



# The demographic projections of the Netherlands Environmental Assessment Agency (PBL) and Statistics Netherlands (CBS)

Dr. Trond Grytli Husby

Netherlands Environmental Assessment Agency (PBL)

[trond.husby@pbl.nl](mailto:trond.husby@pbl.nl)

Amsterdam 24 September, 2018



## Plan for the day

1. Population projections: definition, relevance
2. Components and projection modelling in general
3. Data
4. Uncertainty and scenarios
5. National population projections by CBS
6. Regional population projections by CBS/PBL, focus on internal migration



## Who am I

- Norwegian, living in the Netherlands for 9 years
- PBL for 2.5 years
- Working with the population projections for 0.5 year
- Economist by training (VU Amsterdam)



## Why am I working with this

- It's my job...
- Population change is important for economic questions
- Challenging topic
- Population growth is (sometimes) controversial



Source: [https://img.clipartxtras.com/03843abee543e741be870c4ada22c760\\_free-to-use-public-domain-cauldron-clip-art-witch-cauldron-clipart\\_500-500.png](https://img.clipartxtras.com/03843abee543e741be870c4ada22c760_free-to-use-public-domain-cauldron-clip-art-witch-cauldron-clipart_500-500.png)



## Definition of projections

- Van Dale: *statement about the probable course or the probable outcome or outcome of elections, competitions and the like.*
- Projections ≠ forecast ≠ scenario studies
- Projections are, almost by definition, wrong: population projections with no error are most likely result of coincidence
- Think of projections as an attempt of establishing the *most probable* trajectory of population change
- Anyone working with population projections should have this in mind



## Why do we make demographic projections?

- Population change is important in other settings: house prices and rent in Amsterdam are high because many people want to live in Amsterdam
- Policy makers use them: plans for new residential areas, schools, location of hospitals...
- Important: normative (is population growth good or bad?) versus descriptive (what are the main drivers of population growth?) analysis



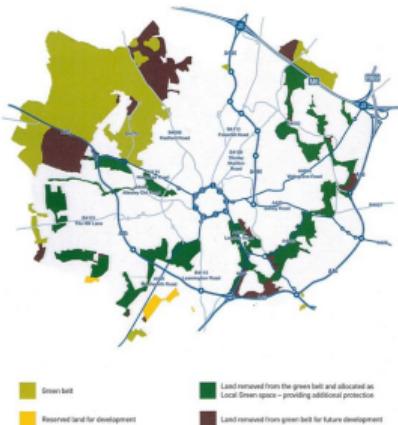
## Why do we make demographic projections?



- Population in the Netherlands is projected to grow
- Currently, growth is driven primarily by net migration [▶ Link](#)
- Big, normative, question about what kind of Netherlands is desirable 40 years from now



## Why do we make demographic projections?



Based on population projections from the UK ONS, the city council of Coventry is planning to develop its Green Belt. Not everyone is happy

*Population projections might have some value in projecting school numbers five years ahead (although recent experience in the Council might cast doubt on that) but projecting twenty years up to 2031 is completely unrealistic.*

▶ [Link](#)

Source: <https://www.coventrytelegraph.net/news/coventry-news/revealed-swathes-green-belt-lost-10674329>



## Demography and projections

- Demography: description of the past
- Projections: best guess of the future
- Population and households:
  - ▶ *Amount*: number of households and people
  - ▶ *Composition*: distribution by gender, age, ethnicity/background, household position
  - ▶ *Spatial distribution*: world, country, region (province, COROP, municipality, neighbourhood)
  - ▶ *Changes*: birth, death, immigration, emigration, departure or arrival, transitions between household positions



## Modelling: static versus dynamic

*Static (trend extrapolation):*  $Pop_{t+1} = Pop_t + \Delta(Pop_t - Pop_{t-1})$

- +: straight-forward implementation, low data requirements
- : no insights into underlying drivers of growth/decline or structural change

*Dynamic (cohort-component):* estimate trends in growth components (birth, death, external and internal migration)

- : high data requirements, sometimes data not available
- +: insight into underlying drivers



## Cohort-component model (national)

$$P_{t+1} = P_t + B - D + I - E$$

$P_t$ : population in  $t$

$B$ : births in the interval  $(t, t + 1)$

$D$ : deaths in the interval  $(t, t + 1)$

$I$ : immigration in the interval  $(t, t + 1)$

$E$ : emigration in the interval  $(t, t + 1)$

X typically calculated as  $\frac{X_t + X_{t+1}}{2}$



## Growth components

### Cross-sectional (calendar year)

- Changes related to the business cycle
- Fluctuations in birth rates
- Stricter immigration policies
- Enlargement of the EU (work-related migration)
- War (asylum seekers)

### Longitudinal (cohort)

- Structural changes
- Average number of children from 3 to 2
- People live longer



## Growth components at different scales

World: births and deaths

National: births, deaths, net migration

Regional: births, deaths, net migration,  
arrivals and departures



Source: <http://4.bp.blogspot.com/-8pNWCR1SJhY/T83V8A2HzXI/AAAAAAAADuM/aRub2XdIBEo/s1600/Planet+Earth+8.jpg>

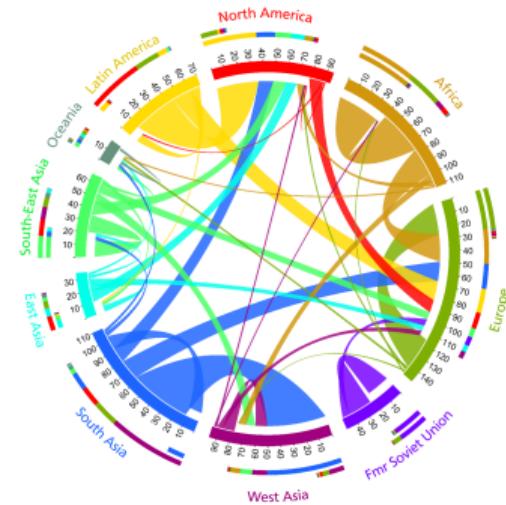


## Growth components at different scales

World: births and deaths

National: births, deaths, net migration

Regional: births, deaths, net migration,  
arrivals and departures



Source: [http://www.slate.com/content/dam/slate/blogs/the\\_world/\\_2014/04/02/world\\_on\\_the\\_move\\_five\\_years\\_of\\_global\\_migration\\_in\\_one\\_chart/circular\\_plot\\_flows\\_between\\_world\\_regions\\_200510\\_1.png.CROP.cq5dam\\_web\\_1280\\_1280.png](http://www.slate.com/content/dam/slate/blogs/the_world/_2014/04/02/world_on_the_move_five_years_of_global_migration_in_one_chart/circular_plot_flows_between_world_regions_200510_1.png.CROP.cq5dam_web_1280_1280.png)



