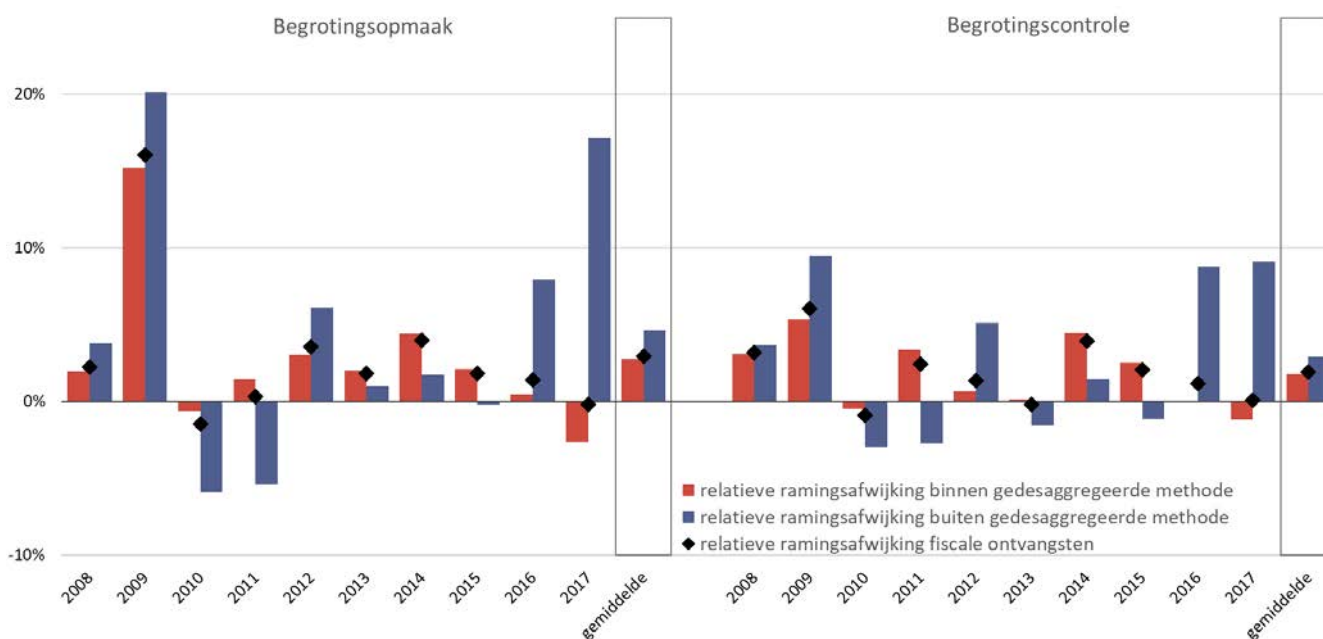


De gemiddelde afwijking over de besproken periode, 2008-2017, bedraagt 2,1 miljard euro voor de ramingen binnen de gedesaggregeerde methode en slechts 684 miljoen euro buiten de gedesaggregeerde methode. In 2009 was de afwijking binnen de gedesaggregeerde methode bijna 11 miljard euro, terwijl dit buiten deze methode ‘slechts’ 2,9 miljard euro was.

Een kleine ramingsafwijking in miljoen euro impliceert een grote relatieve fout wanneer de raming slechts een klein bedrag in ontvangsten betreft. Daarenboven is het logisch dat de afwijkingen van grote te ramen ontvangsten over het algemeen belangrijker zijn dan afwijkingen van ramingen van kleine ontvangstencategorieën. Daarom tonen we in Figuur 4 ook de relatieve afwijking, gedefinieerd als de ratio tussen de afwijking en de realisatie van de te ramen ontvangsten.

Opnieuw komt het linkse paneel in de figuur overeen met de begrotingsopmaak, en het rechtse paneel met de begrotingscontrole. De ruiten geven de relatieve afwijking van de totale fiscale ontvangsten. De rode staven geven de relatieve afwijkingen in de gedesaggregeerde methode, en de blauwe staven de relatieve afwijkingen buiten de gedesaggregeerde methode.

Figuur 4: de relatieve afwijkingen in de ramingen opgesplitst naar de afwijking binnen en buiten de gedesaggregeerde methode (in %)



