



Accounting for uncertainty in London's future population

Ben Corr, Chris Fairless, Marta Lapsley

City Intelligence Unit

Greater London Authority



Contents

- Background
- Reviewing past projections
- Sources of uncertainty
- Coping strategies

Background

- Why is this important?
- Why is this difficult?

Why is this important?

- Long term policy and planning decision based on projections

Do we need to build Crossrail 4?

- Want to test implications of possible scenarios

Fall in EU migration, different levels of housing delivery

- Shorter term planning also relies on projections

How much public transport fare revenue will we have over the next five years?

Why is this difficult?

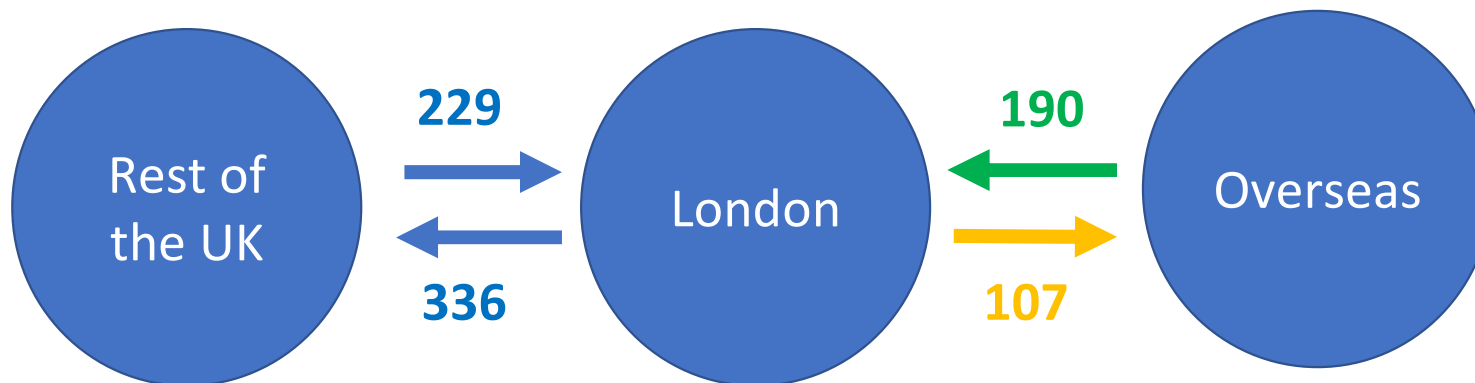
- Annual migration flows in/out of London ~**900k**
- Total net flow to London ~**0** *
- Flows highly asymmetric

*(+/- a few 10s thousands)

International in → Domestic out

Young adults in → Everyone else out

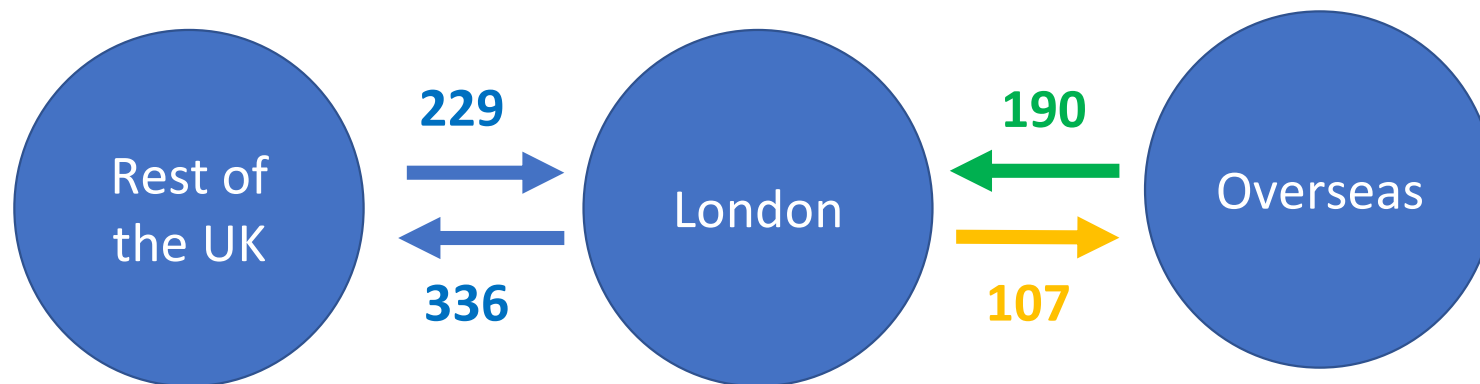
- Flows estimated with different methodologies



Flows mid-2016 to mid-2017: ONS MYE

Why is this difficult?

- Small proportional change in any flow has big impact on net
- Migration is inherently uncertain and hard to measure
- Data sources vary in completeness, accuracy, biases



Flows mid-2016 to mid-2017: ONS MYE

Why is this difficult?

