

Dynamisch spitsgedrag van automobilisten

Bestrijding van filedruk is een van de centrale uitdagingen voor beleidsmakers. Daarbij verdient het dynamische model van spitsverkeer extra aandacht. Regulerende heffingen volgens dit model hebben een gunstiger effect op het gedrag van weggebruikers dan heffingen volgens het veel gehanteerde statische model van Pigou. Ook blijkt dat ICT-oplossingen niet moeten worden overschat.

**ERIK
VERHOEF**

Hoogleraar aan de
Vrije Universiteit
Amsterdam

Een van de belangrijkste grondleggers van de huidige milieu-economie is ongetwijfeld de Cambridge-econoom Arthur Pigou. Zijn theorievorming rond externe kosten, en de economische wenselijkheid van belastingen voor het corrigeren van dit marktfalen, vormt het economisch fundament onder het principe van regulerende milieuheffingen. Toen hij dit idee in 1920 voor het eerst publiceerde, dacht hij overigens niet aan het milieu, maar aan het terugdringen van files.

Inmiddels vormt verkeerscongestie, nu veel meer dan destijds, een uitdaging voor het beleid. Het is daarnaast nog steeds een aansprekende toepassing om de economische principes van regulering van externe effecten te illustreren. Verkeerscongestie kan worden ge-

bruikt om te laten zien hoe belangrijk het is het 'juiste' gedragsmodel te kennen voor het formuleren van beleid gericht op marktfalen, en voor het voorspellen van de welvaartswinsten die daarmee gerealiseerd kunnen worden. We doen dit door de beleidsconclusies van statische en dynamische modellen van verkeerscongestie met elkaar te vergelijken. Dynamische modellen verschillen onder meer van de statische doordat ze tijdstipkeuze expliciet beschrijven, en daardoor ook rekening houden met aanpassingen van deze tijdstippen als er nieuw beleid wordt ingevoerd.

Tijdstipkeuze is op zich al van belang voor het beleid, al is het maar omdat het cruciaal is voor het maken van betrouwbare voorspellingen van de effecten van verschillende maatregelen om spitsdrukke het hoofd te bieden. En het is juist de problematiek rond verkeer op drukke tijden en plaatsen waar in het beleid veel aandacht naar uitgaat. Daarnaast biedt vergelijking van statische en dynamische modellen een goede gelegenheid om in te gaan op de algemenere economische vraag hoe het negeren van relevante *margins of behaviour*, ofwel de verschillende aspecten waarover keuzes worden gemaakt, tot vertekeningen in beleidsmodellen en -adviezen kan leiden.

Dit is van belang omdat het in werkelijkheid slechts zelden het geval is dat consumenten en producenten alleen maar over hoeveelheden (en de bijbehorende prijzen) nadenken, zoals in de tweedimensionale plaatjes van marktevenwichten en -optima. Beleids-

conclusies en adviezen kunnen drastisch wijzigen als we er rekening mee houden dat weggebruikers niet alleen over het aantal spitsritten keuzes maken, maar ook over het moment van die ritten binnen de spits. Het negeren daarvan kan tot forse fouten leiden. Het onderstreept daarmee het belang van het goed kennen van de relevante gedragsmarges voor het verrichten van economisch beleidsonderzoek.