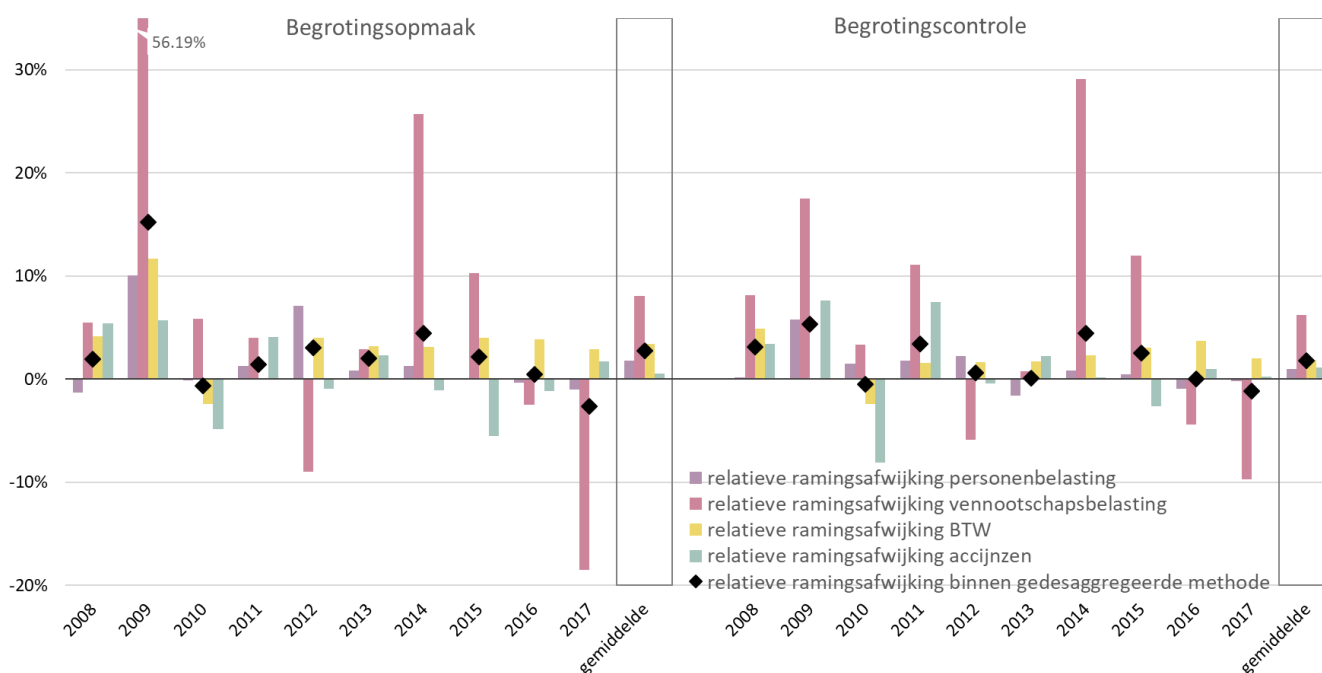
Het is meteen duidelijk dat de ramingen buiten de gedesaggregeerde methode veel grotere relatieve afwijkingen kennen dan de ramingen met de gedesaggregeerde methode. In de initiële begrotingen werden de ontvangsten buiten de gedesaggregeerde methode gemiddeld overschat met 4,6%. Binnen de gedesaggregeerde methode was dit slechts 2,7%. Uitgedrukt in miljard euro bedraagt de afwijking buiten de gedesaggregeerde methode in 2009 slechts een kwart van de afwijking binnen de gedesaggregeerde methode (2,9 miljard euro t.o.v. 11 miljard euro, zie Figuur 2). Maar in verhouding tot de ontvangsten, staat die afwijking voor een overschatting van 20%, tegenover slechts 15% van de raming met de gedesaggregeerde methode.

Om de accuraatheid van beide methoden te peilen, gebruiken we de absolute relatieve afwijkingen (niet weergegeven in de figuur). In de begrotingsopmaak bedraagt de gemiddelde absolute relatieve afwijking 3,4% binnen de gedesaggregeerde methode en 6,9% buiten de gedesaggregeerde methode. In de begrotingscontrole was dit respectievelijk 2,1% en 4,6%. De relatieve accuraatheid is duidelijk groter binnen de gedesaggregeerde methode.

De afwijkingen weergegeven in Figuur 4 zijn nog steeds een resultaat van verschillende ramingen. De afwijkingen voor de gedesaggregeerde methode zijn het gevolg van afwijkingen in de raming van de personenbelasting, de vennootschapsbelasting, de BTW en de accijnzen. Buiten de gedesaggregeerde methode gaat het over een heel aantal kleine ontvangstencategorieën die afzonderlijk geraamd worden. Figuur 5 toont de relatieve afwijkingen voor de vier ontvangstencategorieën van de gedesaggregeerde methode.⁴ Voor elk jaar tonen we zowel de relatieve afwijking van de totale raming met de gedesaggregeerde methode (de ruiten), alsook de relatieve afwijking van iedere categorie (de staven).

⁴ De kwaliteit van de verschillende ramingen buiten de gedesaggregeerde methode is te vinden in hoofdstuk 5 van het eindrapport.

Figuur 5: de relatieve afwijkingen binnen de gedesaggregeerde methode (in %)



De relatieve afwijkingen binnen de vennootschapsbelasting vallen op. Deze zijn veruit het grootst. De afwijking van de raming van de vennootschapsbelasting in 2009 past zelfs niet volledig op de figuur wanneer we willen dat de andere afwijkingen zichtbaar blijven.⁵ Over de hele beschouwde periode 2008-2017 bedroeg de relatieve afwijking bij de begrotingsopmaak 8% voor de vennootschapsbelasting, 3,4% voor de BTW, 1,8% voor de PB en 0,6% voor de accijnzen. Hetzelfde relatieve patroon vinden we bij de ramingen van de begrotingscontrole. Daarnaast valt het ook op dat er grote verschillen zijn in de relatieve accuraatheid over de verschillende belastingen doorheen de tijd. Ze volgen allemaal een zeer verschillend verloop over de jaren heen.