## Growth components at different scales

World: births and deaths

National: births, deaths, net migration

Regional: births, deaths, net migration,  
arrivals and departures



Source: <https://i.ytimg.com/vi/68nvHI10BdA/maxresdefault.jpg>



## Regional population projections

Top-down: disaggregate national numbers to the region

+: straight forward to implement

+: per definition consistent with national numbers

-: little/no insight into determinants of regional population change

Bottom-up: calculate regional numbers consistent with national numbers

+: insights into determinants of regional population change

-: heavy data and time requirements

-: extra step required to ensure consistency with national numbers



## Question

1. In a parallel dimension, a (much smaller) earth is populated by 100 people in year  $t$ . In  $t+15$ , the maximum population aged 15 and above is 100. What else do we need to know in order to project the exact population in  $t+15$ ?
2. The population is evenly distributed between two countries: A and B. What do we need to know in order to project the population in A and B in  $t+15$ ?



## Data availability: the ideal world

To estimate the growth components we need consistent time-series of data with detailed personal information:

- *Migration*: all moves for every individual person during his/her entire life
- *Household*: formation and transitions between positions for every person in every household
- *Construction*: building plans are correct, will be carried out and people will move to the houses that are built
- *Time series*: smooth developments



## Data availability: the real world

- People are registered at a place they do not live...or not at all
- There is no registry (UK!)
- Building plans are not carried out (economic crisis) or no one moves to the houses that are built
- Strange movements in times series: is it noise or structural change?



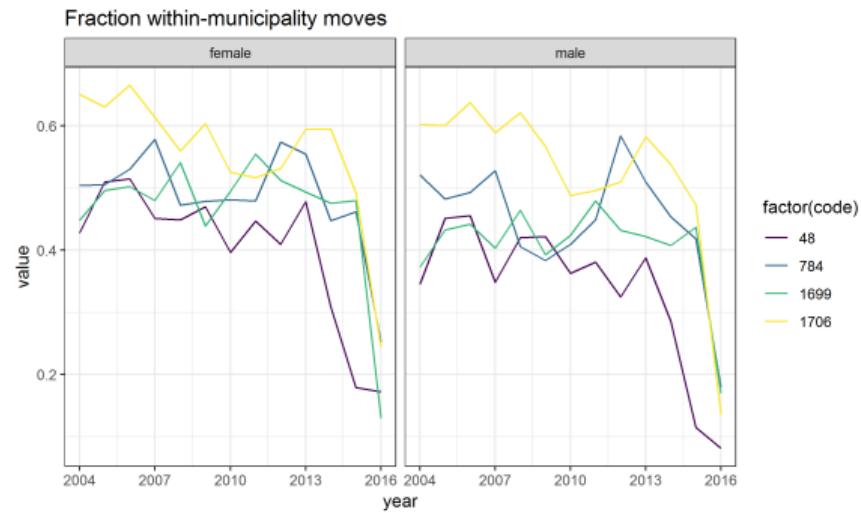
## Missing or no registry data

- UK methodology for subnational population projections, combination of survey and census data [▶ Link](#)
- Example NL: we want to include population by education in the regional population projection
- We only have information about education levels from a survey with a representative sample on a national level
- We can *estimate* the number of people by education levels by scaling the survey data so that it corresponds with registry data (small-area estimation)



## Strange observations in time series data

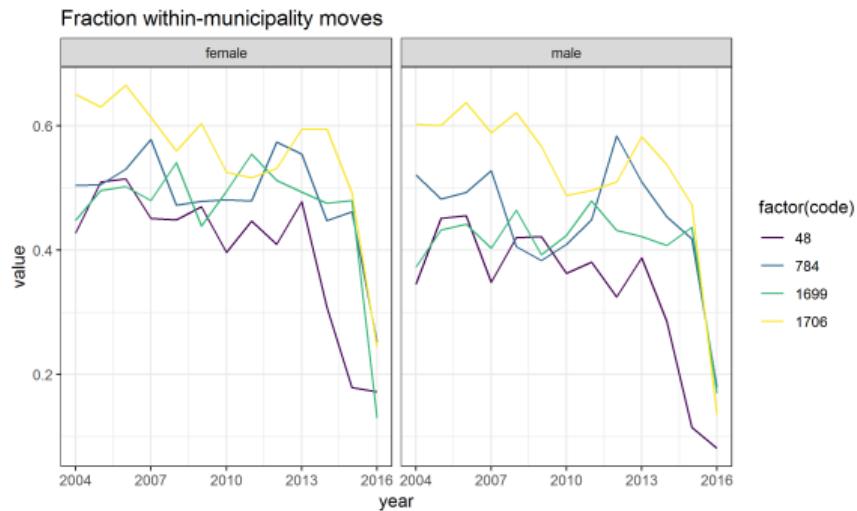
The fraction of within-municipality moves for some municipalities hosting an asylum-seeker centre. What is the fraction in 2028?