- Trend based models offer limited capacity to test scenarios outside of past history
- Models don't typically include direct relationships and interactions between components
- Good at reflecting structural change
- Usually less good reflecting behavioural responses



Reviewing past projections

- How have previous projections stood the test of time?

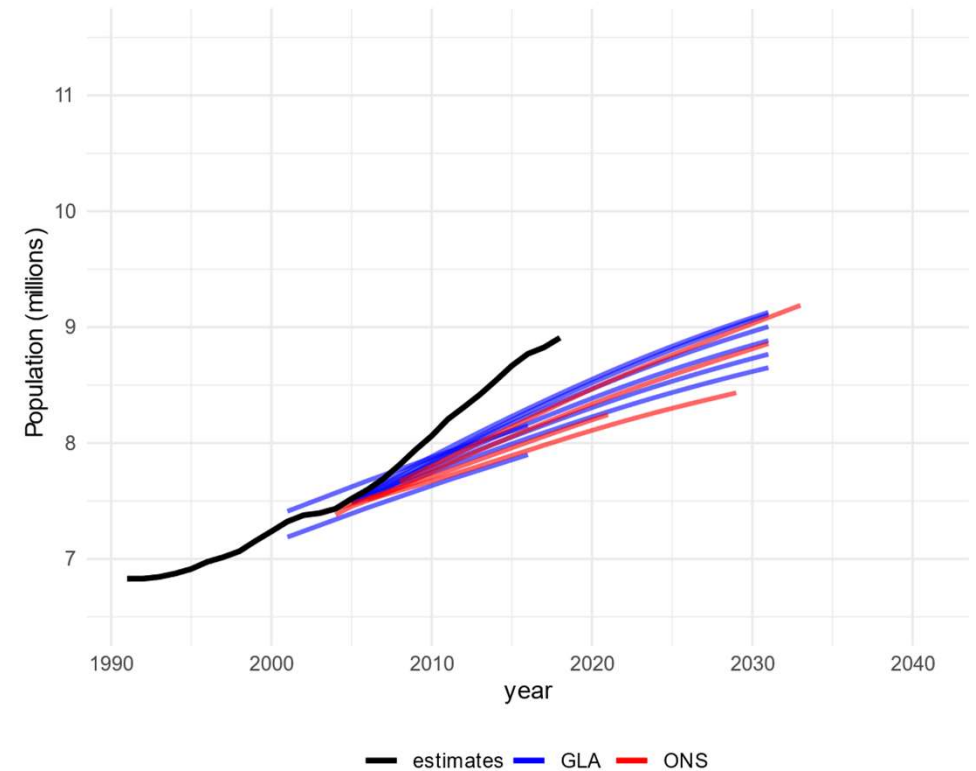


Reviewing past projections

- How have previous projections stood the test of time?
- On the face of it, not great...