Om de accuraatheid van de ramingen van de subcategorieën te bespreken, gebruiken we opnieuw de absolute relatieve afwijking. Gemiddeld kende de vennootschapsbelasting een absolute relatieve afwijking van 14% in de initiële begroting, en 10% tijdens de begrotingscontrole. Voor de andere categorieën ligt deze gemiddelde absolute relatieve fout veel lager, en zijn de ramingen dus accurater. In de personenbelasting bijvoorbeeld was dit voor de initiële ramingen 2,3%, en voor de ramingen in de begrotingscontrole slechts 1,56%. Dit duidt op een grote accuraatheid in de ramingen van de personenbelasting. De gemiddelde absolute afwijking voor BTW en voor accijnzen ligt dicht bij elkaar, namelijk rond 3 à 4% in de initiële ramingen en rond 2 à 3% in de begrotingscontroles. Ook in deze twee laatste categorieën kunnen we spreken van tamelijk grote accuraatheid in de ramingen.

HOUDEN DE RAMINGEN EEN SYSTEMATISCHE OVERSCHATTING IN VAN DE WERKELIJKE ONTVANGSTEN?

Om na te gaan of de ramingsfout systematisch positief (of negatief) is en dus systematisch een overschatting (of onderschatting) van de werkelijke ontvangsten met zich meebrengt, gebruiken we het concept 'vertekening' of '*bias*'. Wanneer er geen vertekening is in de ramingen, zouden de

⁵ Ook in Nederland blijkt de vennootschapsbelasting de moeilijkste ontvangstencategorie om te voorspellen. Bij de doorlichting van de ramingsmethoden rapporteerde de visitatiecommissie (2018) een gemiddelde relatieve absolute ramingsafwijking van 16,6%

afwijkingen elkaar op lange termijn moeten compenseren. De gemiddelde afwijking ligt dan dicht bij nul. Aangezien we met een relatief kleine dataset werken, en het moeilijk te zeggen is wat ‘dicht bij nul’ betekent, gebruiken we een significantietest op het gemiddelde. Dergelijke statistische toets vertrekt van een zogenaamde nulhypothese: onze nulhypothese stelt dat de gemiddelde afwijking nul is, en dat er dus geen vertekening is. In dat geval stelt de alternatieve hypothese dat het gemiddelde van de afwijkingen groter (kleiner) is dan nul.⁶ Wanneer we de nulhypothese van ‘geen vertekening’ verwerpen, concluderen we dat de ramingen vertekening vertonen.

Het resultaat van deze statistische toets wordt weergegeven door een p-waarde, die de kans weergeeft dat we de nulhypothese verkeerdelijk verwerpen. Bij een p-waarde van 0,10 bijvoorbeeld, is er 10% kans dat de nulhypothese onterecht wordt verworpen. De uitspraak dat de ramingen vertekening vertonen, heeft dan een betrouwbaarheid van 90%. Kortom, hoe lager de p-waarde, hoe betrouwbaarder de uitspraak dat er een vertekening is. Een p-waarde van 0,1 is ook de grenswaarde die werd gebruikt in de analyse.

In Tabel 1 geven we het resultaat van deze toets weer voor verschillende onderdelen van de ramingen. Het linkse paneel in de tabel staat voor de initiële ramingen, i.e. de begrotingsopmaak, rechts staan de resultaten voor de ramingen bij de begrotingscontrole. In elk paneel wordt de gemiddelde relatieve afwijking getoond over de jaren 2008-2017, alsook de p-waarde van de significantietest voor een systematische overschatting. We gebruiken kleurschakeringen om snel te zien of er kan geconcludeerd worden dat er vertekening is, en ook hoe betrouwbaar die conclusie is. De cellen met p-waarden boven 0,10 blijven ongekleurd en laten toe te besluiten dat er geen vertekening is. Gekleurde p-waarden wijzen op vertekening en deze conclusie is des te betrouwbaarder (statistisch gesproken) naarmate de kleur donkerder is.

Tabel 1: vertekening in de ramingen van de fiscale ontvangsten

	2008-2017			
	Begrotingsopmaak		Begrotingscontrole	
	gemiddelde relatieve afwijking (%)	p-waarde (gem. rel. afw.=0) ⁷	gemiddelde relatieve afwijking (%)	p-waarde (gem. rel. afw.=0)
totale fiscale ontvangsten	2.95	0.03	1.92	0.01
buiten de gedesaggregeerde methode	4.63	0.06	2.91	0.05
binnen de gedesaggregeerde methode	2.74	0.05	1.79	0.02
personenbelasting	1.78	0.08	1.00	0.08
vennootschapsbelasting	8.04	0.20	6.18	0.12
BTW	3.43	0.00	1.83	0.01
accijnzen	0.56	0.38	1.10	0.26

De bovenste rij in Tabel 1 toont de cijfers voor de totale fiscale ontvangsten. De gemiddelde afwijking in deze raming bedraagt 2,95% en deze raming is duidelijk vertekend, want de p-waarde bedraagt 0,03,

⁶ Het gaat hier om een *one-sided t-test*. Bij een positieve gemiddelde afwijking is de alternatieve hypothese dat het gemiddelde groter is dan nul, by een negatieve afwijking is de alternatieve hypothese een gemiddelde kleiner dan nul.

⁷ gem. rel. afw.=0 is de nulhypothese en staat voor ‘gemiddelde relatieve afwijking = 0’.

wat ver onder de kritieke waarde van 0,10 is. Zowel de ramingen binnen als buiten de gedesaggregeerde methode zijn vertekend met respectievelijke p-waarden van 0,06 en 0,05 tijdens de begrotingsopmaak. Merk bovendien op dat in alle categorieën de gemiddelde relatieve afwijking positief was. Binnen de gedesaggregeerde methode zien we een sterke vertekening voor de BTW en voor de personenbelasting, zowel in de initiële begroting alsook in de aangepaste begroting.

