

UPDATE PRIJSMODULE TEN BEHOEVE VAN AEOLUS

TECHNISCH ACHTERGRONDDOCUMENT

NOTITIE

seo • economisch onderzoek

AUTEURS

CHRISTIAAN BEHRENS & ARNOUT JONGELING

IN OPDRACHT VAN

RWS

AMSTERDAM, 5 NOVEMBER 2024

SEO-notitie nr. 2024-138

ISBN 978-90-5220-466-6

Informatie & Disclaimer

SEO Economisch Onderzoek heeft op de verkregen informatie en data geen onderzoek uitgevoerd dat het karakter draagt van een accountantscontrole of due diligence. SEO is niet verantwoordelijk voor fouten of omissies in de verkregen informatie en data.

Copyright © 2024 SEO Amsterdam.

Alle rechten voorbehouden. Het is geoorloofd gegevens uit deze notitie te gebruiken in artikelen, onderzoeken en collegesyllabi, mits daarbij de bron duidelijk en nauwkeurig wordt vermeld. Gegevens uit deze notitie mogen niet voor commerciële doeleinden gebruikt worden zonder voorafgaande toestemming van de auteur(s). Toestemming kan worden verkregen via secretariaat@seo.nl.

Roetersstraat 29
1018 WB, Amsterdam

+31 20 399 1255
secretariaat@seo.nl
www.seo.nl

Inleiding

Opdracht

Rijkswaterstaat beheert in opdracht van het ministerie van Infrastructuur en Waterstaat (IenW) het nationale luchtvaartmodel AEOLUS dat eigendom is van IenW. Ten behoeve van een actualisatie van AEOLUS heeft Rijkswaterstaat aan SEO Economisch Onderzoek gevraagd om geactualiseerde informatie over ticketprijzen in de luchtvaart op routeniveau op te leveren. De laatste actualisatie van de schatting van de ticketprijzen heeft circa tien jaar geleden plaatsgevonden.

Werkwijze en leeswijzer

SEO maakt voor het actualiseren van de ticketprijzen onder andere gebruik van het door SEO ontwikkelde NetCost-model. Het NetCos-model identificeert alle directe en indirecte reisopties op basis van Official Airline Guide (OAG) dienstregelingsdata. Voor alle directe en indirecte reisopties worden de herkomst, bestemming, luchtvaartmaatschappij, geplande vliegtijd, de wekelijkse frequentie en eventueel de overstapluchthaven gegenereerd over een bepaald kalenderjaar. Dit bestand - het zogenoemde Level of Service-bestand - is een input voor AEOLUS en wordt regelmatig geüpdatet.

Naast de genoemde informatie per reisoptie zijn er als input voor AEOLUS ook data nodig over ticketprijzen per reisoptie. Deze ticketprijzen worden in een aparte module naast het NetCost-model geschat. Deze technische achtergrondnotitie beschrijft de belangrijkste stappen om tot de geactualiseerde geschatte ticketprijzen te komen. Deze zijn vervolgens gekoppeld aan het eerder genoemde Level of Service-bestand en als geheel geleverd aan Rijkswaterstaat. Tijdens het actualiseren is meerdere malen overleg geweest met Rijkswaterstaat en met Significance B.V.. Significance B.V. voert de actualisatie van AEOLUS uit in opdracht van Rijkswaterstaat.

De reden van het gebruik van geschatte ticketprijzen is dat de beschrijvende statistiek niet voor elke reisoptie in het Level of Service-bestand beschikbaar is en er voor het voorspellen binnen AEOLUS behoefte is aan inzicht in het verband tussen de karakteristieken van de reisoptie en de prijzen, zodat ook bij veranderingen in deze karakteristieken er informatie beschikbaar is over het prijsniveau. Deze inzichten zijn gebaseerd op de parameters in het te schatten prijsmodel.

Deze technische achtergrondnotitie is verder als volgt opgebouwd. De eerstvolgende sectie beschrijft de beschikbare data en benoemt de toegepaste databewerkingen. De volgende sectie bespreekt de modelspecificatie en geeft aan welk model wordt toegepast. In de laatste sectie zijn de schattingsresultaten opgenomen, ook geven we een korte toelichting op de interpretatie van het geschatte model.

Beschikbare data en opbouw te gebruiken dataset

Data

Voor dit onderzoek maken we gebruik van drie verschillende data: OAG Schedule Analyser, OAG Traffic Analyser (MIDT) en Schiphol Routes & Profile Monitor. De Official Airline Guide Schedule Analyser (OAG) bevat informatie over wereldwijde vluchtschema's van passagiers- en vrachtluchten over 2023. Deze data gebruiken we als input om te komen tot het Level of Service-bestand door alle directe en indirecte reisalternatieven voor alle wereldwijde combinaties van luchtvaartmaatschappij, herkomst- en bestemmingsluchthavens (markten) met een minimale

frequentie van gemiddeld twee vluchten per maand te identificeren. Incidenteel aanbod laat NetCost buiten beschouwing. Voor de indirecte reisalternatieven geldt dat het gaat om alternatieven met maximaal één overstap op een andere luchthaven en om reële reisopties waarbij de omreistijd (met overstaptijd) niet te veel hoger mag zijn ten opzichte van het directe reisalternatief.

De OAG Traffic Analyser geeft toegang tot Marketing Information Data Transfer (MIDT) data. Deze data bevatten passagiersstromen en ticketprijzen op luchtvaartmaatschappij-marktniveau per maand op basis van daadwerkelijk verkochte tickets. Voor de markt van Amsterdam Schiphol naar, bijvoorbeeld, Boston Logan International geven de data per uitvoerende luchtvaartmaatschappij een inschatting van het aantal reizigers per ticketklasse en de bijhorende gemiddelde ticketprijs. De data zijn zowel voor de directe verbindingen, in dit voorbeeld met Air-France KLM, JetBlue, Delta of Virgin Atlantic als voor indirecte alternatieven via, bijvoorbeeld, Dublin met Aer Lingus of via Kopenhagen met SAS beschikbaar. De ticketklassen zijn uitgesplitst naar first class, businessclass, premium economy, full economy en discounted economy.