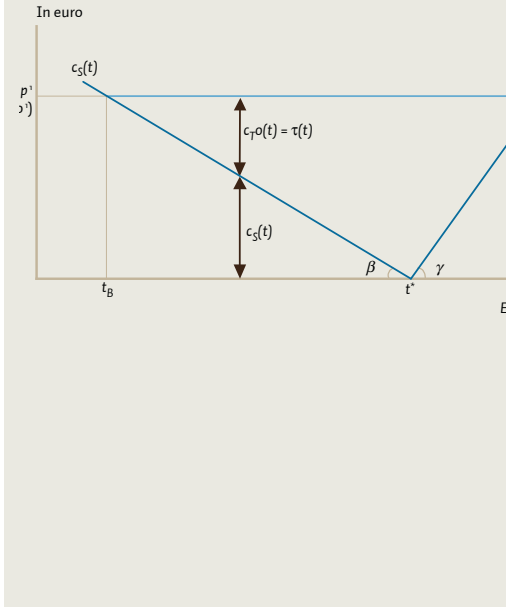
VERKEERSCONGESTIE

Het economische basismodel van verkeerscongestie wordt in figuur 1 samengevat. Het is een statisch model, waarin de marktkuitkomst wordt uitgedrukt in de verkeersstroom – het aantal voertuigen dat per uur een bepaald punt langs de weg passeert, en daarmee het aantal ritten dat per tijdseenheid wordt voltooid. Als congestie relevant is, daalt de snelheid en stijgt de reistijd T met het aantal weggebruikers. Als we andere kosten, zoals brandstof, negeren omdat ze de analyse niet wezenlijk beïnvloeden, en die reistijd met de zogeheten *value of time* – de reistijdwaardering, α – vermenigvuldigen, krijgen we een stijgende gemiddelde kostenfunctie, $gk = \alpha \cdot T$. Het vrijemarktevenwicht F^0 is daar waar de inverse vraagfunctie d – zoals gebruikelijk met hoeveelheid langs de horizontale as – de verticale as – de functie gk snijdt; de zogeheten gegeneraliseerde prijs p^0 is daar gelijk aan de marginale betalingsbereidheid (d) en bestaat louter uit tijdskosten (gk^0). Omdat de gemiddelde kosten stijgen met het aantal weggebruikers, liggen de marginale kosten mk boven gk ; het verschil zijn de marginale externe kosten mek . Het optimale niveau van weggebruik ligt bij het snijpunt van d en mk ; bij F^1 dus. Dat het verschil tussen mk en gk precies de waarde van de marginale externe kosten geeft, volgt ook uit het toepassen van de productregel bij het differentiëren van de totale kosten $F \cdot gk(F)$. Dat levert op $mk = gk + F \cdot gk'$. De tweede term geeft het aantal andere gebruikers (F) maal het effect op de kosten per gebruiker van een extra eenheid verkeer (gk') en geeft daarmee inderdaad de mek .

Pigou zag in dat dit optimum bereikt kan worden door de optimale tol τ in te voeren, gelijk aan de mek in het optimum. Hiermee wordt een toename in het maatschappelijk surplus gerealiseerd dat gegeven wordt door de grijze 'driehoek': de afname in kosten (het gebied onder mk tussen F^1 en F^0), minus de afname in baten (onder d). Pigou gebruikte zo het voorbeeld van een weg met verkeerscongestie om te laten zien dat de kosten aan reistijdverliezen, die een extra

Optimale heffing in het statische model

FIGUUR 1



weggebruiker door lagere snelheid op de weg voor alle andere gebruikers veroorzaakt, voor een efficiënte marktkuitkomst geïnternaliseerd moeten worden.

Het statische model heeft een aantal redelijk intuïtieve beleidsconclusies. Een eerste is dat de gegeneraliseerde prijs p toeneemt als een tol wordt ingevoerd: de waarde van tijdswinsten per gebruiker (de daling in gk) weegt niet volledig op tegen de te betalen tol τ ; anders kan F ook niet afnemen. Een tweede conclusie is een gevolg van de eerste: iedereen, behalve de overheid, is slechter af door invoering van de optimale tol (bij heterogeniteit in tijdswaardering kan deze conclusie overigens veranderen). En een derde conclusie is dat weggebruikers wel een alternatief (bijvoorbeeld reizen buiten de spits, ov, thuiswerken, alternatief vervoer) moeten hebben om daadwerkelijk een welvaartswinst te bereiken: bij een perfect inelastische vraag wordt d vertikaal en wordt het oppervlak van de grijze driehoek nul.