## Strange observations in time series data

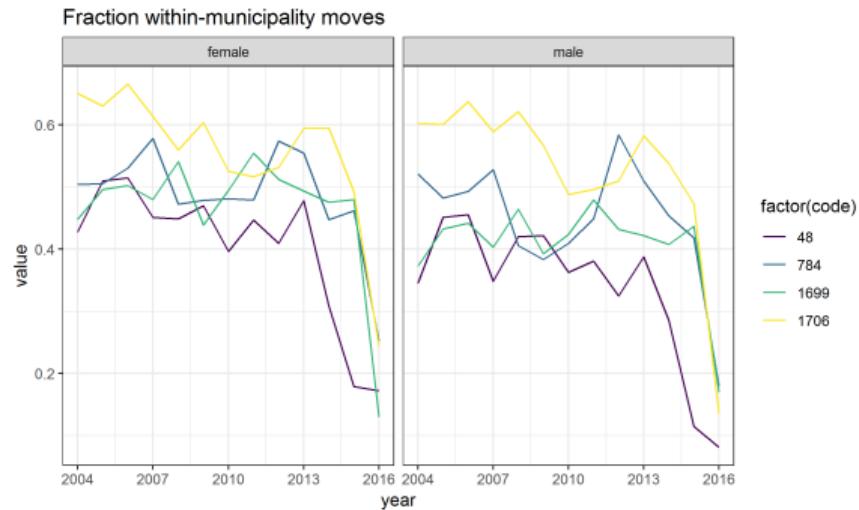
The fraction of within-municipality moves for some municipalities hosting an asylum-seeker centre. What is the fraction in 2028? Relatively small municipalities where a large proportion moved from another municipality and moved quickly to another municipality.





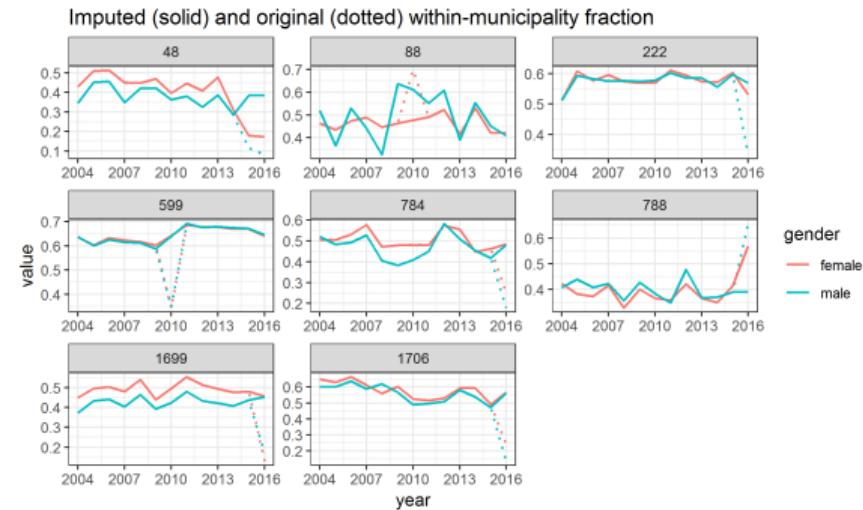
## Strange observations in time series data

The fraction of within-municipality moves for some municipalities hosting an asylum-seeker centre. What is the fraction in 2028? The decline in 2015/2016 is not *structural*, observations can be treated as outliers.



## Strange observations in time series data

One solution: impute the outliers with a regression model. Many other methods available





## The origins and definition of uncertainty

- Typology: Unknown knowns; known unknowns; unknown unknowns.
- Population projections only account for the first type
- The cohort-component model is a book-keeping system; per definition no uncertainty
- However, substantial uncertainty in the trend of the growth components
- Incorporating uncertainty: create projection interval with possible outcomes (according to the model)



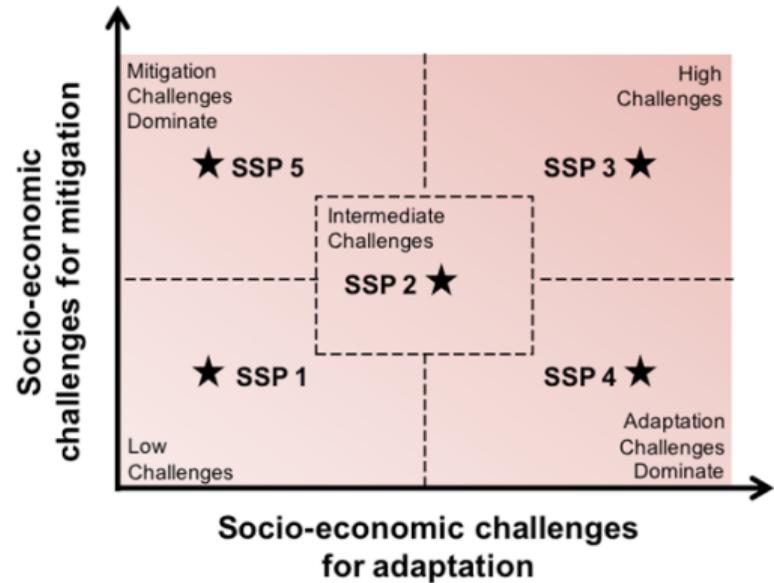
## Uncertainty in the projections

- Circularity: what you are trying to project might be affected by your projection
- Self-fulfilling prophecy: building plans reflect projections, projection becomes true
- ...or self-denying prophecy: policy makers do not like projection and restrict migration, projection is wrong
- Examples: Coventry and population growth in the Netherlands



## Scenarios versus projection

- Scenario's are quantification of future *story lines*, not necessarily the *most likely* outcomes
- A way to structure discussion of possible futures, for example about the impacts of climate change

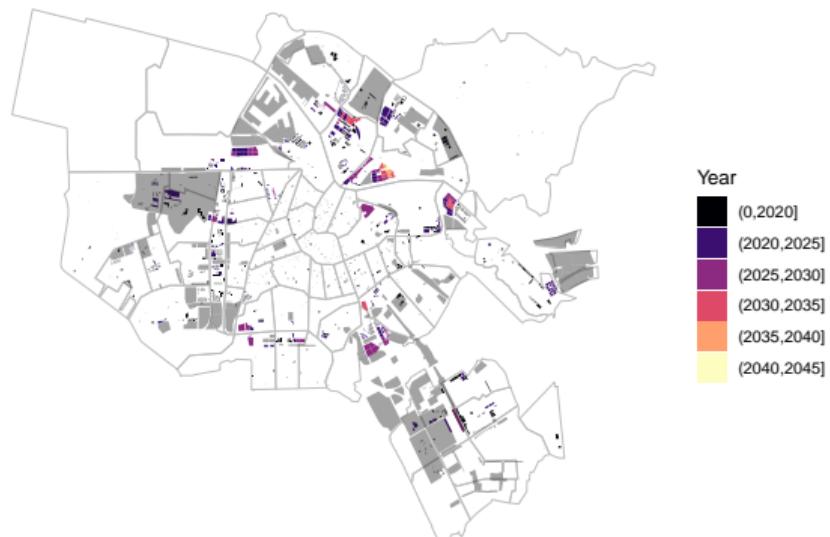


Source: <https://climate4impact.eu/files/SSPs.png>



## Scenarios versus projection

- Scenario models can be used to illustrate spatial or temporal impact of policy (comparison with a baseline scenario)
- Example: investigating the impact on employment and transport demand in Amsterdam with the model TIGRISXL. What happens if the municipality builds more houses instead of apartments?