De ticketprijzen in OAG Traffic Analyser zijn afkomstig uit onder andere MIDT. OAG neemt de MIDT-gegevens als uitgangspunt en past vervolgens verschillende algoritmen toe om op te schalen naar daadwerkelijke passagiersaantallen. Ook voor de ticketprijzen - inclusief de informatie uit de verschillende Global Distribution Systems (zoals Travelport etc.) - hanteert OAG eigen (niet-openbare) algoritmes. Voor de actualisatie van de ticketprijzen nemen we de door OAG aangepaste passagiersschattingen en ticketprijschatting als uitgangspunt. De te gebruiken gemiddelde ticketprijzen zijn exclusief belastingen, toeslagen en brandstoftoeslagen, maar inclusief eventuele provisies aan verkopende partijen.¹

Schiphol Routes & Profile Monitor is gebaseerd op face to face datacollectie op Schiphol, de interviews worden op verschillende locaties op de luchthaven gedurende het hele jaar gehouden onder passagiers en kent zo'n honderdduizend jaarlijkse observaties.² Voor de actualisatie van de ticketprijzen maken we gebruik van de informatie over het gedocumenteerde reismotief per reiziger per herkomst- en bestemmingsluchthaven. Voor Schiphol als geheel geldt in 2023 een verdeling van reismotieven waarbij 27 procent het zakelijk segment betreft (inclusief congres/studie), 48 procent niet-zakelijk (leisure) en 23 procent bezoeken van familie en vrienden (VFR).³

Datadekking

Om inzicht te geven in de omvang en representativiteit van de uiteindelijk te gebruiken dataset, vergelijken we eerst het totaal aantal in OAG ingeschatte passagiers met gerapporteerde aantallen. We werken dit ter illustratie uit voor Amsterdam Schiphol. Over 2023 rapporteert Schiphol een totaal van 61,9 miljoen passagiers.⁴

Tabel 1 laat de opbouw zien van het aantal vanuit OAG ingeschatte passagiers. Er zijn twee dimensies. Ten eerste, het is mogelijk om de OAG-data uit te lezen op het niveau van de uitvoerende luchtvaartmaatschappij of op basis van zowel de uitvoerende maatschappij als de codeshare partner. Voor het aantal ingeschatte passagiers vanuit OAG maakt dit een klein verschil, voor het aantal passagiers met Amsterdam als herkomst of bestemming gaat het

¹ Afgegeven tickets waar geen inkomsten tegenover staan, zoals bijvoorbeeld, voor baby's worden niet meegenomen. Daarnaast is er geen tariefinformatie beschikbaar over eventuele additionele diensten of producten, zoals extra bagage, lounge toegang of (reis)verzekering, die separaat door de luchtvaartmaatschappij in rekening gebracht kunnen worden.

² Zie <https://mcm-research.com/projecten/continu-onderzoek-luchthaven-schiphol/>.

³ Zie <https://www.schipholannualtrafficreview.nl/2023>.

⁴ Zie <https://www.schiphol.nl/nl/schiphol-group/verkeer-en-vervoer-cijfers/>.

om het verschil tussen 35,5 en 36,9 miljoen passagiers. Voor de prijsinformatie, echter, is er geen verschil. De prijsinformatie is namelijk enkel beschikbaar op het niveau van de uitvoerende luchtvaartmaatschappij.

Ten tweede is het aantal passagiers voor Amsterdam op te delen naar passagiers die Amsterdam als herkomst- of bestemmingsluchthaven gebruiken, eventueel met een overstap ergens anders, en naar passagiers die Amsterdam als overstapluchthaven gebruiken (de transferpassagier). In de OAG-data telt deze transferpassagier vanuit het perspectief van de betreffende route één keer mee, terwijl vanuit het perspectief van Schiphol deze reiziger twee keer wordt meegeteld (één aankomst en één vertrek). Tabel 1 laat zien dat het ingeschatte passagiersvolume op Amsterdam afgerond 59,7 miljoen passagiers bedraagt. Daarmee is de dekking binnen OAG voor Amsterdam over het jaar 2023 ongeveer 96 procent (59,7/61,9).

Tabel 1 De dekking van het ingeschatte passagiersvolume in OAG voor Schiphol 2023 bedraagt 96 procent

	Aantal passagiers (in miljoen)	
	Uitvoerende	Uitvoerende en codeshare
Herkomst of bestemming AMS	35,6	36,9
AMS Overstapluchthaven 1	8,6	10,3
AMS Overstapluchthaven 2	0,9	1,1
Totaal (overstap * 2)	35,6+19,0=54,5	36,9+22,8=59,7

Bron: SEO Economisch Onderzoek op basis van OAG

Omdat het voor de prijsanalyse geen verschil maakt zijn de data uit OAG verder verzameld op de uitvoerende luchtvaartmaatschappij. De dekking, gemeten naar aantal passagiers, is daarmee dus gelijk aan 88 procent (54,5/61,9). In de analyse beperken we ons tot vluchten met maximaal één overstap. Na deze selectie blijven er 35,1 miljoen herkomst- of bestemmingspassagiers over en 15,5 miljoen transferpassagiers, oftewel in totaal 82 procent (50,6/61,9).