DYNAMISCH MODEL

In werkelijkheid is spitsdrukke op de weg natuurlijk een dynamisch fenomeen. In plaats van een afgemeten spitsperiode met constante verkeerscondities volgens

De auteur heeft verklaard dit artikel alleen te publiceren in ESB en niet elders te publiceren in wat voor medium dan ook. Het is wel toegestaan om het artikel voor eigen gebruik en voor publicatie op een intranet van de werkgever van de auteur aan te wenden.

figuur 1, en géén congestie daarbuiten, zien we graduele toe- en afnames van verkeersdrukte. Alleen daarom al verdient een dynamische modellering de voorkeur: het is realistischer. Bovendien blijkt dat, op het moment dat met tijdstipaanpassingen rekening wordt gehouden, de conclusies van economische beleidsanalyses kunnen veranderen.

Voordat we daar dieper op ingaan, is het goed te benadrukken dat diverse studies het empirische belang van tijdstipaanpassingen als reactie op tijdstipafhankelijke heffingen onderstrepen. We noemen er twee. Voor vermoedelijk de eerste toepassing van tijdsafhankelijke heffingen in Singapore in 1975 waren de tijdstip-effecten dermate groot, dat vermoed mag worden dat de heffing te hoog was: een afname van het aantal cordon-passages in de *restricted period* met 73 procent en een toename daarbuiten met 23 procent (Santos *et al.*, 2004).

Recenter en dichter bij huis zijn ervaringen met diverse experimenten met ‘Spitsmijden’ in Nederland (Spitsmijden Consortium, 2007), waarbij het vermijden van de spits door automobilisten beloond wordt. Bijvoorbeeld bij de eerste van deze projecten, gehouden in 2006 op de A12, werd grofweg drie vijfde van de spitsmijdingen gerealiseerd door vóór de spits met de auto te rijden, één vijfde door dat na de spits te doen, en één vijfde door met het ov te reizen (omdat de percentages verschillen over bijvoorbeeld hogere en lagere beloningen, houden we het hier bij een grove indicatie).

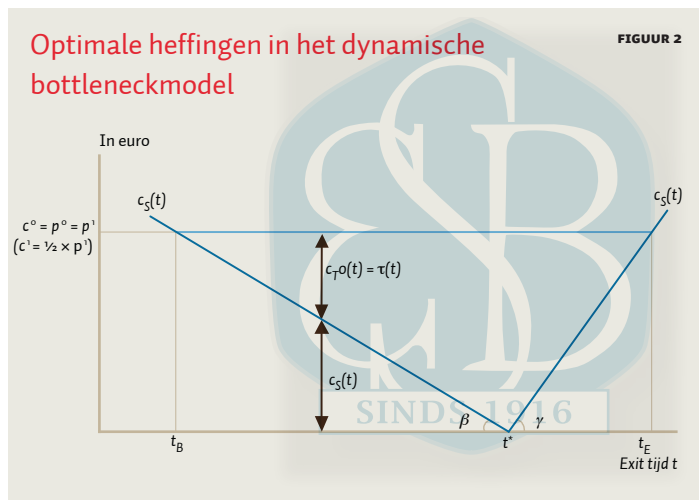
En hiermee in overeenstemming, in een overzicht van het Ministerie van Infrastructuur en Milieu (2011) van verschillende recente projecten wordt gezegd, dat bijna 75 procent van alle gedragsveranderingen tijdstipaanpassingen betreft. Kortom, tijdstipaanpassingen blijken potentieel belangrijk te zijn.

HET BOTTLENECKMODEL VAN VICKREY

De latere Nobelprijswinnaar William Vickrey presenteerde in 1969 een eenvoudig maar overtuigend dynamisch economisch evenwichtsmodel van verkeerscongestie, dat nog steeds veel gebruikt wordt. Essentieel is dat in het model tijdstipkeuze binnen de spits endogeen is, en tot stand komt door een afruil tussen de *schedule delay cost of tijdstipkosten* c_s , en de reistijdkosten c_T (figuur 2). De eerstgenoemde is nieuw ten opzichte van het statische model, en geeft de kosten die gemoed zijn met het eerder of later op de bestemming (zeg: het werk) aankomen dan het gewenste tijdstip t^* . Door daaraan een constante kost per minuut toe te kennen, met wel verschillende waarden β voor vroege en γ voor late afwijkingen, ontstaat de gesegmenteerde-lineaire functie c_s die in figuur 2 getekend staat, met hellingen $-\beta$ en $+\gamma$.