Dergelijke analyse is natuurlijk deels afhankelijk van de geobserveerde periode. Zo speelt de sterke overschatting in 2009 een belangrijke rol. Daarom voerden we voor de ontvangsten binnen de gedesaggregeerde methode de analyse ook uit op de periode 1990-2017, en op verschillende sub-perioden. Tabel 2 toont deze resultaten.⁸

Tabel 2: vertekening in de ramingen in de gedesaggregeerde methode over sub-perioden

	1990-2017		1990-2005		2006-2013		2014-2017	
	gemiddelde relatieve afwijking (%)	p-waarde (gem. rel. afw.=0)	g.r.a.(%)	p-w.	g.r.a.(%)	p-w.	g.r.a.(%)	p-w.
binnen de gedesaggregeerde methode	1.55	0.02	0.81	0.21	3.27	0.04	1.09	0.28
personenbelasting	1.02	0.05	0.45	0.33	2.70	0.04	-0.02	0.48
vennootschapsbelasting	6.86	0.09	6.64	0.15	8.86	0.14	3.76	0.47
BTW	1.43	0.02	0.43	0.42	2.41	0.08	3.44	0.00
accijnzen	0.08	0.43	-0.90	0.15	2.85	0.04	-1.51	0.20

Ook getoetst over de hele periode 1990-2017 vertonen de ramingen van de fiscale ontvangsten, met de gedesaggregeerde methode, een systematische overschatting. Wat wel opvalt, is dat die overschatting zich voor de verschillende categorieën voornamelijk voordoet in de periode 2006-2013. Dit bevestigt het hierboven vermelde vermoeden dat het recessiejaar 2009 een belangrijke rol speelt. Anderzijds is de systematische overschatting niet volledig toe te schrijven aan dat crisisjaar. Zelfs wanneer we 2009 uit de analyse weglaten, vinden we een positieve vertekening (niet weergegeven).

KUNNEN WE DE AFWIJINGEN IN DE RAMINGEN VERKLAREN?

We stellen dus vast dat er wel degelijk afwijkingen zitten tussen de raming van de ontvangsten en de realisatie. Bovendien zijn deze afwijkingen vertekend. De logische volgende stap is in dat geval nagaan wat de oorzaken kunnen zijn van dergelijke afwijkingen. Om die vraag te kunnen beantwoorden moeten we vanzelfsprekend een grondig inzicht hebben in de methodologie van de ramingen. Het eindrapport bevat daarom ook een uitgebreid deel waarin deze methodologie in detail wordt besproken en ontleed. In de volgende paragraaf bespreken we summier deze methodologie. Voor een gedetailleerde beschrijving verwijzen we naar hoofdstuk 2 van het eindrapport voor wat betreft de gedesaggregeerde methode, en naar hoofdstuk 4 voor een beschrijving van de methoden buiten de gedesaggregeerde methode.

⁸ De sub-perioden komen overeen met de grote wijzigingen in de gedesaggregeerde methode. Aangezien de raming van de vennootschapsbelasting in 2015, en niet in 2014 werd herzien, werden de subperioden voor de vennootschapsbelasting anders opgesteld in de analyse. In deze tabel berekenden we de gemiddelde relatieve afwijking en de p-waarden voor de gerapporteerde perioden, om inconsistenties te vermijden.

HOE WORDEN DE RAMINGEN OPGESTELD?

De raming die in de begroting verschijnt, kan opgesplitst worden in een raming die constant beleid veronderstelt, we noemen deze raming de 'baseline voorspelling', en een raming van de impact van beleidsveranderingen. De methoden gebruikt in beide componenten zijn uiteraard zeer verschillend.

De baseline voorspelling, die de veronderstelling van ongewijzigd beleid uitdrukt, start van de ontvangsten in het huidige jaar, en past daar een groeivoet op toe. Binnen de gedesaggregeerde methode wordt deze groeivoet geraamd op basis van voorspellingen van de evolutie van bepaalde macro-economische grootheden. De geraamde groeivoet van de accijnzen bijvoorbeeld wordt geraamd op basis van o.m. de voorspelde groei in de private consumptie. Die voorspellingen komen uit de door het Federaal Planbureau gepubliceerde Economische Begroting. De link tussen de voorspelde groei van de macro-economische variabelen en de geraamde groei in de ontvangsten wordt gecapteerd door partiële elasticiteiten. Zo wordt voor de accijnzen een partiële elasticiteit m.b.t. private consumptie gebruikt van 0,54. Dit betekent dat wanneer de private consumptie groeit met 1% dat de ontvangsten uit accijnzen zullen groeien met 0,54%, ceteris paribus. De elasticiteiten zelf worden geschat op basis van historische tijdreeksen van de fiscale ontvangsten en de macro-economische variabelen. Deze worden regelmatig herschat op recente data.

Buiten de gedesaggregeerde methode is het niet altijd duidelijk hoe de geraamde groeivoet tot stand komt. In de meeste gevallen is deze gebaseerd op een historische trendanalyse; de gemiddelde groeivoet van de betrokken ontvangsten in de voorgaande jaren wordt gebruikt. Daarnaast wordt er in bepaalde categorieën gebruik gemaakt van externe variabelen. Dit is bijvoorbeeld het geval in raming van de roerende voorheffing. Het is dan niet duidelijk hoe deze externe informatie gecombineerd wordt met de geschatte trend.