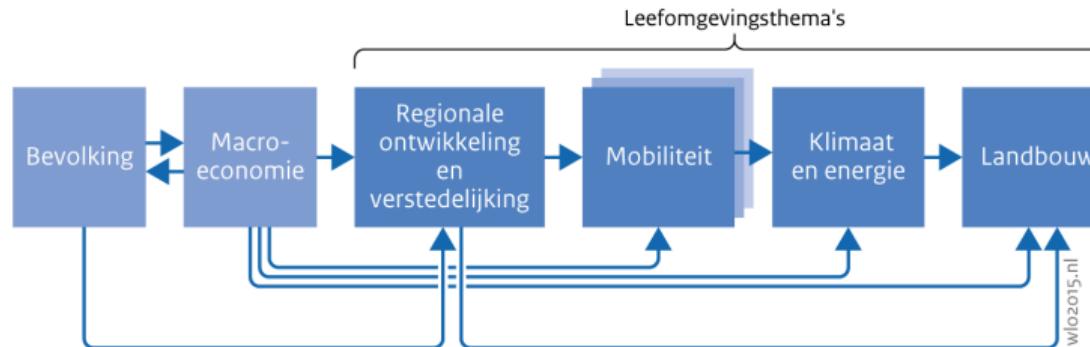




## Example: the WLO scenarios

Toekomstverkenning Welvaart en Leefomgeving = future scenarios for welfare and the environment. Made by CPB and PBL in 2015 [▶ Link](#)

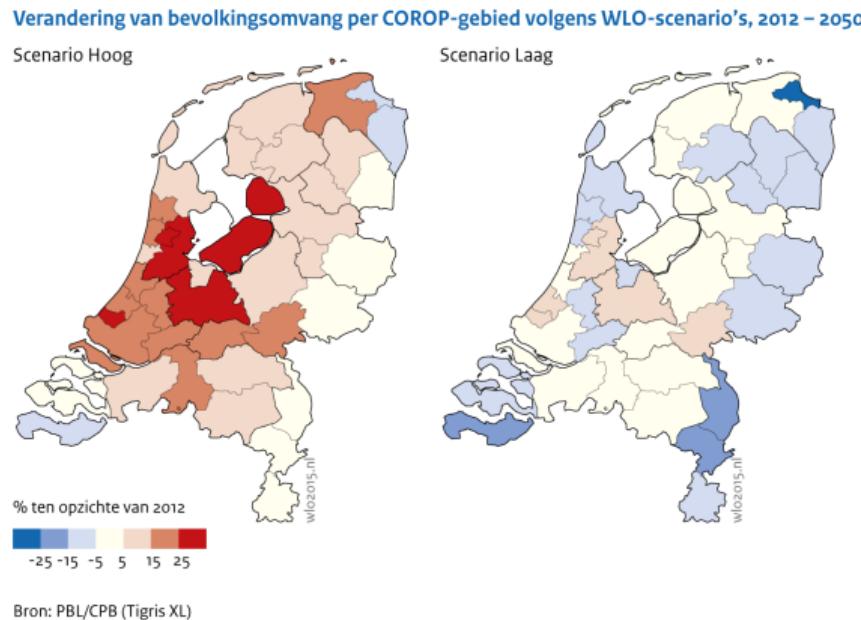
### Samenhang tussen de modules



Bron: PBL/CPB

## Example: the WLO scenarios

- Scenario high: increase in external migration, high increase in life expectancy, high fertility
- Assumptions: high economic growth in Europe, low unemployment and increasing income and productivity in NL
- Scenario low: low net migration, little increase in life expectancy, low fertility
- NL emerges slowly from the recession, economic growth in Europe of 1 % per year





## CBS/PBL stochastic population projections

- Uncertainty incorporated through (estimated) probability of deviance from trend
- Distribution of population and households obtained through assumptions about variance in components (births, death, external migration)
- Generate trajectories of each component with time-series model incorporating deviance from trend
- Monte Carlo simulation of 1000 trajectories for individual components, subsequently calculate 1000 trajectories for population → basis for 67 % (95 %) projection intervals

▶ Link



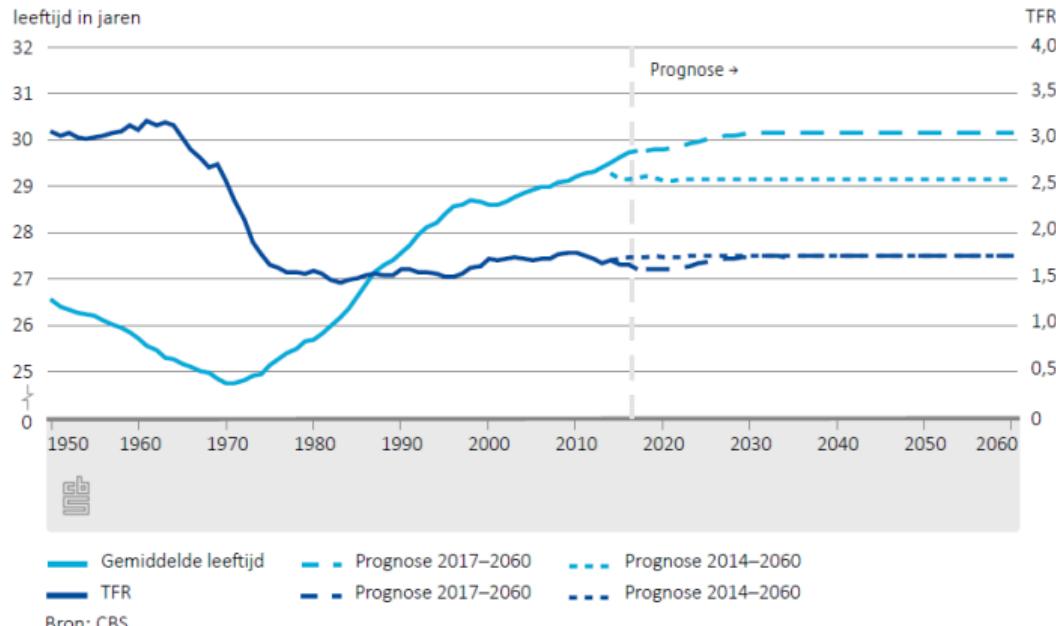
## National population projections

- No 'big theory' about the future (e.g., the demographic transitions)
- Projections until 2060
- Assumptions about components
  - ▶ Birth: development of total fertility rate (average number of children per woman)
  - ▶ Death: development of life expectancy at birth (men and woman)
  - ▶ External migration: development of immigration (number) and emigration (probability)