Vervolgens kijken we enkel naar routes die zowel in het Level of Service-bestand over 2023 als in de OAG Traffic Analyser voorkomen. Voor Amsterdam volgt uit het op deze wijze matchen van de datasets dat 33,6 miljoen herkomst- of bestemmingspassagiers en 10,3 miljoen transferpassagiers gematched kunnen worden, oftewel 71 procent (43,9/61,9).

Er zijn verschillende redenen waarom een alternatief (route en luchtvaartmaatschappij) maar in één dataset voorkomt. Om het Level of Service-bestand te genereren maakt NetCost aannames over reële reisopties kijkend naar de geïmpliceerde reistijd van een indirecte vlucht, terwijl OAG Traffic Analyser deze selectie niet maakt. Ook maakt NetCost een selectie op incidenteel aanbod wat niet noodzakelijk in OAG Traffic Analyser zit. De analyse van de verschillende niet-gematchte alternatieven laat zien dat het in de meeste gevallen om relatief dunne routes met weinig passagiers gaat.⁵ Wel lijkt NetCost een striktere selectie te maken op indirecte verbindingen tussen een herkomst en bestemming die relatief dicht bij elkaar liggen. Het gaat dan bijvoorbeeld om een verbinding van Kopenhagen naar Frankfurt met een overstap op Amsterdam. Beredeneerd vanuit omreistijd is de verwachting dat zulke markten relatief klein zijn of niet bestaan, waardoor dus niet opgenomen in het Level of Service-bestand, terwijl OAG Traffic Analyser suggereert dat er wel een beperkt deel van de vervoersvraag tussen, in dit geval, Kopenhagen en Frankfurt wordt afgewikkeld via Amsterdam.

⁵ Tot slot verwijderen we observaties met een onrealistische reistijd, namelijk met een gemiddelde vliegsnelheid onder de 250 kilometer per uur, en observaties waarvoor vanaf één van de betreffende luchthavens in OAG Traffic Analyser helemaal geen vluchten zijn geregistreerd in 2023. Het gaat in totaal om enkele honderden observaties, oftewel minder dan 0,5 procent van alle observaties.

Hoewel er in deze matching circa 7 miljoen passagiers niet te matchen zijn, is de impact voor de schatting van de tarieven beperkt. Over deze 7 miljoen passagiers was maar voor minder dan 2 miljoen passagiers informatie beschikbaar over de ticketprijzen in OAG Traffic Analyser. Dit is in lijn met de eerdere opmerking dat het met name om dunne routes gaat.

Van type ticket naar reismotief

Het AEOLUS-model werkt met reismotieven. Voor de actualisatie van de ticketprijzen maken we onderscheid naar het zakelijk en niet-zakelijk reismotief. De OAG Traffic Analyser geeft echter geen expliciete informatie over het reismotief, maar geeft voor de verschillende ticketklassen de ingeschatte passagiers en prijzen per alternatief (route en luchtvaartmaatschappij) weer. Om van ticketklasse naar reismotieven te gaan hanteren we de volgende regels:

- Op basis van de Schiphol Routes & Profile Monitor bepalen we per route (dus niet specifiek per luchtvaartmaatschappij) het aandeel zakelijke reizigers (congres/studie en zaken/werk) en niet-zakelijke reizigers (alle overige categorieën);
- Deze informatie is enkel beschikbaar voor routes waarin Amsterdam voorkomt:
 - Als de informatie op routeniveau (herkomst en bestemming) beschikbaar is, wordt dit gekoppeld aan de OAG Traffic Analyser data;
 - Als in de route een luchthaven voorkomt met een verbinding met Amsterdam, wordt voor die luchthaven het aandeel zakelijk gebaseerd op alle routes vanaf Amsterdam naar die luchthaven, bijvoorbeeld voor de route Barcelona-New York is er geen directe informatie over het aandeel zakelijk verkeer, maar voor de route Amsterdam-Barcelona en Amsterdam-New York wel. Het aandeel zakelijk verkeer op de route Barcelona-New York wordt vervolgens geschat als het gemiddelde aandeel van het zakelijk verkeer van de twee losse verbindingen naar Amsterdam;
 - Indien een luchthaven geen verbinding heeft met Amsterdam wordt het gemiddeld aandeel zakelijk verkeer vanaf Schiphol gebruikt, bijvoorbeeld voor de route New York-Fort Lauderdale wordt het aandeel zakelijk bepaald als het gemiddelde van het aandeel zakelijk New York-Amsterdam en het aandeel zakelijk vanaf Schiphol in totaal;
- Op basis van het aandeel zakelijk en niet-zakelijk verkeer delen we de reismotieven in naar de verschillende ticketklassen per alternatief. Hierbij geldt de volgorde dat we eerst alle businessclasstickets toewijzen aan zakelijke reizigers. Als er meer zakelijke reizigers op het alternatief zijn dan dit type tickets, wordt de rest toegewezen aan first class tickets, enzovoort. Als er meer businessclasstickets (of first class tickets) zijn dan zakelijke reizigers, worden deze tickets toegewezen aan de niet-zakelijke reizigers.

Bij het op deze manier toewijzen van de typen tickets naar reismotieven gaan geen data verloren. We hebben wel moeten corrigeren voor onlogische verhoudingen tussen de verschillende gemiddelde tarieven van de ticketklassen. We hanteren hierbij de regel dat als de gemiddelde prijs van een business- of first class ticket op een alternatief lager is dan het gemiddelde discounted economy tarief voor hetzelfde alternatief, we de tariefinformatie over het gemiddelde business- of first class ticket als onbetrouwbaar beschouwen en buiten de verdere analyse laten.⁶

⁶ Dit geldt voor minder dan 0,5 procent van de observaties.