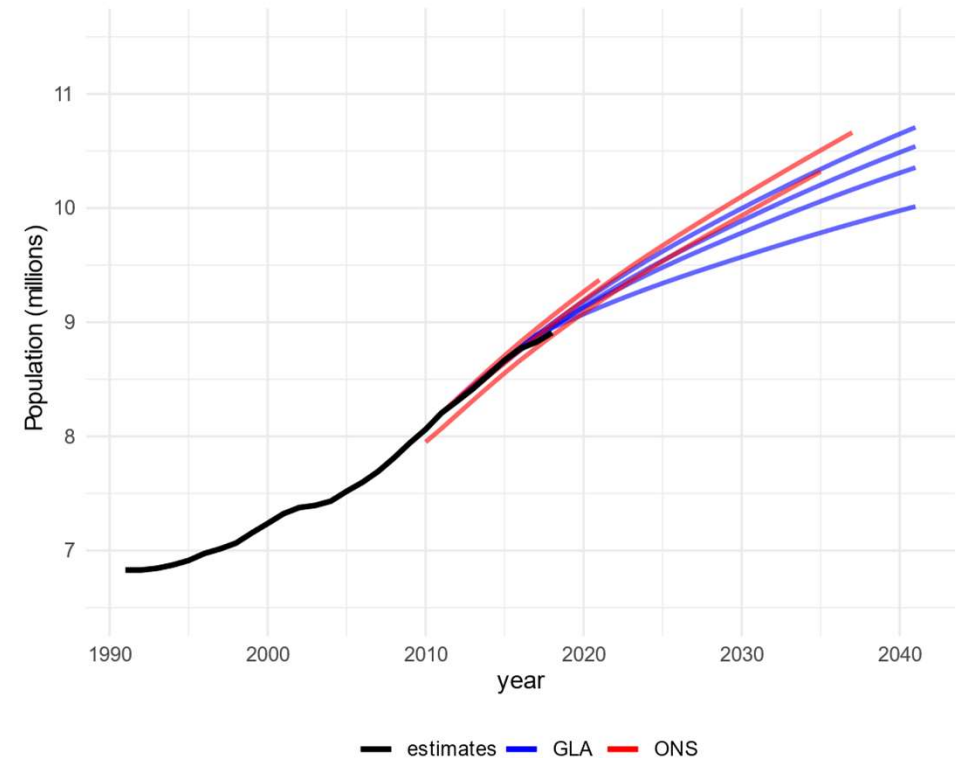
Reviewing past projections

- Projections from the 2000s dramatically undershot



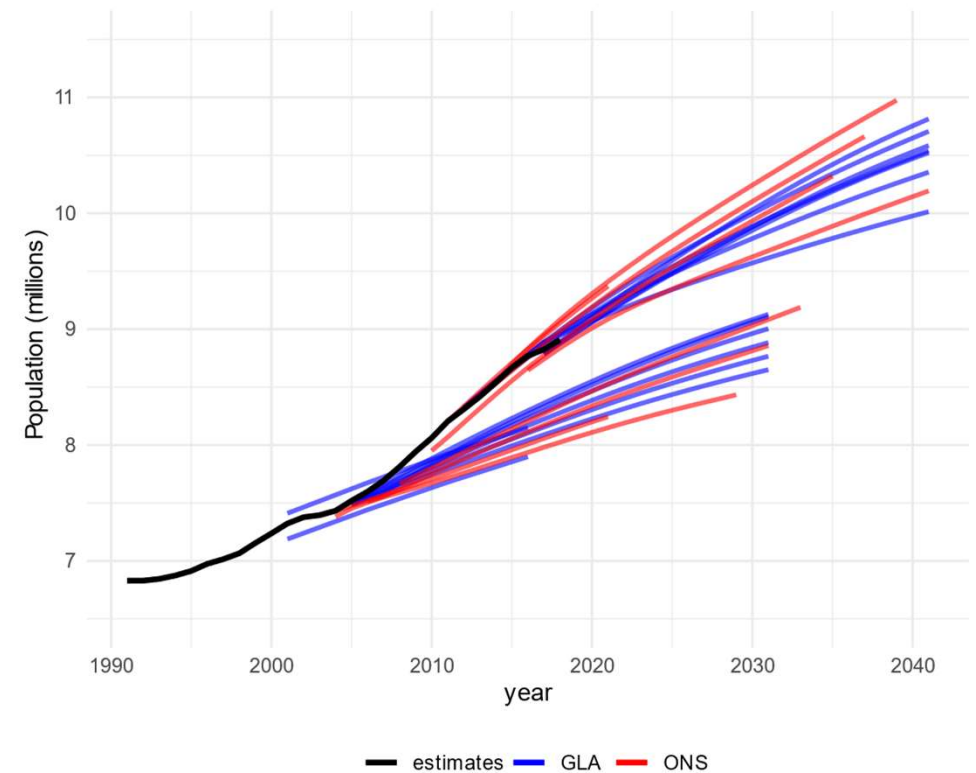
Reviewing past projections

- Projections from the 2000s dramatically undershot
- Projections from 2010s took a much higher trajectory