## National population projections: births

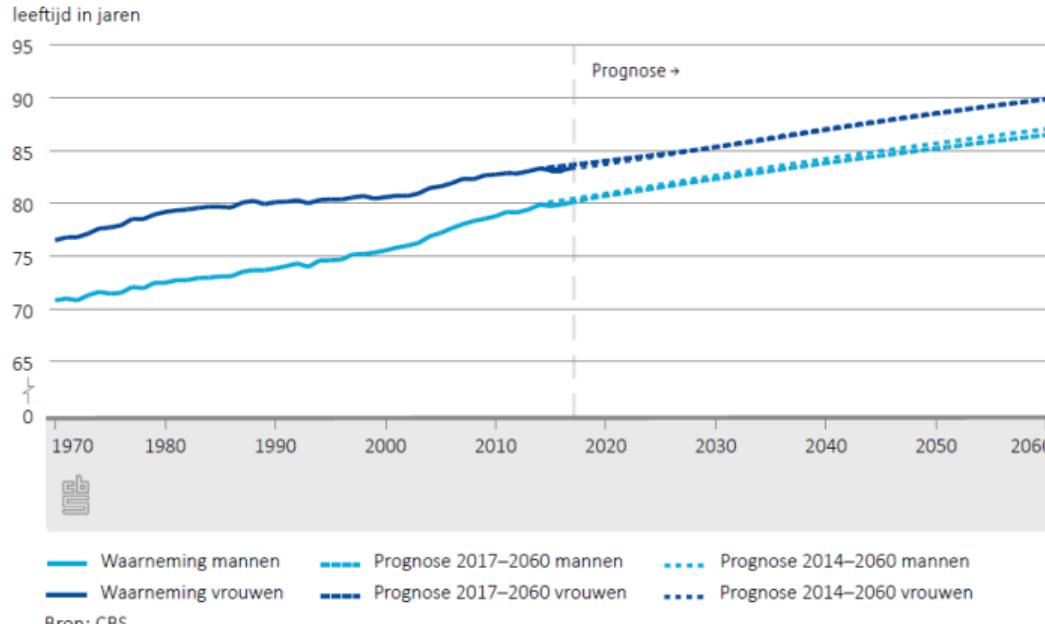
### 2.1.2 Gemiddelde leeftijd waarop vrouwen voor het eerst moeder worden en het totaal vruchtbaarheidscijfer; waarneming, prognose 2017–2060 en prognose 2014–2060





## National population projections: deaths

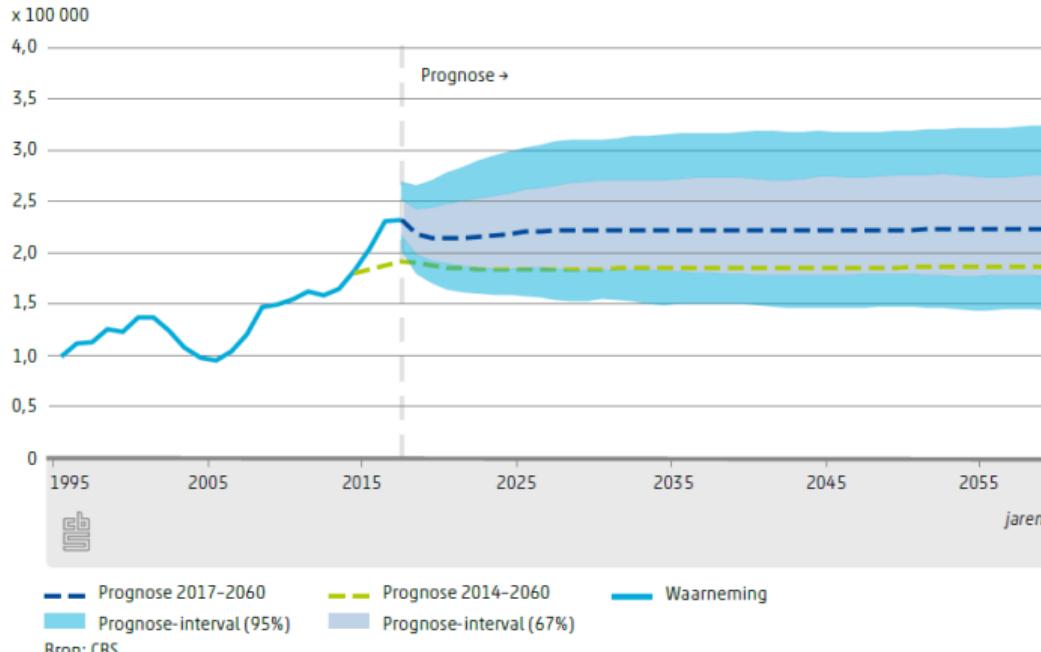
### 2.2.2 Periode-levensverwachting bij geboorte; waarneming, prognose 2017–2060 en prognose 2014–2060