Modelspecificatie en modelselectie

Na de genoemde databewerkingen is het mogelijk om eerst de gemiddelde ticketprijs uit te rekenen. De gemiddelde prijs, gewogen naar het aantal passagiers per alternatief en in euro met prijspeil 2023, voor de hele wereld ligt op €564 (mediaan €346) voor zakenreizigers en €292 (mediaan €225) voor niet-zakenreizigers. Voor Amsterdam geldt €658 (mediaan €340) om €371 (mediaan €246).

In de huidige versie van het ticketprijsmodel worden de directe en indirecte reisopties geschat in hetzelfde model door gebruik te maken van een dummy variabele voor een reisoptie met overstap. Tijdens een tussentijds overleg met Rijkswaterstaat en Significance B.V. is de aanpak besproken om eerst twee aparte modellen te schatten, één voor directe reisopties en één voor indirecte reisopties. Dit maakt een analyse van het effect van reistijd en in welke functionele vorm reistijd in het model opgenomen dient te worden (lineair, kwadratisch, piecewise lineair) overzichtelijker. In het uiteindelijke model kiezen we voor het combineren van de directe en indirecte reisopties en te corrigeren via dummy en interactie variabelen.⁷ Hieronder lichten we de specificaties kort toe.

Modelspecificatie directe en indirecte reisopties apart

De modelspecificatie voor directe en indirecte reisopties apart is als volgt:

$$\ln(\text{fare}_{ijkl}) = \beta_0 + \beta_T \cdot \text{Time}_{ijkl} + \beta_H \cdot \text{Hubpremium}_{ijkl} + \beta_{M0} \cdot \text{Monopoly}_{ijk} + \beta_{ICA} \cdot \text{ICA}_{ij} + D_{Sky} + D_{Star} + D_{Onew} \\ + D_{FSA\ other} + D_{LCC\ Cha} + D_{vertrekregio} + D_{aankomstregio} + \varepsilon,$$

waarbij *fare* de gemiddelde prijs (zakelijk of niet-zakelijk in aparte schattingen) voor de route vertrekluchthaven *i* via eventueel overstapluchthaven *k* naar bestemmingsluchthaven *j* met luchtvaartmaatschappij(groep) *l* betreft. *Time* is de vliegtijd van deze route, waarbij voor indirecte vluchten de vliegtijd van de twee aansluitende vluchten bij elkaar zijn opgeteld. Hubpremium en Monopoly corrigeren voor de marktstructuur. De Hubpremium is gebaseerd op de twintig belangrijkste hubluchthavens in Europa en op de twintig belangrijkste hubluchthavens buiten Europa in 2023.⁸ Voor deze luchthavens is vervolgens vastgesteld welke luchtvaart(groep) deze luchthaven als (thuis)hub gebruikt, bijvoorbeeld Frankfurt is de thuishub van Star Alliance. De Hubpremium is een dummy die aangeeft of een directe vlucht vertrekkend is vanuit de eigen thuishubluchthaven. Monopoly is een dummy die gelijk is aan 1 als de betreffende route de enige beschikbare route is tussen de vertrekluchthaven en de bestemmingsluchthaven. ICA is een dummy variabele gelijk aan 1 voor routes waarbij de vertrekluchthaven en de bestemmingsluchthaven niet op hetzelfde continent liggen. We controleren ook voor eventuele (gemiddelde) prijsverschillen tussen de luchtvaart(groepen) via de Sky, Star, Oneworld, FSA other en LCC Charter dummies. De laatste normaliseren we (gelijk aan nul). Tot slot houden we rekening met eventuele prijsverschillen tussen vertrek- en bestemmingsregio's.⁹

De *Time* variabele testen we zowel lineair als lineaire splines met zeven intervallen (en dus zes knooppunten).¹⁰ Hiermee beslaat elke interval ongeveer vijf uur.

⁷ Daarnaast speelt er ook de dimensie van zakelijk versus niet-zakelijk. Hierbij maken we de keuze om de modelspecificatie niet te laten verschillen.

⁸ Zie ACI Connectivity Report 2024, <https://www.aci-europe.org/air-connectivity.html>.

⁹ Deze zijn gedefinieerd in lijn met de OAG wereldregio's, zie Appendix A.

¹⁰ De lineaire spline is een specifieke vorm van piecewise lineair waarbij gegarandeerd is dat de waarde van de x-variabele op de knooppunten in de twee betreffende intervallen (links en rechts van het knooppunt) gelijk is.

Modelspecificatie directe vluchten en indirecte vluchten in één model

De modelspecificatie voor directe en indirecte reisopties in één model is als volgt:

$$\ln(\text{fare}_{ijkl}) = \beta_0 + \beta_T \cdot \text{Time}_{ijkl} + \beta_H \cdot \text{Hubpremium}_{ijkl} + \beta_{M0} \cdot \text{Monopoly}_{ijk} + \beta_{ICA} \cdot \text{ICA}_{ij} + D_{Sky} + D_{Star} + D_{Onew} \\ + D_{FSA\ other} + D_{LCC\ Char} + D_{vertrekregio} + D_{aankomstregio} + \beta_{DIRECT} \cdot \text{Direct}_{ijkl} + \beta_{int} \text{Time}_{ijkl} \cdot \text{Direct} + \varepsilon$$

De interpretatie is gelijk aan de vorige modelspecificatie met dat verschil dat we expliciet een dummy variabele opnemen voor routes van vertrekluchthaven i naar bestemmingsluchthaven j waarbij k leeg is, oftewel waarbij geen overstapluchthaven is en het dus gaat om een directe vlucht. Ook testen we de interactie tussen Time en de Direct dummy.