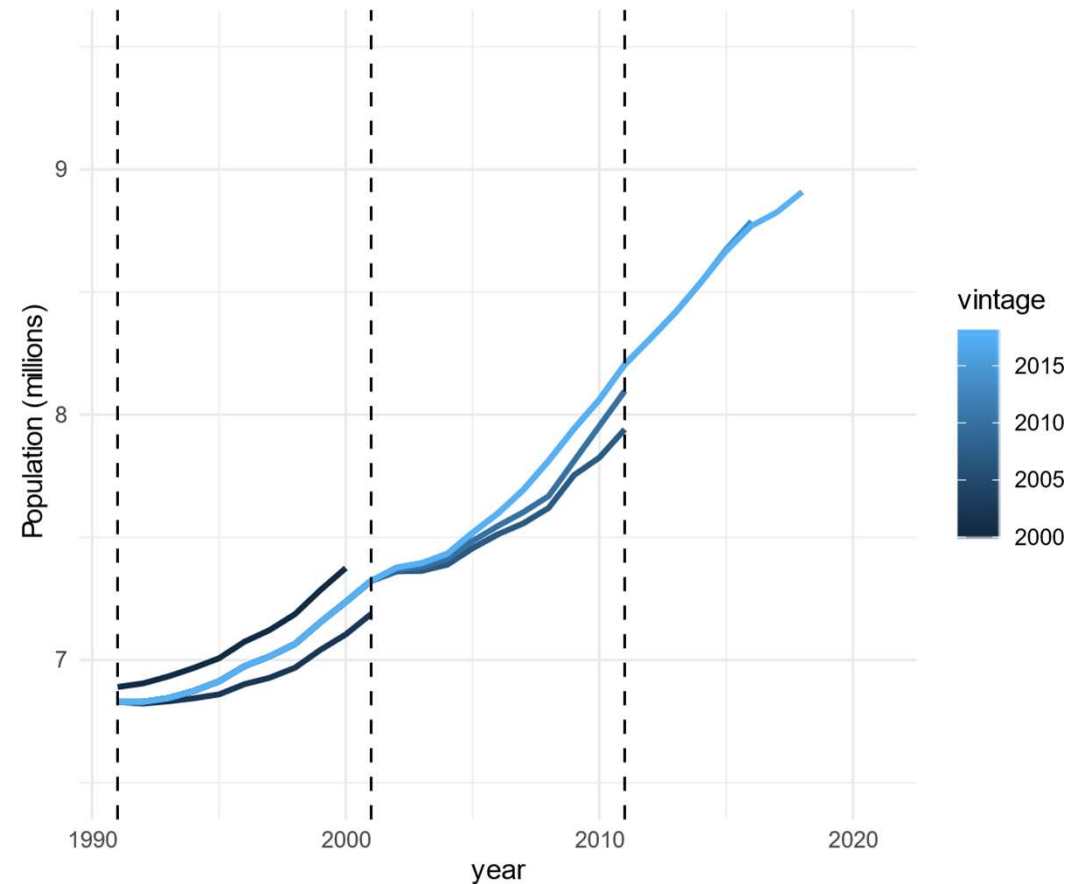
Reviewing past projections

- Projections from the 2000s dramatically undershot
- Projections from 2010s took a much higher trajectory
- Why the sudden change?



Reviewing past projections

- Have to view projections in the context they were produced
- Estimates revised multiple times over the years
- Projections generally look plausible *given the inputs*





Sources of uncertainty

- Variation in components
- Error in input data
- Assumptions
- Model incompleteness
- Complexity

Variation in components

- Year-on-year fluctuations in underlying behaviours
- Random noise in measurement
- Cyclical changes – e.g. with economic cycles

- Probably the most tractable form of uncertainty
- Can usually be adequately quantified by:
 1. Considered use of variants
 2. Probabilistic methods

Error in input data

- i.e. **estimated population** and **migration**
- Migration estimates prone to **systematic** error
- Accuracy of population estimates is cyclical with census
 - Population rolled forward from most recent census
 - Systematic errors **compound** over time
- We're approaching peak uncertainty...

Assumptions

- Projections incorporate a lot of assumptions about the future:
- Whether they are *explicit* or *implicit* in data
- The uncertainty of assumptions often not readily quantifiable
e.g.

Future migration policy and the relative attractiveness of the country

Future housing policy, delivery of new homes, affordability



Model incompleteness

- Our models are informed simplifications of reality
- Capture what we think is most important
- But always incomplete

Model incompleteness

Example:

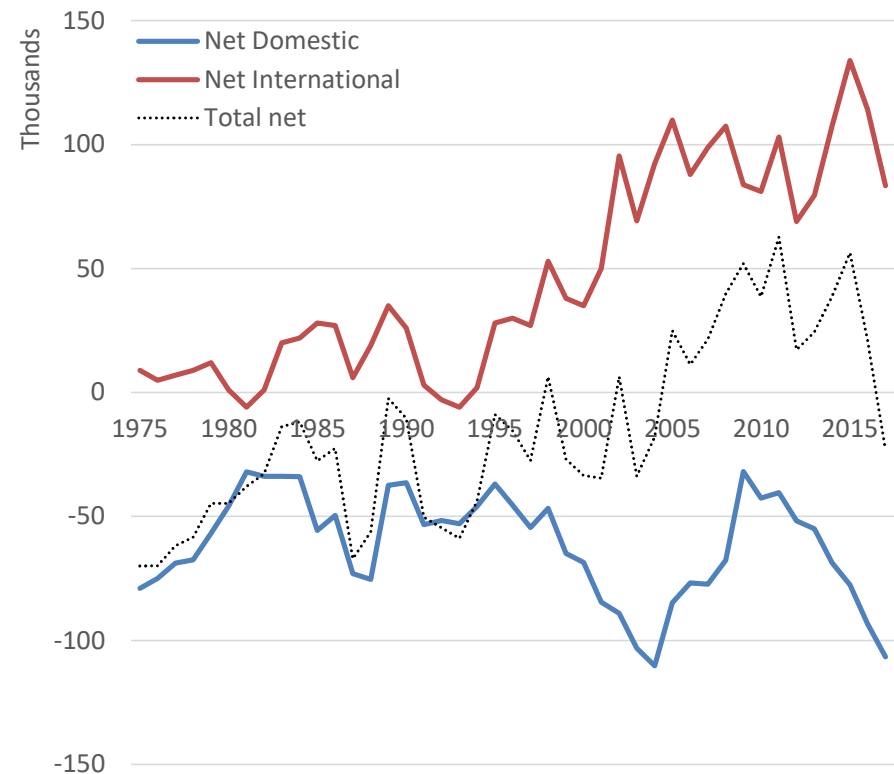
- Domestic migration and fertility reflect recent patterns and structure changes in age structure
- But do not capture **direct** impacts of international migration
 - Domestic flows from displacement by/onward moves of migrants*
 - Tempo fertility effects **

* Gordon et al: *Migration Influences and Implications for Population Dynamics in the Wider South East*