## National population projections: immigration

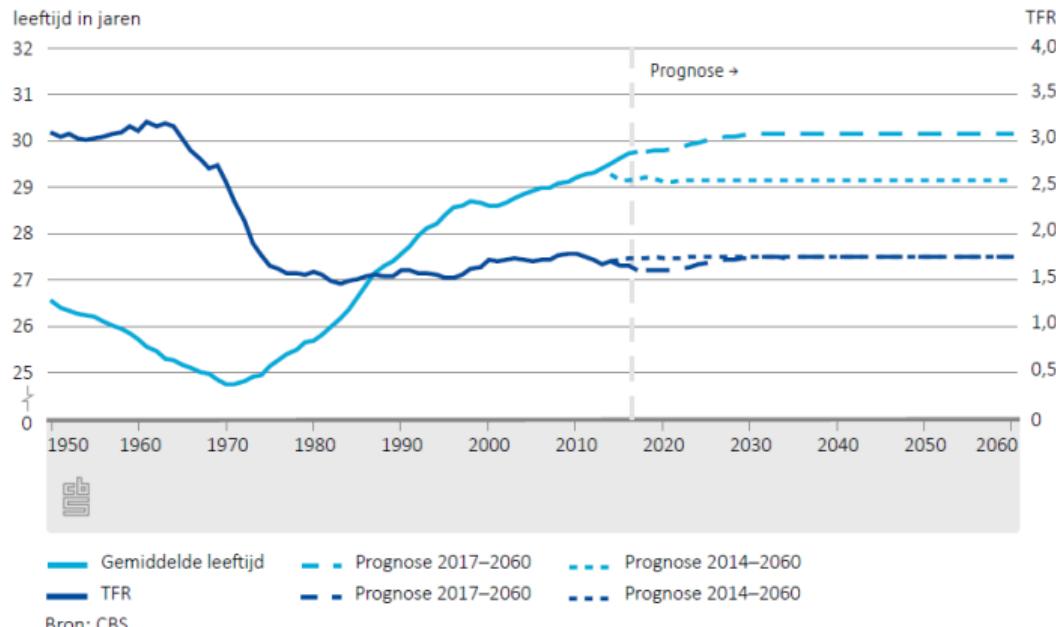
### 2.4.5 Immigratie; waarneming, prognose 2017-2060 (inclusief intervallen) en prognose 2014-2060





## National population projections: emigration

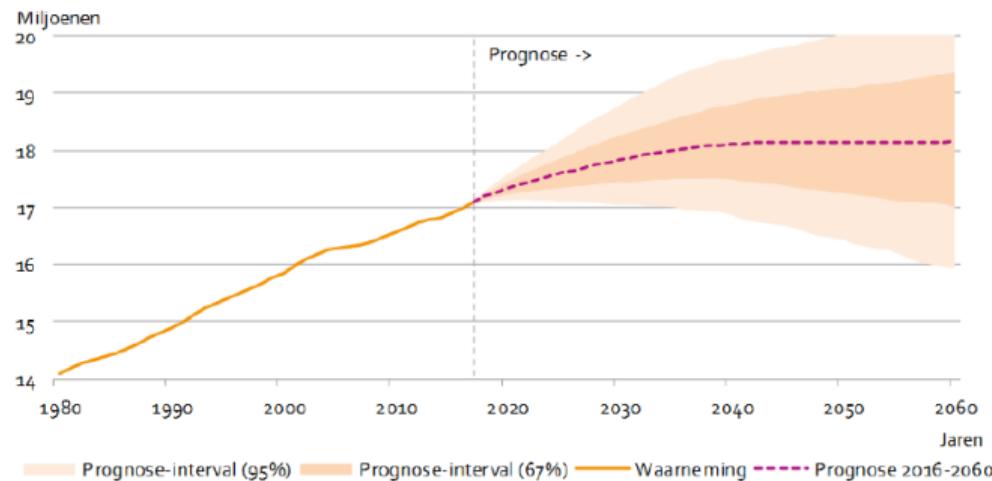
### 2.1.2 Gemiddelde leeftijd waarop vrouwen voor het eerst moeder worden en het totaal vruchtbaarheidscijfer; waarneming, prognose 2017–2060 en prognose 2014–2060





## National population projections: population

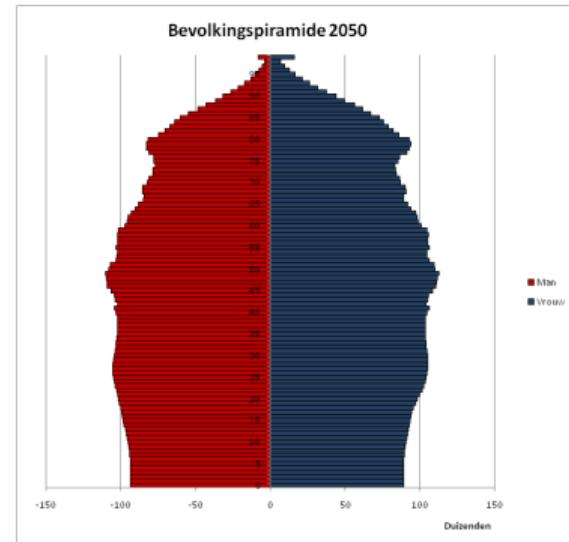
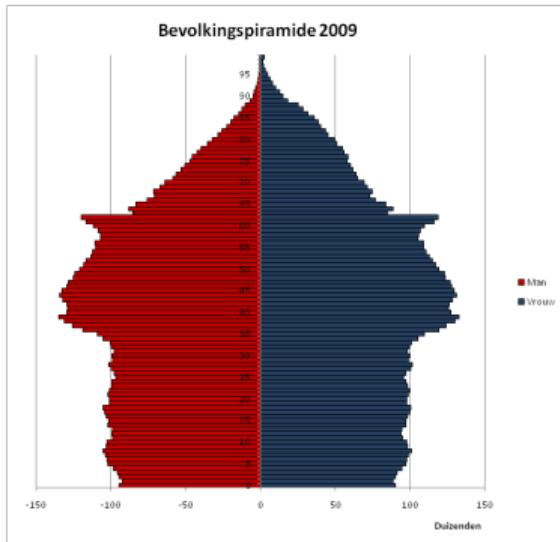
### 3.1 Inwonertal, prognose 2016-2060