Als door de beperkte wegcapaciteit niet iedereen precies om t^* kan aankomen, zullen reistijden $T(t)$ in het evenwicht zodanig over aankomsttijden moeten variëren dat de reistijdkosten $c_T(t) = \alpha \cdot T(t)$ precies compenseren voor verschillen in c_s (waarbij α weer de reistijdwaardering geeft). Alleen dan heeft geen enkele weggebruiker nog een prikkel om zijn of haar aankomsttijdstip t te veranderen. Dit is het evenwicht dat getekend is in figuur 2, aangegeven met superscript 0. Het gegeneraliseerde kostenniveau zonder heffingen, c^0 , is voor iedereen even hoog en bestaat uit de som van c_s en c_T .

Om het model volledig te maken, dienen nog enkele zaken gespecificeerd te worden. Eén daarvan is de vraagfunctie, en daarvan veronderstelde Vickrey dat deze perfect inelastisch is: het aantal weggebruikers N is gegeven. Een ander is de congestiefunctie. Hiervoor gebruikte hij pure *bottleneckcongestie*: reistijdverliezen bestaan daarbij uit wachten vóór een knelpunt, en het aantal auto's in de wachtrij verandert met een groeivoet gelijk aan de stroom van ‘aansluiters’ achter aan de rij, minus de capaciteit s waarmee het knelpunt auto's laat uitstromen. Als er op het moment van aansluiten in de rij een wachtrij van $Q(t)$ auto's staat, is het tijdverlies door de file $T(t) = Q(t)/s$. Voor het gemak kunnen alle



De auteur heeft verklaard dit artikel alleen te publiceren in ESB en niet elders

te publiceren in wat voor medium dan ook. Het is wel toegestaan om het artikel voor eigen gebruik en voor publicatie op een intranet van de werkgever van de auteur aan te wenden.

andere (constant veronderstelde) reistijden genegeerd worden; de reistijd $T(t)$ geeft dus alleen het tijdverlies door wachten in de file.

Om het evenwichtspatroon van reistijdverliezen $T(t)$ te krijgen waarbij de tijdsvariatie in c_T precies compenseert voor die in c_s , moet dan voor vroege aankomsten de toestroom richting de file constant zijn en hoger dan de capaciteit, zodat de file lineair over de tijd groeit, terwijl deze voor late aankomsten constant en lager dan de capaciteit moet zijn, zodat de file lineair afneemt. Omdat dan de aankomststroom op de bestemming s constant over de tijd is, laat figuur 2 zien dat in het evenwicht de totale *schedule delay*-kosten (s maal het oppervlak onder c_s tussen het begin en het eind van de spits, t_b en t_e) gelijk is aan de totale reistijd-kosten (s maal het oppervlak tussen de horizontale lijn op c^0 en c_s over dezelfde periode). Met andere woorden, het beschouwen van alleen de twee laatstgenoemde, zoals nogal eens gebeurt als het over de kosten van files gaat, geeft dan een aanzienlijke onderschatting van die kosten.

HALVERING KOSTEN

Omdat reistijdverliezen in dit model een puur (*dead-weight*) verlies zijn, is het optimum de situatie waarin de toestroom naar het knelpunt constant over de tijd is, en gelijk aan s . De totale kosten worden dan gehalveerd: de *schedule delay*-kosten blijven bestaan, maar de reistijdverliezen zijn verdwenen. Zonder aanvullend beleid is dat natuurlijk geen evenwicht, omdat dan de aankomst op t^* aantrekkelijker zou zijn dan een aankomst op elk ander mogelijk tijdstip: de gegeneraliseerde prijs is niet constant over de tijd.

Vickrey liet zien dat een tijdsvariërende heffing τ , die op elk moment gelijk is aan wat de waarde van de reistijdverliezen in het oorspronkelijke evenwicht op dat moment was, dit optimum ook als evenwicht voor de gebruikers ondersteunt; de gegeneraliseerde prijs, die nu bestaat uit *schedule delay*-kosten plus tol, is dan weer constant over de tijd.