** Robards and Berrington: *The fertility of recent migrants to England and Wales*

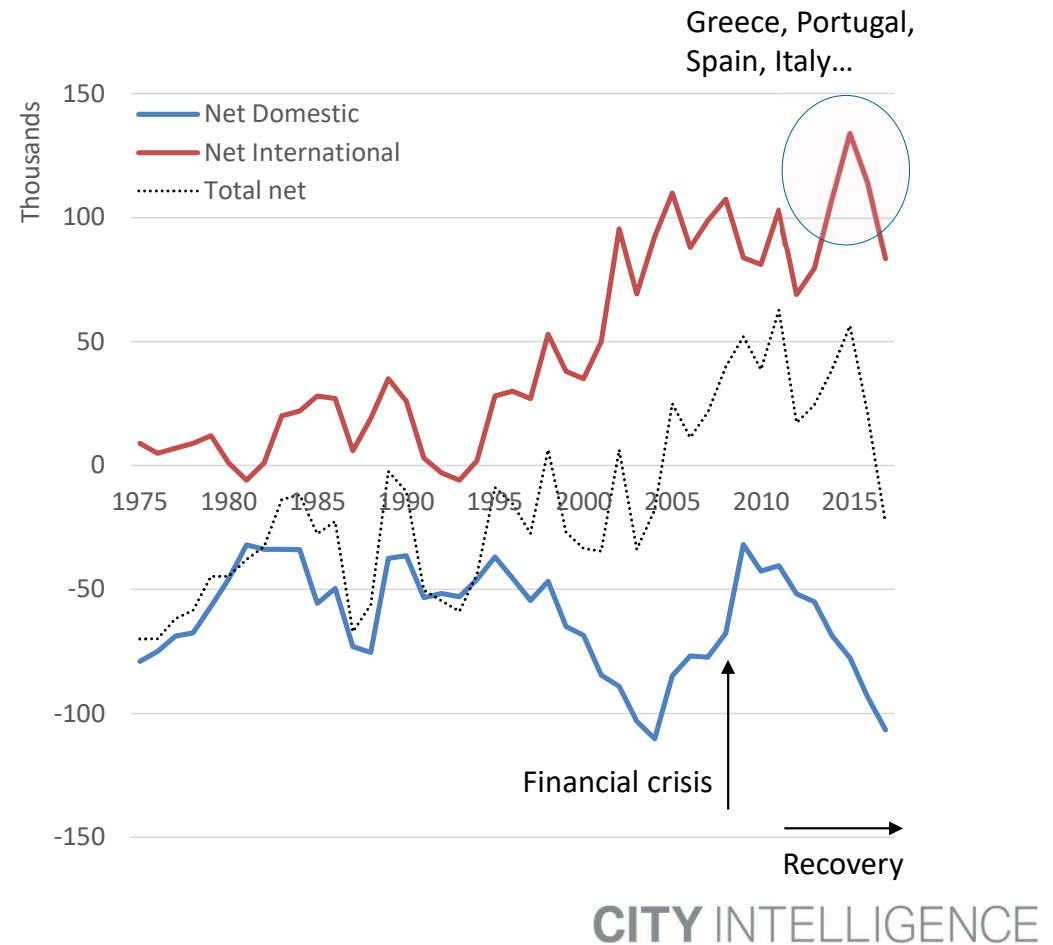
Model incompleteness

- Domestic outflows mirror net international inflows



Model incompleteness

- Domestic outflows mirror net international inflows
- Some additional complexity to consider too



Complexity

- Reality is complex and full of interactions and feedbacks between individuals, institutions and the wider environment
- Unexpected patterns can emerge as a result
- Simple aggregate-level cohort component models don't reflect such dynamics

No trend model based on recent data gives alternative to continued population growth – chance of reversal is zero?

*What kind of model that would have predicted young people would **willingly** move to Peckham?*



Coping strategies

- Understand limitations
- Diversifying models and data
- Communication

Understand limitations

“I do not fancy I know what I do not know.”

- Important to think critically about the elements of the projection and their limitations:
 - Input data, methodology, assumptions
- Data sources are too often treated naïvely by users
- Understand the provenance of input data
 - How it is collected/processed
 - Quality / coverage / bias
 - Consistency over time

■ Diversifying models and data

The 'standard' approach is to attempt to identify and overcome weaknesses – i.e. *do the same but better*

- Increase fidelity of the model
- Improve data sources or work around their limitations
- Refine assumptions or produce variants

Important, but don't solve underlying problems

- Projections vulnerable to a single weakness
- Difficult to provide a meaningful indication of uncertainty

Diversifying models and data

Strength in numbers

An alternative strategy is to diversify the methods and data we use

- Base inputs on multiple data sources to improve robustness
- Test a wide range of assumptions and scenarios
- Use multiple *independent* models and interpret their results in combination

■ Diversifying data

- Identify additional information that can be used to *monitor, validate, improve, or quantify uncertainty* in the principal data sources
- Alternative sources invariably have more limitations than, e.g. official estimates
- But we can still improve the robustness of our inputs by borrowing from their various strengths

Diversifying inputs

Example: adjusting migration estimates for children

- Minor change in estimates methodology after 2011 has big effect on subsequent estimates of children in central London
- Issue spotted through monitoring estimates against:
 - Births by cohort, school census data, GP registration data
- Used timeseries of administrative data to estimate regional patterns of cohort change
- Account for domestic migration and deaths
- Remainder assumed result of net international migration
- Use as basis for revised international migration flows

Diversifying models

Strength in numbers

- The *Diversity Prediction Theorem** suggests we can reduce the error of predictions by combining results of multiple models
- More models = better predictions!

