



Paper

Prijsindex Nieuwbouw Koopwoningen

Methodebeschrijving

Juli 2018

Inhoud

1. Inleiding 3

- 1.1 Aanleiding 3
- 1.2 Doel PNK 3
- 1.3 Afkortingen 3

2. Berekeningsmethode PNK 4

- 2.1 Hedonische regressiemethode 4
- 2.2 Kenmerken van hedonische regressiemethode 5
- 2.3 Data 5
- 2.4 Uitwerking PNK 7

3. Berekeningsmethode aantal nieuwbouwtransacties 8

- 3.1 Samenwerking CBS-Kadaster 8
- 3.2 Data 8
- 3.3 Methode afleiding nieuwbouwtransacties 8
- 3.4 Berekening aantallen en totale verkoopwaarde 9
- 3.5 Verkochte nieuwbouwkooptoningen t.o.v. opgeleverde nieuwbouwwoningen 9

Referenties 10

1. Inleiding

1.1 Aanleiding

Voor het CBS geldt een verplichting (EU-verordening) om een prijsindex voor nieuwbouwoopwoningen (PNK) te publiceren. Het was daarom noodzakelijk om een prijsindex voor nieuwbouwoopwoningen te ontwikkelen. In dit document wordt beschreven hoe de nieuwe PNK wordt berekend voor Nederland totaal, welke voldoet aan de richtlijnen vanuit Eurostat [1], het statistische bureau van de Europese Unie.

1.2 Doel PNK

De Prijsindex Nieuwbouw Koopwoningen (PNK) heeft als doel het meten van de prijsontwikkeling per kwartaal van verkochte nieuwbouwwoningen in Nederland. De woning moet gebouwd zijn op Nederlands grondgebied en moet verkocht zijn aan een particulier. Ook is het belangrijk dat de woning is of wordt opgericht op een nieuw fundament en voor de eerste keer wordt toegevoegd aan de woningvoorraad. Huur-, renovatie- en transformatiewoningen en woningsplitsingen behoren in dit geval niet tot de populatie. Ook geldt dat er geen sprake moet zijn van particulier opdrachtgeverschap (ook wel zelfbouw genoemd). De PNK wordt opgenomen in de OOH (Owner Occupied Housing) en vanuit de richtlijnen van Eurostat voor OOH wordt een onderscheid gemaakt tussen nieuwbouwoopwoningen en zelfbouw.¹⁾

Gebruiksdoelen van PNK:

- Opname in de OOH. De OOH brengt de kostenontwikkelingen van huizenbezitters in kaart.
- Opname in de House Price Index (HPI). De HPI brengt de prijsontwikkelingen van nieuwbouwoopwoningen en bestaande koopwoningen (PBK) in kaart.
- Zelfstandige publicatie op Statline. De tabel wordt aangevuld met waarden en aantallen van nieuwbouwtransacties.

1.3 Afkortingen

De volgende afkortingen zijn in deze notitie gebruikt:

HDIF	Hedonische Double Imputation Fisher
HDIL	Hedonische Double Imputation Laspeyres
HDIP	Hedonische Double Imputation Paasche
HPI	House Price Index
OOH	Owner Occupied Housing
PNK	Prijsindex Nieuwbouw Koopwoningen

¹⁾ Voor meer informatie over het onderscheid tussen nieuwbouwoopwoningen en zelfbouw, zie ook Technical manual on Owner-Occupied Housing [2]

2. Berekeningsmethode PNK

2.1 Hedonische regressiemethode

De PNK wordt berekend op basis van de hedonische regressiemethode. Dit is een meervoudige regressieanalyse, waarbij meerdere onafhankelijke variabelen van invloed zijn op de afhankelijke variabele. Voor de PNK betekent dit dat de samenhang wordt berekend tussen de verkoopprijs van de woning en verschillende woningkenmerken. Hierdoor kan de prijsontwikkeling gecorrigeerd worden voor kwaliteitsverschillen tussen de verkochte woningen in opeenvolgende perioden.