De beleidsconclusies van dit model contrasteren sterk met die van het statische model. In de eerste plaats neemt de gegeneraliseerde prijs p niet toe door invoering van de optimale heffing: die heffing heeft dezelfde waarde als de reistijdverliezen hadden die in het optimum zijn verdwenen. In de tweede plaats zijn de weggebruikers niet slechter af door invoering van de tol. En in de derde plaats is de omvang van de welvaartswinsten onafhankelijk van de vraagelasticiteit.

Dat wil zeggen: bij inelastische vraag N zijn er al forse welvaartswinsten (de reiskosten worden gehalveerd), en als de vraag elastisch zou zijn, zouden deze niet veranderen door invoering van de optimale heffing, juist omdat de gegeneraliseerde prijs en dus het totaal aantal gebruikers niet verandert.

Twee veelgehoorde bezwaren tegen congestieheffingen zijn volgens dit model feitelijk irrelevant. Een eerste is dat alternatieven voor weggebruik, zoals het ov, onvoldoende aantrekkelijk zijn of niet genoeg capaciteit hebben. In dit model is het optimale aantal weggebruikers hetzelfde als zonder congestieheffingen, en is er dus ook geen alternatief voor weggebruik nodig om het optimum te bereiken. Een tweede bezwaar is dat mensen nu eenmaal op een bepaald moment op hun werk moeten zijn, en dus het tijdstip niet kunnen aanpassen. In dit model is de uitstroom van het knelpunt in zowel het evenwicht als het optimum gelijk aan de capaciteit s , en vindt het gedurende hetzelfde tijdsvenster tussen t_b en t_e plaats. Aankomsttijden hoeven dus ook niet te veranderen: het zijn alleen de vertrektijden van huis die aangepast worden.

CAPACITEIT EN MOBIELE TELEFOONS

Het dynamische model heeft niet alleen gevolgen voor de beleidsanalyse van prijsbeleid. Small en Verhoef (2007) bekeken bijvoorbeeld welke fout een beleidsanalist maakt als hij de baten van capaciteitsuitbreiding probeert te berekenen wanneer hij zich ten onrechte niet realiseert dat hij te maken heeft met een dynamisch knelpunt zoals beschreven wordt door het bottleneckmodel van Vickrey. De beleidsanalist analyseert alsof het een statisch model betreft, waarbij alleen de reistijdverliezen worden meegenomen, tijdstipaanpassingen worden genegeerd, en hij de meest gebruikte statische congestiefunctie hanteert: de BPR-functie (*Bureau of Public Roads*). Deze functie veronderstelt dat reistijdverliezen via een *power law* afhangen van de ratio weggebruik/capaciteit. De relatieve fout die de analist maakt, blijkt dan proportioneel te zijn met de gebruikte exponent in de BPR-congestiefunctie. Voor de gebruikelijke waarde 4 overschat hij de marginale baten van capaciteitsuitbreiding met een factor 2, terwijl hij de marginale baten met vijftig procent onderschat als hij de exponent 1 gebruikt; een lineaire functie dus.

Een ander voorbeeld betreft de analyse van 'ICT aan boord' die de waarde van reistijdverliezen laat dalen, omdat reistijd immers beter besteed kan wor-

den. Mobile telefoons zijn hiervoor natuurlijk een goed voorbeeld.

In het statische model laat een afname van de tijdswaardering beide kostencurves gk en mk naar beneden schuiven en minder steil worden: dezelfde reistijdverliezen leveren immers lagere gebruikerskosten op. Het gevolg is dat zowel het vrijemarktevenwicht als het optimum bij hogere gebruikersniveaus en lagere kosten komt te liggen, en dat de welvaartswinst van optimale heffingen normaal gesproken kleiner zal worden (tenzij de inverse vraagfunctie sterk convex is). ICT aan boord zou dus een goede mogelijkheid kunnen zijn de maatschappelijke kosten van files te verlagen.