Bron: Centraal Bureau voor de Statistiek

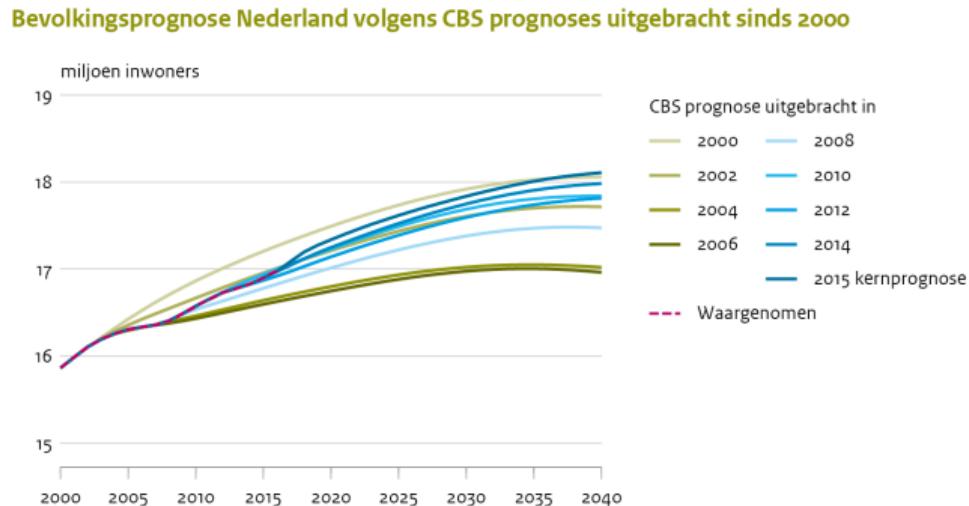


## National population projections: age pyramid





## National population projections: uncertainty





## Regional population projections

- Projections of population, households and demographic events in Dutch municipalities until 2040
- Carried out every three years: previous edition was in 2016, next in 2019
- Regional projections (PBL/CBS) are made consistent with the national projections (CBS)
- The projections made with the cohort-component model PEARL (Projecting population Events at Regional Level)
- Trends in components of population growth and transition rates between household positions are projected separately as inputs to the model



## Migration and regional population growth

- PEARL distinguishes long- and short distance migration (cut-off 35 km)
- Long distance migration probability matrix extrapolated from past observations
- Trends in short-distance migration predicted with a spatial interaction model (right hand side)

$$M_{i,j} = O_i \frac{P_j^{\alpha_i} D_{ij}^{\beta_i} C_{ij}^{\gamma_i} H_j^{\delta_i}}{\sum_k P_k^{\alpha_i} D_{ik}^{\beta_i} C_{ik}^{\gamma_i} H_k^{\delta_i}}$$

(i,j): municipalities

M: moves

O: departures

P: population

D: distance

I: centrality

H: net new dwellings

$\alpha, \beta, \gamma, \delta$ : estimated parameters

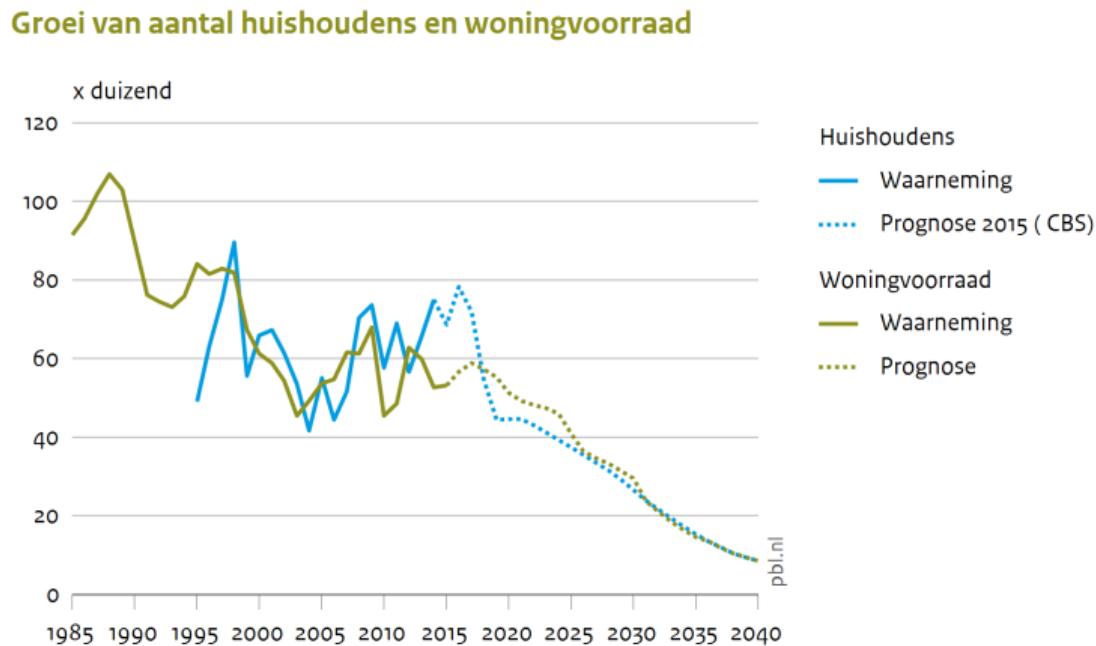


## Short-distance migration and the dwelling stock

- Strong relation between population growth and the number of dwellings: effect of new construction is particularly strong in municipalities with a shortage of housing
- Short-distance migration in PEARL:
  1. Potential moves estimated with spatial interaction model
  2. Actual migration calculated by taking into account housing availability (potentially moving back to 1...)
- Building plans are obtained through a combination of desk research and consultation with municipalities and provinces