Het volgende regressiemodel wordt gebruikt om een hedonische prijsindex te berekenen:

$$\log(p_i^t) = \beta_0^t + \beta_1^t x_{i1} + \dots + \beta_k^t x_{ik} + \varepsilon_i^t$$

In dit geval is $\log(p_i^t)$ in periode t het natuurlijke logaritme van de verkoopprijs en x_{ik} is het k e kenmerk van de i e woning die verkocht is in periode t . Welke kenmerken in het model worden opgenomen is bepaald aan de hand van hun bijdrage aan het verklaren van de prijs. Alleen significant bevonden kenmerken zijn opgenomen.

Op basis van de coëfficiënten die de regressieanalyse oplevert, wordt een verkoopprijs geschat. De geschatte verkoopprijs is vervolgens input voor het berekenen van een prijsontwikkeling. In dit onderzoek is gekozen om de verkoopprijs van zowel de referentieperiode als van de verslagperiode te schatten. Deze wijze van hedonische regressie wordt ook wel de double imputation-methode genoemd [1].

De double imputation-methode kent drie varianten van hedonische regressie: Laspeyres, Paasche en Fisher. Alle drie zijn gebaseerd op dezelfde bovenstaande regressie, maar verschillen in hoe de prijsindex hier vervolgens uit wordt geconstrueerd. Gezien de resultaten van het onderzoek bleek de double imputation-methode van Fisher het meest geschikt, om een index voor Nederland totaal te berekenen. Deze index is het geometrische gemiddelde van de Laspeyres-index en de Paasche-index.

De volgende formule is een weergave van de Hedonische Double Imputation Laspeyres Index (I_{HDIL}):

$$I_{HDIL}(t,0) = \frac{\exp\left(\sum_{k=1}^K \hat{\beta}_k^t \bar{x}_k^0\right)}{\exp\left(\sum_{k=1}^K \hat{\beta}_k^0 \bar{x}_k^0\right)}$$

De volgende formule is een weergave van de Hedonische Double Imputation Paasche Index (I_{HDIP}):

$$I_{HDIP}(t,0) = \frac{\exp\left(\sum_{k=1}^K \widehat{\beta}_k^t \bar{x}_k^t\right)}{\exp\left(\sum_{k=1}^K \widehat{\beta}_k^0 \bar{x}_k^t\right)}$$

De volgende formule is een weergave van de Hedonische Double Imputation Fisher Index (I_{HDIF}):

$$I_{HDIF}(t,0) = (I_{HDIL}(t)I_{HDIP}(t))^{1/2}$$

In deze formules geeft $\widehat{\beta}_k^t$ voor de geschatte coëfficiënt weer voor het k^e kenmerk in periode t , \bar{x}_k^0 staat voor het gemiddelde van het k^e kenmerk in periode 0, \bar{x}_k^t staat voor het gemiddelde van het k^e kenmerk in periode t en K geeft het aantal woningkenmerken weer dat is gebruikt in het hedonisch model.

Middels de hedonische methode wordt een verkoopprijs geschat. Om te controleren of deze geschatte prijs dicht bij de werkelijke prijs ligt en of het meervoudige regressiemodel dus geschikt is om een index te berekenen, zijn verschillende aannames getoetst. Onder andere is onderzocht of er sprake is van een lineair verband tussen verkoopprijs en de woningkenmerken en of de woningkenmerken onderling niet te veel correleren. Om de index verder te kunnen verbeteren zijn met behulp van Cook's distance uitschieterende waarden (outliers) bepaald. De prijsindex wordt berekend op basis van de data zonder deze uitschieterende waarden.

2.2 Kenmerken van hedonische regressiemethode

Een belangrijk kenmerk van een double imputation hedonische regressie-methode is dat een prijsontwikkeling berekend kan worden zonder een woning te volgen in de tijd of identieke woningen met elkaar te vergelijken. In het eerste geval is een woning geen nieuwbouwwoning meer als deze opnieuw wordt verkocht. Het betreft dan een bestaande koopwoning. In het tweede geval is het bijna onmogelijk om woningen in de tijd te vergelijken met exact dezelfde woningkenmerken, zoals exact dezelfde locatie, woningtype, enzovoorts. Bij een hedonische prijsindex wordt namelijk gecorrigeerd voor kwaliteitsverschillen tussen woningen.