Aan de raming bij ongewijzigd beleid wordt daarna een beleidsimpact toegevoegd wanneer zich beleidswijzigingen voordoen. De raming van de beleidsimpact steunt op een scala van methoden, waaronder microsimulatiemodellen en historische statistieken. Deze methoden worden zelden tot nooit toegelicht in de begroting. De uiteindelijke raming van de beleidsimpact in een bepaalde categorie, bijv. de ontvangsten uit de personenbelasting, wordt opgeteld bij de baseline voorspelling van die categorie.

Ten slotte worden naast de beleidsimpact ook bepaalde 'verschuivingen en technische factoren' geraamd die worden opgeteld bij de 'baseline voorspelling'. Hierbij gaat het veelal om toevallige ontvangsten of een verschuiving in inning door factoren die ook als toevallig kunnen worden beschouwd. In deze executive summary tonen we deze in de analyse als beleidsimpact. In de studie en het eindrapport worden de verschuivingen en technische factoren wel afzonderlijk behandeld van de beleidsimpact.

Met dit inzicht in de methode kunnen we de vorige vraag opnieuw behandelen:

KUNNEN WE DE AFWIJINGEN IN DE RAMINGEN VERKLAREN?

In de studie hebben we drie van vier mogelijk oorzaken voor de vastgestelde afwijking in de ramingen geïsoleerd: afwijkingen in de raming van de ontvangsten in het huidige jaar, afwijkingen in de macro-economische voorspellingen, en afwijkingen in de raming van de beleidsimpact. Het deel van de totale afwijking dat niet kan worden toegewezen aan één van deze drie componenten van de ramingsfout, de residuele afwijking, interpreteren we als de afwijking die volgt uit de specificatie van de methode. Het slaat op de manier waarop de elasticiteiten geschat worden: de functionele specificatie, de keuze

van verklarende macro-economische variabelen, enz. Deze residuele fout meet hoe groot de afwijking zou zijn als we met deze elasticiteiten de raming hadden uitgevoerd op basis van perfecte informatie omtrent de ontvangsten in het huidige jaar, het toekomstige verloop van de macro-economische variabelen en de beleidsimpact.

Het was niet vanzelfsprekend om de vier verschillende componenten af te zonderen. Een eerste stap bestond erin het effect van de raming van beleidsimpact af te zonderen. De afwijking die overblijft nadat we deze component geneutraliseerd hebben, volgt uit afwijkingen in de raming van de ontvangsten in het huidige jaar, uit afwijkingen in de macro-economische voorspelling en uit de residuele afwijking. We noemen dit de baseline ramingsafwijking.

Om deze baseline afwijking te construeren, moeten we dus zowel de raming als de realisatie van de beleidsimpact kennen. Maar in de meeste gevallen is de realisatie van beleidsimpact niet of heel moeilijk te observeren. In de analyse hebben we deze realisatie benaderd door een 'best mogelijke schatting' van de beleidsimpact, die ex post geschat werd door FOD Financiën.

Tabel 3 vergelijkt deze baseline afwijkingen (in het linkse paneel) met de historisch geobserveerde afwijkingen (in het rechtse paneel). Het betreft de ramingen voor de initiële begroting.

Tabel 3: de vergelijking van de historische en baseline voorspelingsafwijkingen

	2008-2017					
	baseline voorspelling			historische voorspelling		
	gemiddelde absolute relatieve afwijking (%)	gemiddelde relatieve afwijking (%)	p-waarde (gem. rel. afw.=0)	gemiddelde absolute relatieve afwijking (%)	gemiddelde relatieve afwijking (%)	p-waarde (gem. rel. afw.=0)
totale fiscale ontvangsten	3.02	1.99	0.06	3.29	2.95	0.03
buiten de gedesaggregeerde methode	7.41	5.20	0.04	6.93	4.63	0.06
binnen de gedesaggregeerde methode	2.68	1.53	0.10	3.29	2.74	0.05
personenbelasting	2.09	0.82	0.17	2.34	1.78	0.08
vennootschapsbelasting	12.19	2.71	0.35	14.04	8.04	0.20
BTW	3.17	2.46	0.01	3.93	3.43	0.00
accijnzen	2.50	1.08	0.18	3.29	0.56	0.38

In absolute waarde zijn de baseline afwijkingen gemiddeld gezien kleiner dan de afwijkingen zonder neutralisatie van de beleidsimpact. De accuraatheid van de ramingen neemt dus toe als we de beleidsimpact buiten beschouwing laten. Voor de totale fiscale ontvangsten gaat de fout van 3,29% tot 3,02%. Bemerkt wel dat dit enkel geldt voor de ramingen op basis van de gedesaggregeerde methode. De ramingen buiten de gedesaggregeerde methode worden *minder* accuraat als de beleidsimpact weggelaten wordt.