Verskil met eerdere modelspecificatie

In vergelijking met het geschatte ticketprijsmodel in de huidige versie van AEOLUS, neemt de bovenstaande modelspecificatie enkele andere verklarende variabelen mee en is er een andere manier om met directe en indirecte reistijd om te gaan:

- Frequentie wordt niet meer als verklarende factor opgenomen vanwege de twee kanten opgaande relatie tussen de prijzen en de frequentie, dit maakt het model inclusief frequentie minder stabiel en daardoor minder geschikt voor het beoogde doel binnen AEOLUS;
- De marktstructuur wordt niet langer via de marktaandeelen (HHI) op routeniveau meegenomen, maar met de combinatie van hubpremium en de monopoly dummy. De hubpremium volgt rechtstreeks uit de recente literatuur en van de verschillende manieren om de verdere marktstructuur te includeren presteerde de monopoly dummy het best. Een dummy voor oligopolie, een variabele met het aantal aanbieders op de route, het eigen marktaandeel van een airline op de route en/of de luchthaven of de HHI zijn getoetst, en de combinatie van hubpremium met een monopoly dummy levert de meest robuuste resultaten in termen van te verwachten tekens, economische en statistische significantie;
- In plaats van enkel onderscheid te maken naar full service carriers en legacy carriers, maken we met de huidige modelspecificatie onderscheid naar de vijf categorieën zoals deze zijn opgenomen in NetCost en AEOLUS: de drie allianties, overige full service carriers en low cost-/chartercarriers;
- Om het model te schatten maken we gebruik van alle routes in het Level-of-Service-bestand waarvoor prijsdata beschikbaar zijn, in eerdere versies werd een selectie gemaakt van routes van, naar en via luchthavens rondom de Nederlandse luchthavens. Om te corrigeren voor prijsniveauverschillen per wereldregio zijn in de huidige modelspecificatie regiodummies opgenomen voor zowel de herkomst- als bestemmingsregio, een dummy voor de combinatie van de herkomst- en bestemmingsregio is onwenselijk omdat het het onmogelijk maakt om apart nog verschillen tussen continentale en intercontinentale vluchten te schatten en het een sterke interactie heeft met de tijdsvariabele;
- Het definiëren van tijd als non-stop reistijd (de reistijd van het directe alternatief) – zoals in het verleden is toegepast – voor zowel directe als indirecte verbindingen heeft als groot nadeel dat in de vergelijking van (toekomstige) directe en indirecte verbindingen van dezelfde vertrekluchthaven naar dezelfde bestemmingsluchthaven (de route, minus k) het onderscheid niet te maken is, de reistijd is dan namelijk hetzelfde ongeacht of het een directe of indirecte route betreft. Dit is niet in lijn met de data, daaruit volgt duidelijk dat een indirecte route (combinatie van twee aansluitende vluchten) een langere reistijd kent. Vandaar dat we in de huidige actualisatie ervoor kiezen om de daadwerkelijke reistijd te gebruiken en niet de nonstop reistijd.
- De reistijd is niet kwadratisch als verklarende opgenomen, maar getoetst als piecewise (spline) lineair om zo een flexibeler effect van reistijd op de gemiddelde ticketprijs toe te staan.

Modelselectie en schatter

Het uiteindelijke model combineert de directe en indirecte alternatieven, maar is wel apart geschat voor het zakelijke en niet-zakelijke reismotief. Het model is geschat via de standaard kleinste kwadratenmethode (OLS). Hierbij geldt wel dat de schatter weegt naar het aantal passagiers op het alternatief, hierdoor weegt de prijsinformatie op de dikkere routes meer mee dan deze informatie op dunnere routes. Tot slot zijn de herkomst- en bestemmingsregio's dummies expliciet als dummies geschat (en niet als fixed effect) om eventuele verschillen tussen regio's in de te schatten prijzen te kunnen duiden.

Schattingsresultaten gecombineerde modellen

Toelichting tabel

Tabel 2 toont de schattingsresultaten. Er zijn per reismotief twee modellen getoond. Het model 1 (voor zakelijk en niet-zakelijk) is het volledige model met alle interactietermen, terwijl in model 2 (voor zakelijk en niet-zakelijk) de niet statistisch significante interactietermen tussen de direct dummy en reistijd op nul zijn gesteld. Dit laatste model labelen we als het 'constrained model' en is op verzoek van Significance B.V. toegevoegd met het oog op het gebruik van de resultaten voor het voorspelmodel in AEOLUS.

De observatie bestaat uit de informatie voor het alternatief (route en luchtvaartmaatschappij) voor het jaar 2023. Het model verklaart circa 70 procent van de variantie in ticketprijzen in het zakelijk segment en circa 63 procent in het niet-zakelijk segment. De standaardfouten (en daarmee indicatie van statistische significantie) zijn geclusterd op de vertrekluchthaven. De afzonderlijke herkomst- en bestemmingsregio-parameters zijn opgenomen in Tabel A.1 in Appendix A. De interactietermen tussen directe verbinding en vliegtijd zijn niet te schatten voor het vijfde en zesde tijdsinterval omdat in dat interval er geen directe verbindingen plaatsvinden.

Interpretatie

De modelresultaten zijn plausibel en daar waar relevant in lijn met de verwachtingen op basis van economische theorie. De hubpremium is positief en geeft dus aan dat de gemiddelde ticketprijs *ceteris paribus* hoger ligt als het alternatief vertrekt uit de thuishub van een luchtvaartmaatschappij.

Door het combineren van directe en indirecte alternatieven in één model is een directe interpretatie van de tekens van de tijdsvariabelen, de direct dummy en de interacties daarvan niet eenvoudig. De andere verklarende variabelen kennen een simpeler interpretatie. De intercontinentale routes kennen gemiddeld een hogere ticketprijs, hoewel dit voor het niet-zakelijke segment niet statistisch significant blijkt. Is er geen concurrentie voor een alternatief, dan liggen de prijzen ook gemiddeld hoger. In de zakelijke markt is het duidelijk dat de ticketprijzen van de niet-low cost carriers gemiddeld hoger liggen, voor de niet-zakelijke markt zien we hetzelfde effect voor luchtvaartmaatschappijen behorende tot Star Alliance en Oneworld, maar niet voor de leden van SkyTeam.