2.3 Data

2.3.1 Primaire waarneming

Om de PNK te kunnen berekenen, is een proces voor primaire waarneming opgezet. Er is een aselechte steekproef getrokken uit een geschatte populatie projectontwikkelaars en ontwikkelende bouwondernemingen. De steekproef is getrokken per stratum (regio X woningtype) en wordt regelmatig (gedeelte)lijk ververs. Aan de ondernemingen in de steekproef wordt elk kwartaal gevraagd om data aan te leveren van de verkochte nieuwbouwwoningen. De PNK is beschikbaar vanaf het eerste kwartaal van 2015.

2.3.2 Variabelen

Om de PNK te kunnen berekenen wordt gebruik gemaakt van adres- en garantiecertificaatgegevens (om de woningen te identificeren), de verkoopprijs en verschillende woningkenmerken.

Verkoopprijs (afhankelijke variabele)

Onder de verkoopprijs wordt het volgende verstaan: de overeengekomen (totaal)prijs inclusief grondkosten, kosten om de woning op naam te krijgen en eventueel andere bijkomende kosten, inclusief BTW. Indien bekend op het moment van verkoop, worden ook de kosten voor meerwerk meegenomen in de verkoopprijs.

Verkoopperiode (variabele om te stratificeren)

Het verkoopmoment is de datum waarop de (koop)overeenkomst door de laatste partij wordt ondertekend. Er is gekozen om het verkoopmoment aan te houden en niet het oplevermoment, omdat de prijs wordt bepaald en akkoord bevonden op het moment van tekenen. De oplevertermijn van een woning is daarnaast variabel per project en veelal ook per woning binnen een project. De tijd tussen het verkoopmoment en het oplevermoment kan daarbij behoorlijk oplopen in tegenstelling tot bij bestaande koopwoningen, waar deze termijn redelijk stabiel is.

De volgende woningkenmerken worden gebruikt om de PNK te berekenen:

Kenmerk 1: woonoppervlakte (onafhankelijke variabele)

Onder woonoppervlakte wordt verstaan: de gebruiksoppervlakte van de woning volgens NEN 2580, bestaande uit de gebruiksoppervlakte van de woonfunctie aangevuld met de gebruiksoppervlakte van overige inpanidige ruimte. De woonoppervlakte wordt opgegeven in vierkante meters.

Kenmerk 2: woningtype (onafhankelijke variabele)

Eén van de onafhankelijke variabelen is woningtype. Er worden vijf woningtypen onderscheiden:

- Hoekwoning
- Tussen- of geschakelde woning
- Twee-onder-een-kap woning
- Vrijstaande woning
- Appartement

Kenmerk 3: locatie (onafhankelijke variabele)

Ondernemingen geven op het enquêteformulier aan wat de woonplaats is van de woning en de woonplaats is gebruikt om de provincie uit af te leiden. De provincienaam is vervolgens gebruikt in de berekening van de prijsindex.

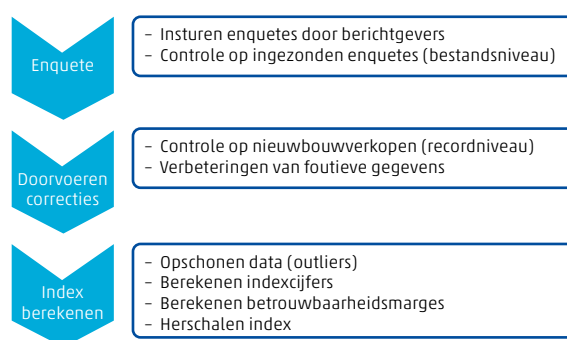
Kenmerk 4: garage (onafhankelijke variabele)

Ondernemingen geven per woning aan of er een garage aanwezig is bij de woning en of deze garage inpandig of uitpandig is. Afhankelijk hiervan is de woonoppervlakte inclusief of exclusief de oppervlakte van de garage.

De overige uitgevraagde variabelen worden gebruikt ter validatie van de uitkomsten.