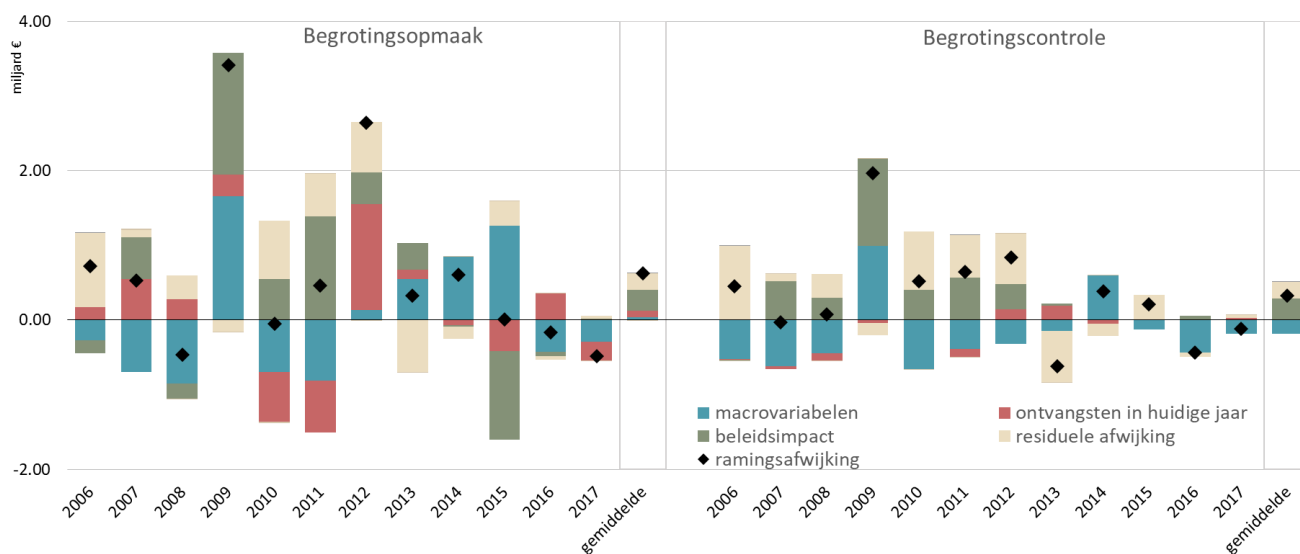
Ook de vertekening wordt kleiner na neutralisering van de afwijkingen in de beleidsimpact: de gemiddelde relatieve afwijking is voor de baseline voorspelling telkens kleiner dan de gemiddelde relatieve afwijking in de historische voorspelling (1,99% i.p.v. 2,95%). Ook hier stellen we vast dat dit enkel geldt voor de ontvangsten geraamd met de gedesaggregeerde methode. In de personenbelasting, waar vertekening werd vastgesteld in de historische ramingen, blijken de baseline voorspellingen niet langer vertekend. De systematische overschattingen in de personenbelasting zijn dus te wijten aan de schatting van de impact van het beleidswijzigingen. Deze vaststelling geldt mutatis

mutandis ook voor de vennootschapsbelasting en BTW, zodat we kunnen stellen dat de vertekening in deze voorspellingen op zijn minst deels het resultaat is van de vertekening in de geraamde beleidsimpact. Alleen voor de raming van de accijnsopbrengst neemt de gemiddelde relatieve fout toe van 0,56% naar 1,08%. Dit wijst erop dat de ramingsfouten in de beleidsimpact op accijnsopbrengsten vaak tegengesteld waren aan de afwijking in de baseline voorspelling.

Voor een verdere opsplitsing van de afwijking in de resterende drie componenten – de raming van ontvangsten in het huidige jaar, de macro-economische voorspelling, en de residuele afwijking – hebben we gedetailleerde informatie nodig. Dat zorgt ervoor dat we de volledige decompositie enkel konden uitvoeren voor de ramingen met de gedesaggregeerde methode, en voor een beperkt aantal jaren. Voor de personenbelasting en BTW konden we de decompositie berekenen vanaf 2006, voor de vennootschapsbelasting en accijnzen slechts vanaf respectievelijk 2014 en 2015.

Figuur 6 toont de volledige decompositie van de afwijking in de ramingen voor de personenbelasting in de vier componenten. Het linkse deel van de grafiek ontbindt de ramingsafwijkingen van de begrotingsopmaak, en het rechtse deel van de figuur deze uit de begrotingscontrole. In elk jaar wordt de totale ramingsafwijking, de ruit, opgesplitst in de afwijking te wijten aan de beleidsimpact, in de afwijking in de raming van ontvangsten in het huidige jaar, in de afwijking in de voorspelling van de macrovariabelen, en in een residuele component die de methode ‘an sich’ capteert.