In het dynamische model, daarentegen, verandert er aan de reiskosten helemaal niets als de tijdswaardering α lager wordt. Dat wil zeggen, de lengte van de wachtrij in het evenwicht zal op elk moment in dezelfde proportie toenemen als dat α is afgenomen, omdat het product van wachttijden en tijdswaardering, $c_t(x) = \alpha \cdot T(x)$, nog steeds zal moeten compenseren voor verschillen in *schedule delay*-kosten over de tijd. ICT aan boord zal dus de maatschappelijke kosten van files niet kunnen verminderen. Op zijn hoogst kunnen tijdelijk reizigers profiteren die het wél hebben terwijl anderen het niet hebben. Maar, als uiteindelijk iedereen het heeft, zijn de reiskosten voor iedereen weer even hoog als ze oorspronkelijk waren, en zullen dus ook de potentiële welvaartswinsten van optimale prijzen weer even hoog zijn geworden.

LITERATUUR

- Ministerie van Infrastructuur en Milieu (2011) *Resultaten Mobiliteitsprojecten*. Den Haag: Ministerie van IenM.
- Pigou, A.C. (1920) *The economics of welfare* (first edition). Londen: Macmillan.
- Santos, G., W.W. Li en W.T.H. Koh (2004) Transport policies in Singapore. In: Santos, G., *Road pricing: theory and evidence*, Amsterdam: Elsevier, 209–235.
- Small, K. en E.T. Verhoef (2007) *The economics of urban transportation*. Londen: Routledge.
- Spitsmijden Consortium (2007) *Effecten van Belonen*. Den Haag: Spitsmijden Consortium.
- Vickrey, W.S. (1969) Congestion theory and transport investment. *American Economic Review*, 59(2), 251–260.

CONCLUSIE

Tijdstipkeuze is een belangrijk element in het spitsgedrag van weggebruikers. Het niet meenemen daarvan in beleidsanalyses kan tot substantiële vertekeningen leiden in de conclusies. Waar het statische model bijvoorbeeld zegt dat invoering van optimale congestieheffingen leidt tot hogere gegeneraliseerde kosten van de gebruiker en geen effecten zal hebben bij een perfect inelastische vraag, voorspelt het dynamische model gelijkblijvende gegeneraliseerde kosten en zijn de welvaartswinsten onafhankelijk van de vraagelasticiteit. Het gebruik van het statische model bij kosten-batenanalyses van wat in werkelijkheid dynamische congestie is – bepaald niet ongebruikelijk in de praktijk – kan tot zowel een substantiële onder- als overschatting leiden van de marginale baten van capaciteitsuitbreiding. En waar het statische model voorspelt dat ICT aan boord de kosten van congestie kan laten dalen doordat tijd beter gebruikt kan worden, maakt het dynamische model duidelijk dat er vervolgens tijdstipaanpassingen zullen plaatsvinden die dergelijke winsten volledig teniet zullen doen.

Het dynamische model levert een gunstiger beeld op van de potentie van prijsbeleid ter bestrijding van congestie dan het statische model dat Pigou zelf impliciet in gedachten had. Daarbij moet wel bedacht worden dat die gunstige effecten zijn afgeleid voor optimale, over de tijd variërende heffingen. Met name voor de regulering van congestie bij knelpunten – files dus – is het model het best toepasbaar, en de wenselijkheid van dynamische heffingen daarmee het meest aannemelijk. Tegelijkertijd zijn dat de situaties waar de baten van ICT-oplossingen niet overschat moeten worden vanwege mogelijke *rebound-effecten* door tijdstipaanpassingen, en waar het voor kosten-batenanalyses van capaciteitsuitbreidingen heel belangrijk is om het juiste gedragsmodel te gebruiken.

Het belang van het op de juiste wijze modeleren van *scheduling*-gedrag in de analyse van beleid gericht op verkeerscongestie is daarmee afdoende onderbouwd. Maar dat ook daadwerkelijk doen is in de huidige praktijk nog een zeldzaamheid.