2.4 Uitwerking PNK

De berekening van de PNK op basis van de hedonische regressiemethode, wordt in drie stappen uitgevoerd. Het onderstaande schema geeft een duidelijk overzicht van de verschillende processen die worden doorlopen om de prijsindex voor nieuwbouwoopwoningen te berekenen. In de eerste fase wordt de data binnengehaald en gereed gemaakt om in te lezen. In de tweede fase worden verschillende handelingen verricht om de kwaliteit van de data te verhogen. De derde fase heeft betrekking op het berekenen van de prijsindex. De laatste stap in deze fase, het herschalen van de index, is belangrijk omdat de referentieperiode nog aangepast moet worden. Het gemiddelde van de vier kwartalen van 2015 wordt bepaald en dit niveau wordt in de indexreeks gelijkgeschakeld aan 100. Hierdoor wordt 2015 de referentieperiode (2015=100).



Naast de prijsindex wordt ook de gemiddelde verkoopprijs van nieuwbouwoopwoningen berekend. Dit wordt gedaan na het opschonen van de data aan de hand van de onderstaande formule.

$$\bar{p} = \frac{\sum p}{n}$$

Hierbij staat $\sum p$ voor de som van alle waargenomen prijzen, n voor het aantal waargenomen prijzen en \bar{p} voor het berekende gemiddelde.

3. Berekeningsmethode aantal nieuwbouwtransacties

3.1 Samenwerking CBS-Kadaster

De prijsindex nieuwboukoopwoningen (PNK) maakt gebruik van data van verkopende ondernemingen. Omdat deze waarneming gebaseerd is op een steekproef, worden binnen dit onderzoek niet alle transacties van nieuwboukoopwoningen geregistreerd. Het Kadaster registreert wel alle woningtransacties. Uit deze registratie is echter niet direct af te leiden wanneer het een nieuwbouwwoning betreft. Voor het in kaart brengen van het aantal verkochte nieuwbouwwoningen is daarom in samenwerking met het Kadaster een aparte methode ontwikkeld om deze transacties te lokaliseren. Deze methode wordt hieronder toegelicht.

3.2 Data

De basis voor de methode is een combinatie tussen de Basisregistratie Kadaster (BRK) en de Basisregistratie Adressen en Gebouwen (BAG). Het transactiebestand van het Kadaster, opgebouwd vanuit de BRK, bevat alle woningtransacties in Nederland vanaf 1993. De database is gevuld aan de hand van informatie die is opgenomen in de overdrachtsakten.²⁾ De BAG bevat informatie over alle gebouwen in Nederland. Waaronder het gebruiksdoel van het object, het bouwjaar en de status van het pand (bijvoorbeeld bouwvergunning verleend, bouw gestart, in gebruik, etc.).

3.3 Methode afleiding nieuwbouwtransacties

De indicatie 'nieuwbouw' is bij het Kadaster geen authentiek gegeven, daarom is er een methodiek bedacht om de nieuwbouwtransacties te identificeren. Het is een stapsgewijze methode waarbij op basis van transactiedata van het Kadaster gecombineerd met data uit de BAG de nieuwboukoopwoningen worden geïdentificeerd.

Bij het filteren worden de onderstaande vragen bij elke transactie gesteld:

- Is er een koppeling tussen de transactie en een woning in de BAG?
- Dateert het perceel van vóór of na de start van de registratie door het Kadaster in 1993?
- Is het de eerste transactie van het perceel aan een natuurlijk persoon?
- Is het de eerste transactie van het object (woning) aan een natuurlijk persoon?
- Hoe groot is het verschil tussen het bouwjaar en het moment van verkoop?
- Heeft de transactie bij het Kadaster de identificatie nieuwbouwgrond bestemd voor een woning gekregen?
- Was de woning voorheen een huurwoning?

²⁾ De koopsom die het Kadaster registreert heeft betrekking op de grond, exclusief de bouwkosten van de nieuw te bouwen woningen. Om deze reden wordt de prijsindex niet gebaseerd op de Basisregistratie van het Kadaster.

Een combinatie van de antwoorden op deze vragen leidt tot de indicatie of het gaat om een transactie van een nieuwbouwoopwoning of niet.