Om een duiding te geven van de economische significantie van de verschillende parameters, namelijk met hoeveel verschilt de gemiddelde ticketprijs *ceteris paribus*, kijken we naar een voorbeeldroute waarbij we de tijdsvariabele vastprikken. De gemiddelde vliegtijd over alle directe en indirecte routes in de voor de schattingen relevante dataset, gewogen naar het aantal passagiers, is circa 4 uur. Voor enkel de directe vluchten bedraagt dit gemiddelde 3,2 uur en voor indirecte vluchten 8,4 uur.

Tabel 2 OLS-regressieresultaten zakelijk en niet-zakelijk

	Zakelijk Non-constrained	Zakelijk Constrained	Niet-zakelijk Non-constrained	Niet-zakelijk Constrained
Reistijd in interval 1	0,208*** (0,0115)	0,208*** (0,0115)	0,127*** (0,00853)	0,133*** (0,0117)
Reistijd in interval 2	0,162*** (0,00879)	0,162*** (0,00880)	0,113*** (0,00558)	0,108*** (0,00560)
Reistijd in interval 3	0,0459*** (0,00745)	0,0458*** (0,00745)	0,0721*** (0,00572)	0,0756*** (0,00568)
Reistijd in interval 4	0,142*** (0,00780)	0,142*** (0,00786)	0,103*** (0,00847)	0,101*** (0,00894)
Reistijd in interval 5	-0,0423*** (0,0148)	-0,0429*** (0,0148)	-0,0315** (0,0157)	-0,0305* (0,0158)
Reistijd in interval 6	-0,240* (0,135)	-0,239* (0,135)	-0,112 (0,0797)	-0,114 (0,0797)
Directe verbinding	-0,151*** (0,0543)	-0,151*** (0,0543)	-0,285*** (0,0393)	-0,263*** (0,0472)
Directe verbinding * reistijd int 1	0,118*** (0,0129)	0,118*** (0,0129)	0,107*** (0,00888)	0,100*** (0,00969)
Directe verbinding * reistijd int 2	-0,0338*** (0,0109)	-0,0340*** (0,0109)	-0,00711 (0,00818)	0 (0,0102)
Directe verbinding * reistijd int 3	0,0509*** (0,0170)	0,0517*** (0,0166)	0,0434*** (0,0129)	0,0364*** (0,0102)
Directe verbinding * reistijd int 4	0,0178 (0,0519)	Constrained	-0,00153 (0,0644)	Constrained
Directe verbinding * reistijd int 5	NA	NA	NA	NA
Directe verbinding * reistijd int 6	NA	NA	NA	NA
Hub premium	0,148*** (0,0479)	0,148*** (0,0479)	0,103*** (0,0389)	0,103*** (0,0389)
Intercontinentale route	0,0812** (0,0358)	0,0812** (0,0358)	0,0278 (0,0189)	0,0275 (0,0190)
Monopolie route	0,105*** (0,0279)	0,105*** (0,0279)	0,183*** (0,0193)	0,183*** (0,0193)
SkyTeam	0,344*** (0,0470)	0,344*** (0,0470)	0,0218 (0,0455)	0,0209 (0,0452)
Star	0,475*** (0,0379)	0,475*** (0,0379)	0,223*** (0,0349)	0,223*** (0,0347)
Oneworld	0,457*** (0,0511)	0,457*** (0,0511)	0,129*** (0,0397)	0,129*** (0,0394)
Andere legacy carriers	0,181*** (0,0430)	0,181*** (0,0430)	0,0246 (0,0395)	0,0242 (0,0394)
Low cost carriers / charters	Ref cat	Ref cat	Ref cat	Ref cat
Constante	3,803*** (0,115)	3,803*** (0,115)	3,973*** (0,111)	3,954*** (0,118)
Herkomstregio (dummies)	Ja	Ja	Ja	Ja
Bestemmingsregio (dummies)	Ja	Ja	Ja	Ja
Adjusted r-kwadraat	0,696	0,696	0,632	0,632
AIC	143.454	143.452	237.314	237.320
Aantal observaties	99.125	99.125	208.853	208.853

Bron: SEO Economisch Onderzoek

Bron: * p < 0,1, ** p < 0,05, *** p < 0,01, Cluster robuuste standaardfouten (vertrekluchthaven) tussen haakjes. Ttime1 0-5,5, Ttime2 5,5-10,2, Ttime3 10,2-15,0, Ttime4 15,0-19,8, Ttime5 19,8-24,6 & Ttime6 > 24,6. De observaties (route alternatief) zijn gewogen naar het aantal reizigers.

De economische significantie illustreren we door te kijken naar model 1 voor het zakelijke segment. Ter illustratie kijken we naar een vlucht vertrekkend vanuit de regio EU1 en met een bestemming in de regio EU2.

- Voor deze vlucht geldt, als uitgevoerd als directe vlucht door een low cost carrier zonder monopolie, een geschatte gemiddelde prijs van:

$$\ln(\text{fare}) = 3,803 - 0,151(\text{Direct}) + 0,208 * 4\text{uur} + 0,118 * (4 \text{ uur} \cdot \text{Direct}) + 0,343 (EU1) + 0,190(EU2) = 5,49$$

$$\text{fare} = \exp(5,49) = \text{€}242$$

Als de vlucht als indirecte vlucht plaatsvindt, dan verandert de geschatte ticketprijs in:

$$\ln(\text{fare}) = 3,803 + 0,208 * 4\text{uur} + 0,343 (EU1) + 0,190(EU2) = 5,17$$

$$\text{fare} = \exp(5,17) = \text{€}176$$

De combinatie van de dummy variabelen voor directe vluchten en de interactie met reistijd garandeert dus dat *ceteris paribus* de geschatte gemiddelde ticketprijs van het directe alternatief hoger ligt dan van het indirecte alternatief.