**For N independent models:*

$$(\bar{M} - V)^2 = \sum_i^N \frac{(M_i - V)^2}{N} - \sum_i^N \frac{(M_i - \bar{M})^2}{N}$$

V = true value, M_i = i^{th} model's prediction, \bar{M} = average of predictions

■ Diversifying models

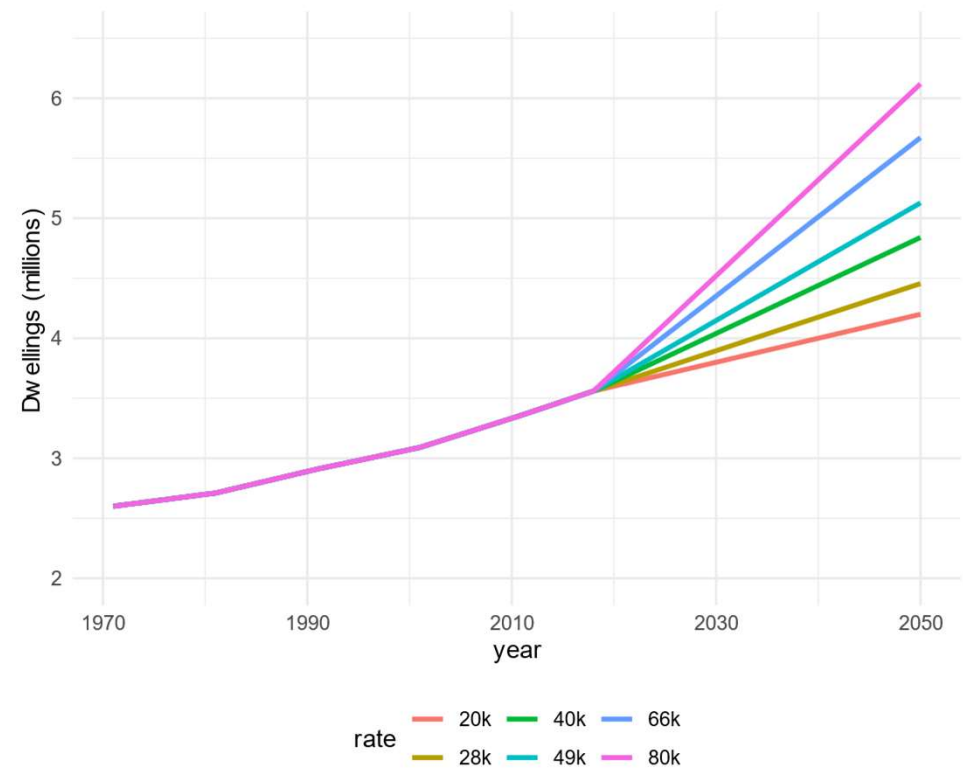
- In practice, usually limited by the number of good, independent, models we can create
- It's ok for the models to have limitations – we'd just like them to have *different* limitations
- The GLA currently makes use of two alternative approaches

■ Diversifying models – housing-led projections

- Adapt cohort component model to give results consistent with scenarios of future housing development
- Capacity of housing stock used as mechanism to govern migration/population
- Range of plausible outcomes for both:
 - Number of new homes that will be built in future
 - Future density of occupation