3.4 Berekening aantallen en totale verkoopwaarde

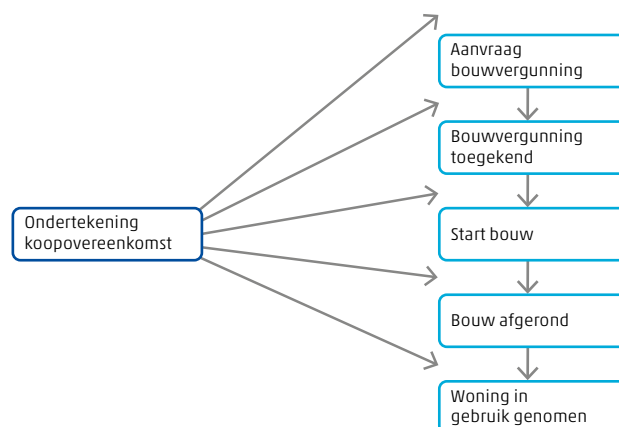
Nadat de nieuwbouwoopwoningen zijn afgeleid uit de Kadasterregistratie, wordt de som van het aantal transacties per periode berekend. Dit vormt het aantal dat weergegeven wordt in de tabel op Statline. De aantallen worden vervolgens vermenigvuldigd met de gemiddelde verkoopprijzen die berekend worden vanuit de primaire waarneming (hoofdstuk 2) aan de hand van de onderstaande formule. Dit vormt de totale verkoopwaarde die wordt weergegeven in de tabel op Statline.

$$P = N \cdot \bar{p}$$

Hierbij staat \bar{p} voor de gemiddelde verkoopprijs vanuit de primaire waarneming, N voor het berekende aantal uit de Kadasterregistratie en P voor de berekende totale verkoopwaarde.

3.5 Verkochte nieuwbouwoopwoningen t.o.v. opgeleverde nieuwbouwwoningen

Het CBS publiceert naast aantallen van verkochte nieuwbouwoopwoningen ook aantallen m.b.t. opgeleverde nieuwbouwwoningen³⁾. Het voornaamste verschil tussen beide statistieken is het meetmoment: bij de opgeleverde nieuwbouwwoningen wordt een nieuwbouwwoning meegeteld op het moment dat de bouw van een huis is afgerond en wordt toegevoegd aan de woningvoorraad. Bij de verkochte nieuwbouwoopwoningen wordt een woning meegeteld zodra er een koopovereenkomst is getekend. Zoals in onderstaande figuur is weergegeven, kan dit plaatsvinden op meerdere momenten van het bouwproces.



³⁾ Zie: <https://opendata.cbs.nl/statline/#/CBS/nl/dataset/81955NED/table?dl=E4E5>

Daarnaast worden bij de aantallen voor de PNK alleen koopwoningen meegenomen. De statistiek 'opgeleverde nieuwbouwwoningen' bevat naast koopwoningen ook huurwoningen.

Referenties

[1] Eurostat (2013). Handbook on Residential Property Prices Indices (RPPIs).

[2] Eurostat (2012). Technical manual on Owner-Occupied Housing for Harmonised Index of Consumer Prices (version 2.0).

Verklaring van tekens

Niets (blanco)	Een cijfer kan op logische gronden niet voorkomen
.	Het cijfer is onbekend, onvoldoende betrouwbaar of geheim
*	Voorlopige cijfers
**	Nader voorlopige cijfers
2017-2018	2017 tot en met 2018
2017/2018	Het gemiddelde over de jaren 2017 tot en met 2018
2017/'18	Oogstjaar, boekjaar, schooljaar enz., beginnend in 2017 en eindigend in 2018
2015/'16-2017/'18	Oogstjaar, boekjaar, enz., 2015/'16 tot en met 2017/'18

In geval van afronding kan het voorkomen dat het weergegeven totaal niet overeenstemt met de som van de getallen.

Colofon

Uitgever

Centraal Bureau voor de Statistiek
Henri Faasdreef 312, 2492 JP Den Haag
www.cbs.nl

Prepress

Centraal Bureau voor de Statistiek

Ontwerp

Edenspiekermann

Inlichtingen

Tel. 088 570 70 70
Via contactformulier: www.cbs.nl/infoservice

© Centraal Bureau voor de Statistiek, Den Haag/Heerlen/Bonaire, 2018.
Verveelvoudigen is toegestaan, mits het CBS als bron wordt vermeld.