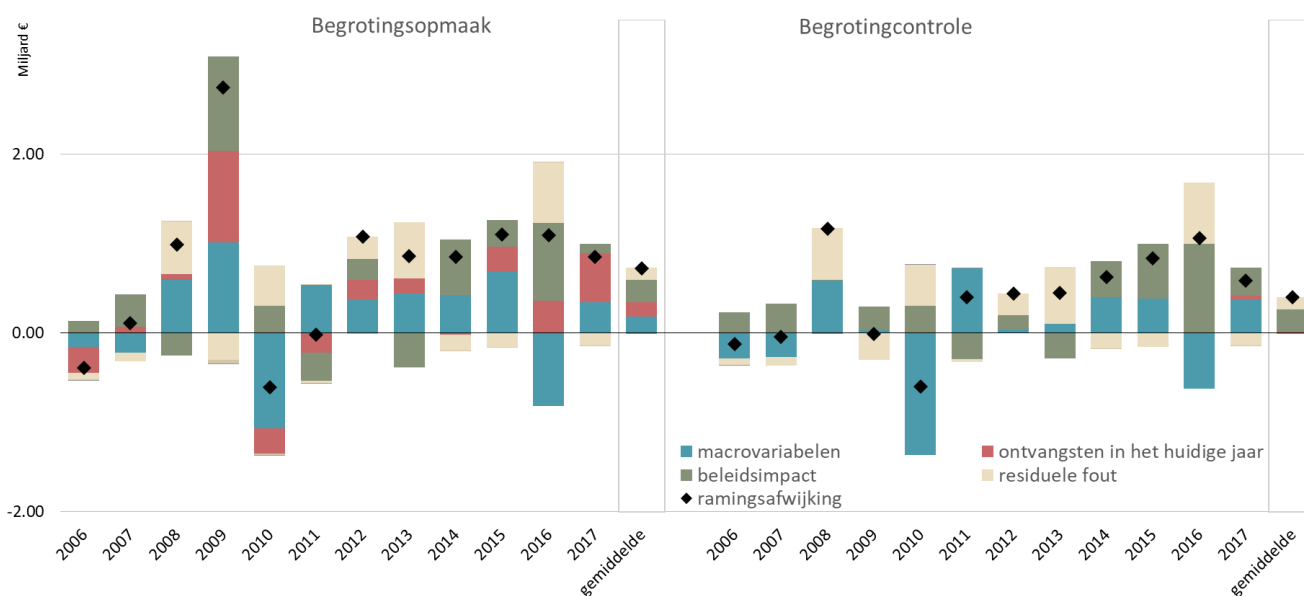
Figuur 6: decompositie van de afwijkingen in de personenbelasting



Figuur 6 toont heel duidelijk dat een niet onbelangrijk deel van de ramingsafwijking te wijten is aan de projecties van de macro-economische variabelen (blauw), dit zowel in de ramingen voor de initiële begroting als in deze voor de begrotingscontrole. Ook de rol van de inschatting van de beleidsimpact en vooral verschuivingen en technische factoren (groen) zijn in bepaalde jaren van groot belang. Overigens zien we dat deze afwijkingen eerder positief zijn. Dat de bijdrage tot de ramingsfout van de inschatting van de raming van de ontvangsten in de lopende jaar (rood) enkel belangrijk is bij de initiële ramingsafwijking, en niet bij de begrotingscontrole, is logisch. Op het moment van de begrotingscontrole heeft men al een veel preciezer beeld van de ontvangsten in dat lopende jaar: met uitzondering van de inkohieren zijn alle ontvangsten in de personenbelasting reeds geobserveerd.

De residuele afwijking (licht beige) is eerder beperkt. Dit betekent dat de gedesaggregeerde methode relatief goed de evolutie van de ontvangsten voorspelt, als ze beschikt over perfecte ramingen van macro-economische variabelen, en over perfecte ramingen van de beleidsimpact. Deze fout wordt trouwens heel klein na 2014. Dit wijst erop dat de herziening van de methode die in dat jaar werd geïmplementeerd, de ramingen zeker ten goede kwam.

Figuur 7: decompositie van de afwijkingen in de BTW



Figuur 7 toont dezelfde decompositie voor de afwijking in de raming van de BTW-ontvangsten. Dat de historische afwijkingen (de ruiten) bijna altijd positief zijn, bevestigt de vertekening in de ramingen die we al vaststelden in Tabel 2. Nu wordt duidelijk dat die vertekening in de ramingen in grote mate wordt veroorzaakt door de afwijkingen in de macro-economische cijfers, en door de ramingsafwijking bij de inschatting van de beleidsimpact. De residuele fout is klein in vergelijking met de andere fout-componenten, en is niet systematisch positief.

In accijnzen en in de vennootschapsbelasting konden we de historische raming slechts vanaf resp. 2014 en 2015 opsplitsen naar de verschillende componenten. De besluiten die hieruit kunnen worden getrokken, zijn dan ook beperkt. We verwijzen naar hoofdstuk 3 paragraaf 4.4.2 en paragraaf 4.4.4 voor deze decomposities.

PRESTEERT DE HUIDIGE SPECIFICATIE VAN DE GEDESAGGREGEERDE METHODE BETER?

De specificatie van de gedesaggregeerde methode werd sinds de ingebruikname een aantal keer gewijzigd en geüpdatet. De voornaamste wijzigingen werden doorgevoerd in 2006 en in 2014. In 2014 werd onder meer het voorspellen van de ontvangsten van de personenbelasting opgesplitst naar het voorspellen van de bedrijfsvoorheffing, de voorafbetaling, en de inkohieren. Ook in andere belastingcategorieën veranderde de specificatie. Zo werden in de vennootschapsbelasting nieuwe macro-economische variabelen meegenomen in de ramingsvergelijking, net zoals bij de accijnzen. Bovendien werden in 2017 een aantal van de partiële elasticiteiten, die gebruikt worden om de macro-economische variabelen te linken aan de fiscale ontvangsten, herschat.

De analyse van de historische afwijkingen toonde al aan dat de ramingen sinds 2014 accurater zijn, en minder vertekend. Dit is een eerste indicatie dat de herziening van de methoden verbetering inhield. Vier jaar is vanzelfsprekend wat kort om deze voorlopige conclusie als voldoende robuust te beschouwen. Daarom vonden we het nuttig om de specificatie van de op dit moment gebruikte methode, met de huidige elasticiteiten, te testen op basis van extra observaties. We deden dit door de huidige specificatie van de methode toe te passen op historische input, zoals de historische projecties van de macrovariabelen, en de historische ramingen van de beleidsimpact.