## Building plans country-level

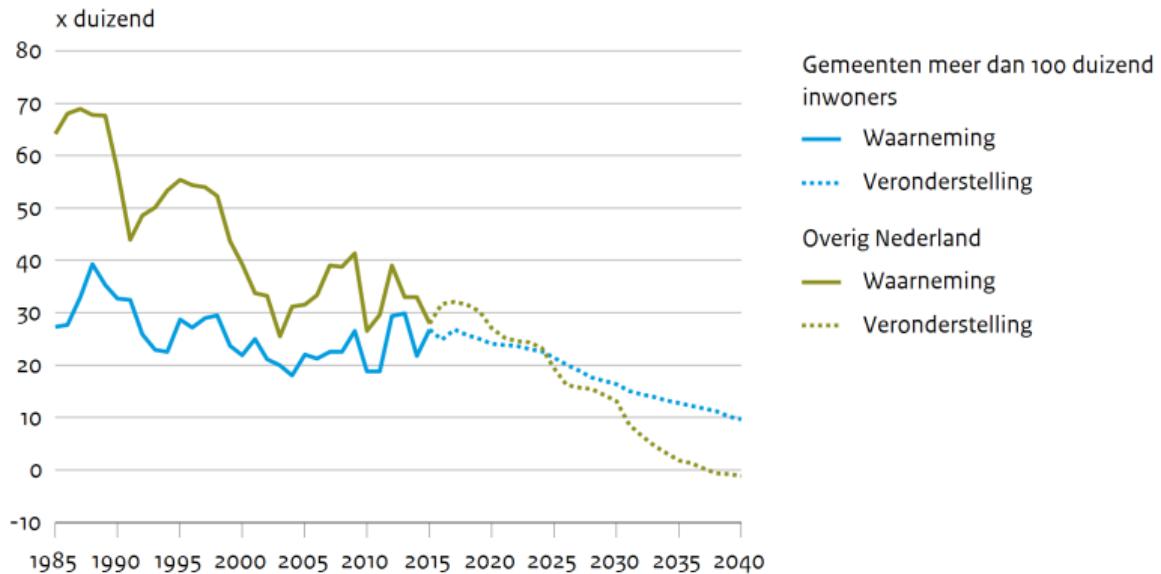


Bron: CBS



## Building plans large and small municipalities

### Netto woningbouw





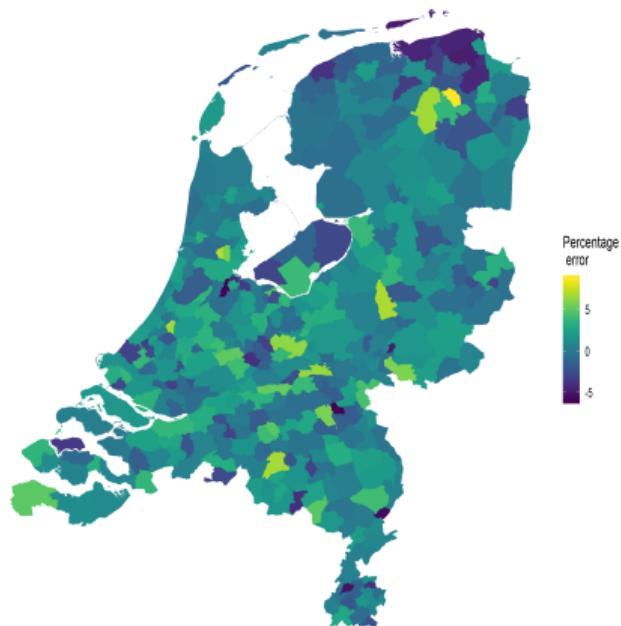
## Regional population 2016 edition

- Population in the Netherlands will grow until 2040 (consistent with national projections)
- Growth is unevenly spread: strong growth in urban agglomerations, population decline in peripheral regions



## How accurate were the projections? <sup>1</sup> <sup>2</sup>

Mean absolute percentage error: 1.85



---

<sup>1</sup>

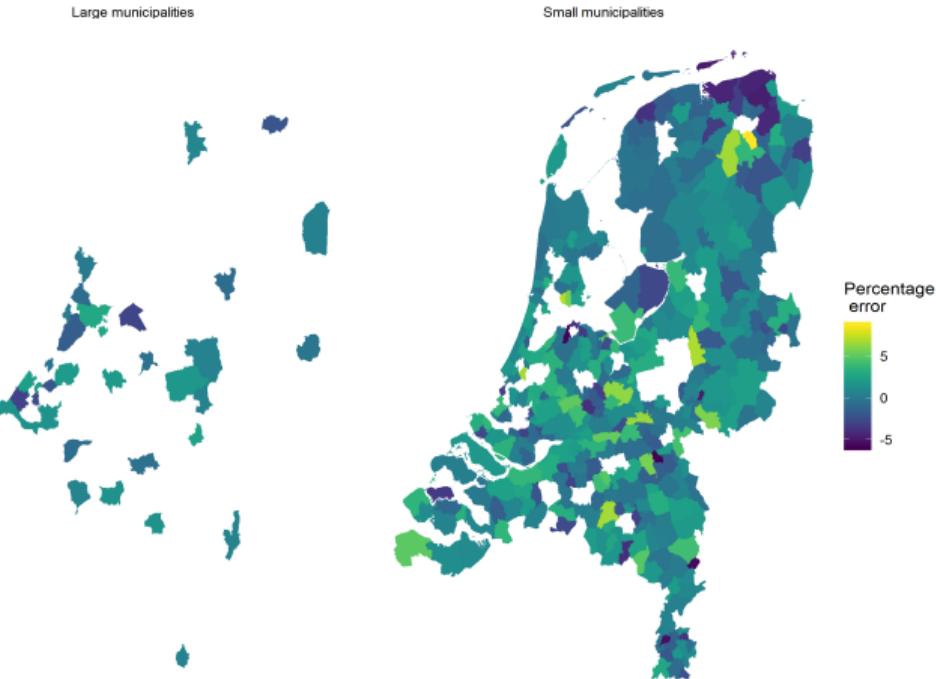
Population in 2017 according to the 2013 edition,  
compared with the data.

<sup>2</sup>

$$\text{error} = \frac{\text{actual} - \text{prediction}}{\text{actual}} * 100$$



## How accurate were the projections?<sup>3</sup>





## Explaining deviances

- 2012 was, in terms of the housing market, still very much a crisis year
- Plans for future building by municipalities and provinces were likely pessimistic
- The sour mood worked through the projections of short-distance migration into the population projections!
- The deviations were larger in small municipalities (still PEARL did better than a naive forecast)

▶ Link



## Question

A furious property developer from Almere calls you up. Right before the projections of 2013 the municipality had agreed to several large construction plans. However, now it appears the projections were off the mark: population grew much slower than the projections. The property developer blames you for the houses that were not sold. How do you respond?



## Conclusions

- *Prediction is difficult, especially about the future* Niels Bohr (?)
- *Essentially all models are wrong, but some are useful* George Box
- Population projections are central for policy making, both for large-scale normative and local practical questions
- Data: difference between ideal world and reality
- Uncertainty is inherent in any projection, and they are almost never exactly accurate
- National population projections by CBS
- Regional population projections by CBS/PBL: internal migration key driver of population change