- Het geschatte effect van de hubcarrier is gelijk aan $\exp(5,49 + 0,148) - \exp(5,49) = \text{€} 38$, voor het voorbeeld van het directe alternatief gaat het dus om een effect op de geschatte ticketprijs van circa 15 procent.
- Het geschatte effect van de intercontinentale dummy is gelijk aan $\exp(5,49 + 0,081) - \exp(5,49) = \text{€} 20$, oftewel circa 8 procent.
- Het geschatte effect van de marktstructuur monopolie is gelijk aan $\exp(5,49 + 0,105) - \exp(5,49) = \text{€} 27$, oftewel circa 11 procent.
- Het geschatte effect van een SkyTeam, in plaats van een low cost carrier, vlucht is gelijk aan $\exp(5,49 + 0,344) - \exp(5,49) = \text{€}100$, oftewel circa 40 procent. Ook voor de andere allianties ligt de gemiddelde geschatte ticketprijs, *ceteris paribus*, hoger dan die van de low cost carriers. Voor Star Alliance geldt bijvoorbeeld een verschil van € 150.

Tot slot illustreren we de werking van de lineair spline schatter. Als we naar dezelfde vlucht kijken maar enkel de vluchttijd aanpassen naar 8 uur, dan geldt de volgende gemiddelde geschatte prijs voor de directe verbinding:

$$\ln(\text{fare}) = 3,803 - 0,151(\text{Direct}) + 0,208 * 5,6\text{uur} + 0,162 * 2,4\text{uur} + 0,118 * (5,6 \text{ uur} \cdot \text{Direct}) - 0,0338$$

$$* (2,4 \text{ uur} \cdot \text{Direct}) + 0,343 (EU1) + 0,190(EU2) = 6,31$$

$$\text{fare} = \exp(6,31) = \text{€}555$$

Voor de indirecte verbinding geldt weer een lagere gemiddelde geschatte ticketprijs:

$$\ln(\text{fare}) = 3,803 + 0,208 * 5,6\text{uur} + 0,162 * 2,4\text{uur} + 0,343 (EU1) + 0,190(EU2) = 5,9$$

$$\text{fare} = \exp(5,9) = \text{€}361$$

Voor het zakelijke segment (de duurdere tickets) zijn de ticketprijzen op deze voorbeeldroute *ceteris paribus* dus circa dertig procent goedkoper voor indirecte alternatieven.

Van modelschatting naar geschatte prijzen voor alle alternatieven

Om te komen tot de finale set aan geschatte prijzen op basis van het empirische model rekenen we voor elke route in het Level of Service-bestand de geschatte prijs uit op dezelfde wijze als hierboven gedemonstreerd. Het Level of Service-bestand inclusief de schatters voor elk van de vier modellen over alle alternatieven is aan Rijkswaterstaat geleverd.