Housing-led projections

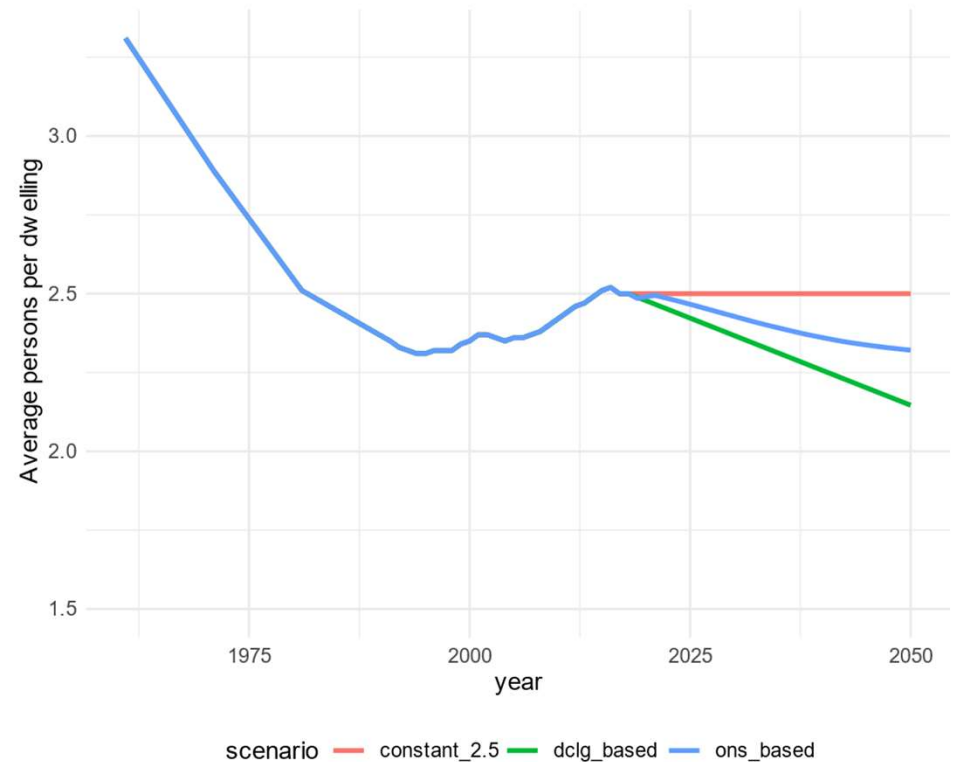
- Range of possible house building scenarios
- Consideration of past trends and policy goals



Housing-led projections

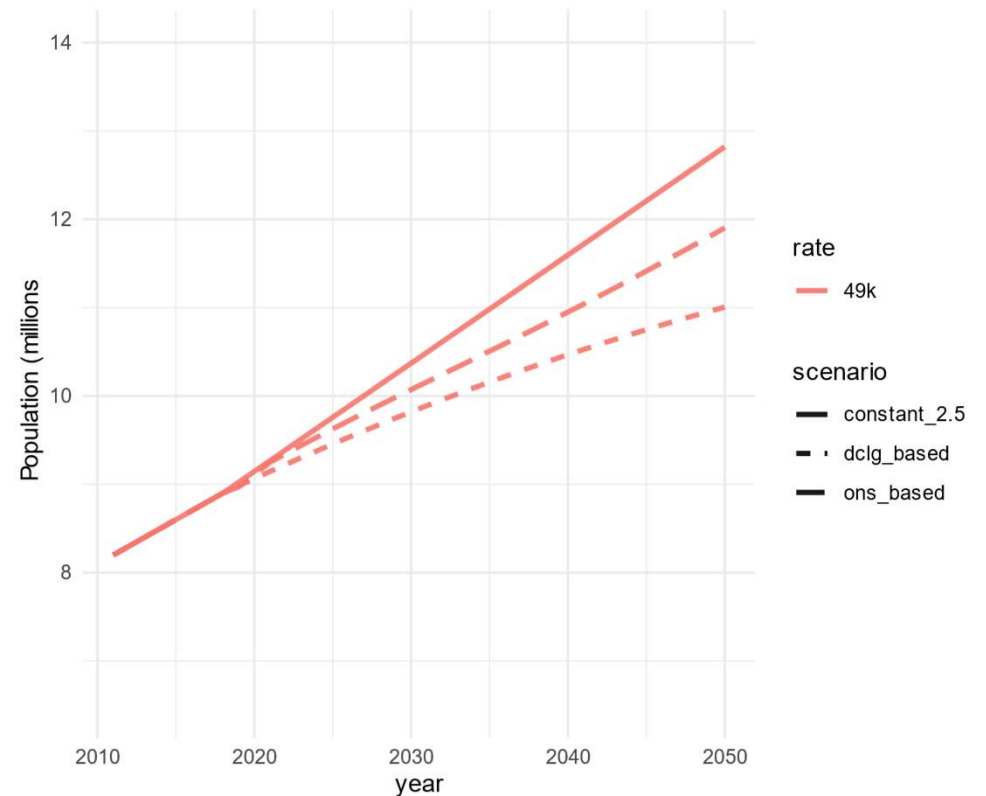
Person per dwelling scenarios

- Application of age-specific household formation rates
 - ONS 2016-based
 - DCLG 2014-based
- Constant at current number