Zonder deze analyse hier in detail te herhalen, vermelden we wel de conclusie ervan: de huidige specificatie van de methode blijkt ook volgens deze analyse betere voorspellingen op te leveren dan de voorheen gebruikte specificaties. We verwijzen naar hoofdstuk 3, paragraaf 5 van het eindrapport, voor een gedetailleerde bespreking van deze analyse.

IS DEZE METHODE VERGELIJKBAAR MET DE METHODE GEHANTEERD IN DE BUURLANDEN?

De methoden die in België gebruikt worden, zijn zeer vergelijkbaar met de methoden gebruikt in de buurlanden. Dit geldt althans voor de methoden die gekend zijn. Het gaat dan vooral om de gedesaggregeerde methode, en de algemene methode van de historische trendanalyse buiten de gedesaggregeerde methode. In hoofdstuk 2 van het eindrapport bespreken we de vergelijking met de methoden gehanteerd in Duitsland, Nederland, Frankrijk en het Verenigd Koninkrijk.

In al deze landen wordt de voorspelling opgesteld met een gedesaggregeerde aanpak, d.w.z. dat de ontvangsten eerst opgesplitst worden in een aantal categorieën alvorens ze worden geraamd voor het volgende jaar. In al deze landen is de raming, net zoals in België, een optelsom van een baseline voorspelling en een raming van de beleidsimpact. Die baseline voorspelling is meestal gebaseerd op de raming van een baseline groei die wordt toegepast op de ontvangsten in het huidige jaar. De baseline groei wordt op zijn beurt geraamd op basis van geprojecteerde macro-economische ontwikkelingen, of op basis van een trendanalyse, wat overeenkomt met de Belgische methode.

Voor de raming van de impact van beleidsveranderingen vinden we wel een aantal grote verschillen. Voornamelijk het feit dat in het Verenigd Koninkrijk en in Nederland er een onafhankelijk begrotingsinstituut (resp. *the Office for Budget Responsibility* en het Centraal Planbureau) instaat om de kwaliteit van die ramingen te garanderen, of op z'n minst te bespreken, is een belangrijk verschil met de Belgische praktijk.

ZIJN ER VERBETERINGEN MOGELIJK?

Over het algemeen zijn de voorspellingsmethoden kwalitatief goed. De afwijkingen die geobserveerd worden zijn moeilijk te vermijden. Ze volgen zowel uit de voorspellingen in de macro-economische cijfers, als uit de inschatting van de impact van nieuw beleid, en uit de zogenaamde "verschuivingen en technische factoren". De restterm, die we in dit rapport interpreteren als de fout die volgt uit *de methode zelf*, is eerder beperkt.

We zien één duidelijke en noodzakelijk verbetering mogelijk, en volgen daarin het Rekenhof (2017): we raden aan om toereikende documentatie te voorzien omtrent de methoden buiten de gedesaggregeerde methode. Om ook deze ramingen grondig te evalueren, is het nodig dat een document wordt opgesteld waarin de methode wordt toegelicht.

Daarnaast zijn er ongetwijfeld verbeteringen mogelijk wat betreft transparantie. Die zijn misschien minder noodzakelijk, maar ze zouden zonder twijfel de kwaliteit van de ramingen ten goede komen.

Een eerste punt waarin de transparantie beter kan, betreft de raming van de beleidsimpact. Het lijkt aangewezen dat bij de ramingen die in de begroting worden opgenomen een aantal kenmerken van de methodologie worden toegelicht. Dit is vooralsnog niet het geval en hindert de evaluatie van dergelijke ramingen. Een tweede punt waarop de transparantie kan toenemen is binnen de gedesaggregeerde methode. Zo zijn er een aantal zeer specifieke zaken die niet aan bod komen in Valenduc & Lenoir (2014) of in de Algemene Toelichting bij de begroting, maar die wel nuttig kunnen zijn voor een externe gebruiker van de begrotingsramingen. Het gaat in het bijzonder om de constructie van een reeks van belastingontvangsten onder de assumptie van constant beleid, en de recurrente herschatting van de partiële elasticiteiten.

In het laatste hoofdstuk van het eindrapport stellen we nog een aantal bijkomende paden tot verbetering voor. Ook al weten we van die voorstellen niet a priori of ze tot kwalitatief betere ramingen zullen leiden, de investering om de mogelijke kwaliteitswinst te onderzoeken lijkt ons eerder beperkt.

REFERENCES

- Lenoir, T. & C. Valenduc (2006) 'Herziening van de macro-economische methode voor de raming van de fiscale ontvangsten.' Documentatieblad, FPS Finance, Belgium, 66, 1, pp. 97-220.
- Rekenhof (2017) "Raming van de fiscale ontvangsten – organisatie en processen." 49 pp.
- Valenduc, C. & T. Lenoir (2014) 'Révision de la méthode de prévision des recettes fiscales.' Documentatieblad, FPS Finance, Belgium, 74, 4, pp. 5-56.
- Visitatiecommissie (2018) 'Raamwerk aan de winkel, 20 juli. Rapport van de commissie visitatie raming belasting- en premieontvangsten' published as annex 6 to the Miljoenennota 2019.