Appendix A Regiodummies

Tabel A.1 Regiodummies

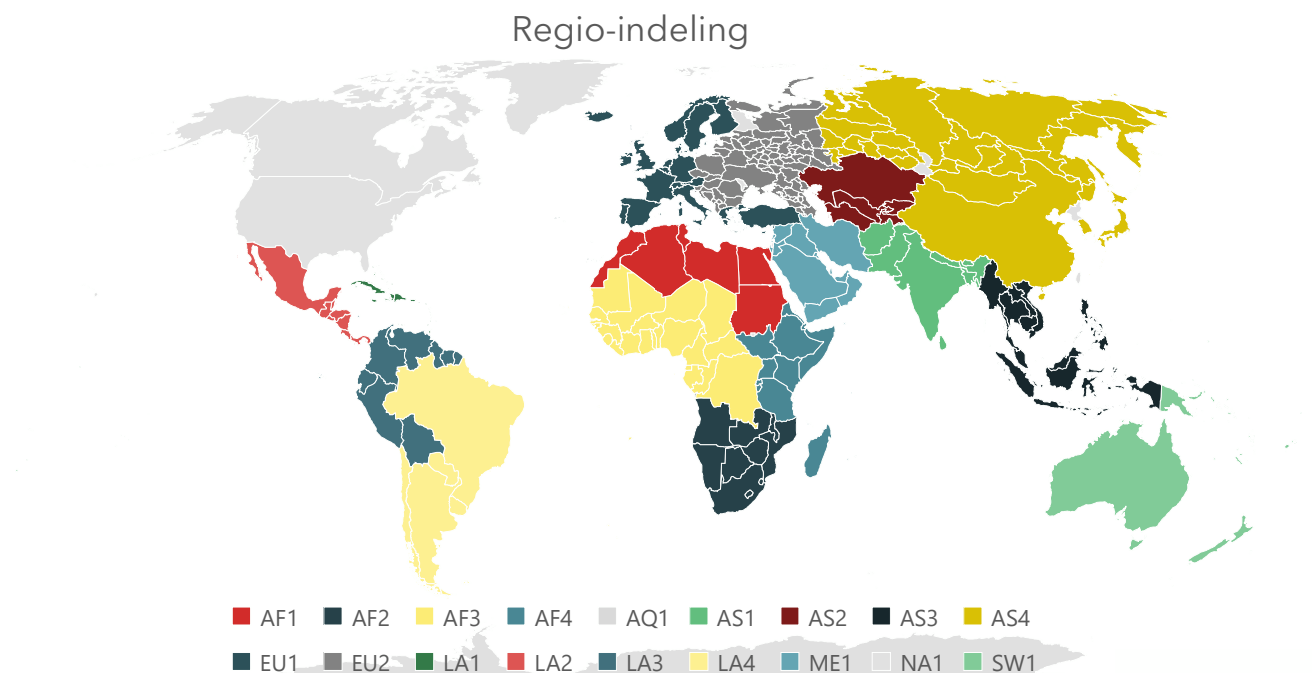
	Zakelijk 1	Zakelijk 2	Niet-zakelijk 1	Niet-zakelijk 2
Parameter schattingen overige verklarende variabelen	Zie Tabel 2 in hoofdttekst			
Herkomst				
Afrika 1	0 (.)	0 (.)	0 (.)	0 (.)
Afrika 2	0,0754 (0,138)	0,0755 (0,138)	0,182* (0,0957)	0,181* (0,0956)
Afrika 3	0,394*** (0,129)	0,394*** (0,129)	0,394*** (0,100)	0,394*** (0,100)
Afrika 4	-0,0619 (0,177)	-0,0619 (0,177)	0,0841 (0,126)	0,0840 (0,126)
Azië 1	-0,0550 (0,117)	-0,0550 (0,117)	0,0877 (0,0969)	0,0881 (0,0970)
Azië 2	0,118 (0,134)	0,118 (0,134)	0,311*** (0,116)	0,311*** (0,116)
Azië 3	0,176 (0,136)	0,176 (0,136)	0,248** (0,0990)	0,248** (0,0991)
Azië 4	0,361*** (0,128)	0,361*** (0,128)	0,398*** (0,0971)	0,398*** (0,0971)
Europa 1	0,343*** (0,120)	0,343*** (0,120)	0,396*** (0,0891)	0,396*** (0,0891)
Europa 2	0,200* (0,120)	0,200* (0,120)	0,330*** (0,0902)	0,331*** (0,0902)
Latijns-Amerika 1	0,462*** (0,126)	0,462*** (0,126)	0,594*** (0,0924)	0,594*** (0,0924)
Latijns-Amerika 2	0,403*** (0,120)	0,403*** (0,120)	0,517*** (0,0926)	0,517*** (0,0926)
Latijns-Amerika 3	0,115 (0,125)	0,115 (0,125)	0,435*** (0,0965)	0,434*** (0,0965)
Latijns-Amerika 4	0,475*** (0,116)	0,475*** (0,116)	0,536*** (0,100)	0,536*** (0,100)
Midden-Oosten	0,562*** (0,134)	0,562*** (0,134)	0,520*** (0,104)	0,520*** (0,104)
Noord-Amerika	0,430*** (0,113)	0,430*** (0,113)	0,484*** (0,0886)	0,484*** (0,0886)
Zuidoost-Azie	0,457*** (0,119)	0,457*** (0,119)	0,653*** (0,0955)	0,653*** (0,0956)
Bestemming				
Afrika 1	0 (.)	0 (.)	0 (.)	0 (.)
Afrika 2	0,241** (0,120)	0,241** (0,120)	0,176*** (0,0516)	0,176*** (0,0517)
Afrika 3	0,409*** (0,104)	0,409*** (0,104)	0,306*** (0,0700)	0,306*** (0,0700)
Afrika 4	0,0726 (0,0987)	0,0726 (0,0987)	0,0249 (0,0719)	0,0247 (0,0720)
Azië 1	-0,160* (0,0852)	-0,160* (0,0852)	-0,123*** (0,0413)	-0,123*** (0,0415)
Azië 2	0,133 (0,121)	0,133 (0,121)	0,176** (0,0855)	0,176** (0,0856)
Azië 3	0,146* (0,0873)	0,146* (0,0873)	0,150*** (0,0420)	0,150*** (0,0420)
Azië 4	0,369*** (0,0814)	0,369*** (0,0814)	0,335*** (0,0365)	0,335*** (0,0366)

Europa 1	0,395*** (0,0760)	0,395*** (0,0760)	0,370*** (0,0346)	0,369*** (0,0346)
Europa 2	0,190** (0,0758)	0,190** (0,0758)	0,276*** (0,0421)	0,276*** (0,0421)
Latijns-Amerika 1	0,368*** (0,113)	0,368*** (0,113)	0,477*** (0,0484)	0,476*** (0,0487)
Latijns-Amerika 2	0,358*** (0,0885)	0,358*** (0,0885)	0,431*** (0,0423)	0,431*** (0,0424)
Latijns-Amerika 3	0,0879 (0,0924)	0,0879 (0,0924)	0,296*** (0,0525)	0,295*** (0,0525)
Latijns-Amerika 4	0,445*** (0,0881)	0,445*** (0,0881)	0,480*** (0,0553)	0,480*** (0,0554)
Midden-Oosten	0,650*** (0,0701)	0,650*** (0,0701)	0,498*** (0,0412)	0,498*** (0,0412)
Noord-Amerika	0,480*** (0,0822)	0,480*** (0,0822)	0,463*** (0,0354)	0,462*** (0,0355)
Zuidoost-Azië	0,526*** (0,0824)	0,526*** (0,0824)	0,610*** (0,0504)	0,609*** (0,0504)
Adjusted r-kwadraat	0,696	0,696	0,632	0,632
AIC	143.454	143.452	237.314	237.320
Aantal observaties	99.125	99.125	208.853	208.853

Bron: SEO Economisch Onderzoek

Noot: * $p < 0,1$, ** $p < 0,05$, *** $p < 0,01$, Cluster robuuste standaardfouten (vertrekluchthaven) tussen haakjes. De observaties (route alternatief) zijn gewogen naar het aantal reizigers.

Figuur A.1 De regio-indeling zoals gehanteerd in OAG



Bron: SEO Economisch Onderzoek