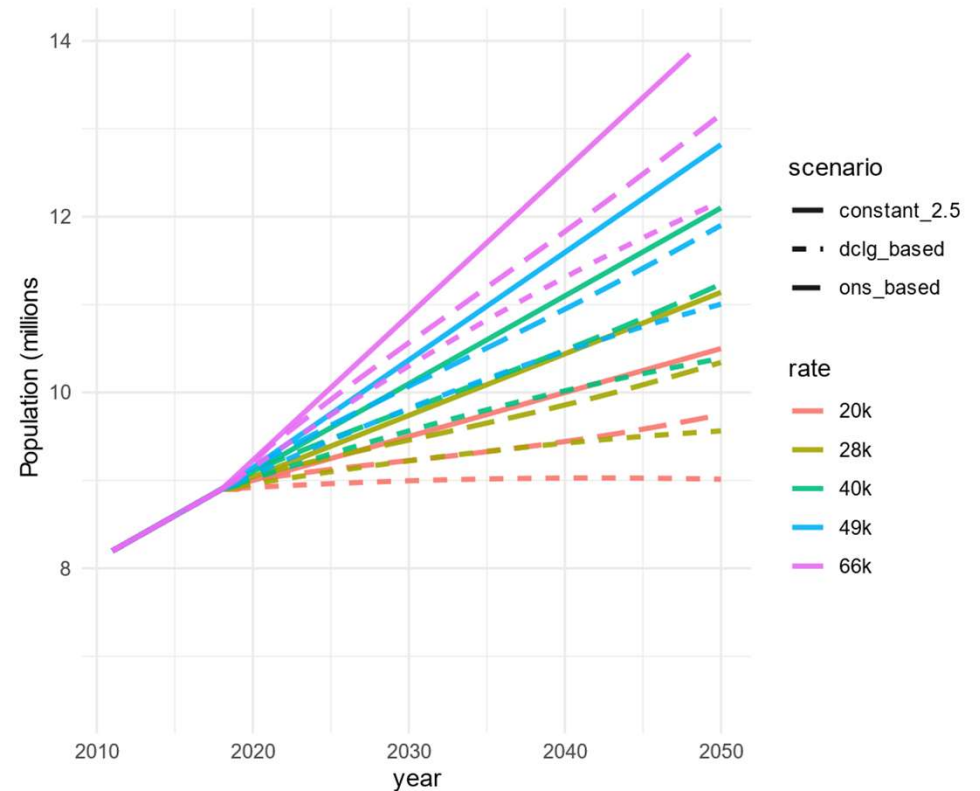
Housing-led projections

- Obtain projected population for each combination of housing trajectory and occupation assumption



Housing-led projections

- Obtain projected population for each combination of housing trajectory and occupation assumption



Diversifying models – employment-led projections

- Similar to housing-led approach, but availability of employment used as driver of population change
- Consistent projections of future GVA and jobs growth produced by colleagues in GLA Economics
- GVA growth scenarios
 - Central: 2.5% → 2.0%
 - High: 3.0% → 2.5%
 - Low: 2.5% → 1.5%

<https://www.london.gov.uk/business-and-economy-publications/london-labour-market-projections-2017>

GLA Experimental Employment-led Projections (<https://data.london.gov.uk/dataset/projections-documentation>)

Employment-led projections

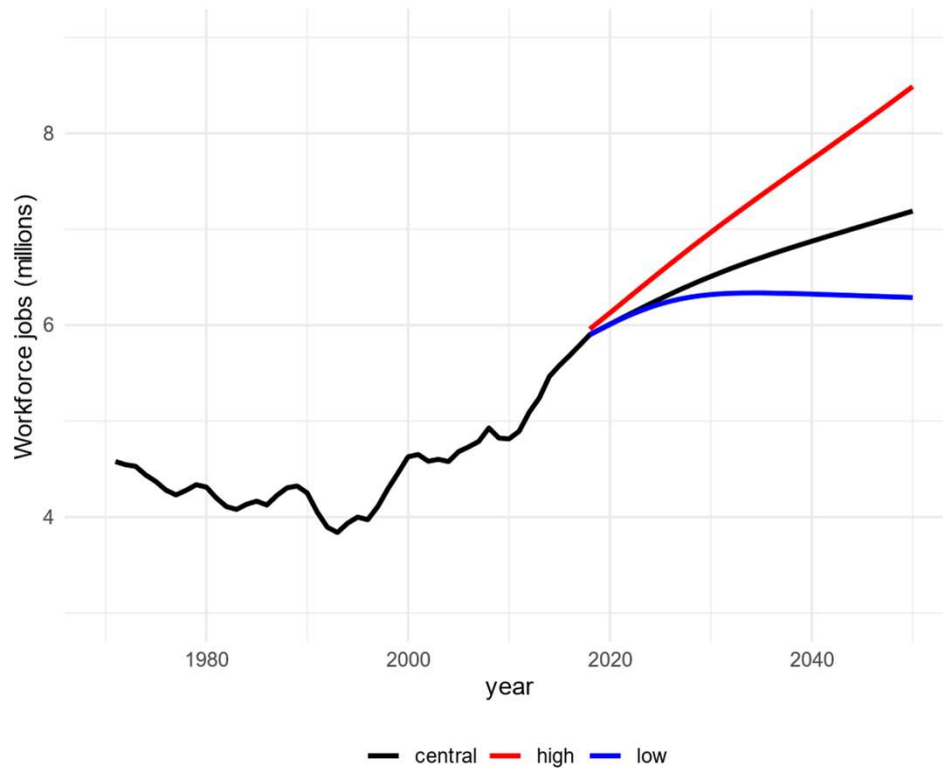
Model iteratively adjusts population for each year until unemployment consistent with target rate

$$u_y = (e \cdot p_y) - l_y^s - o_y^s$$

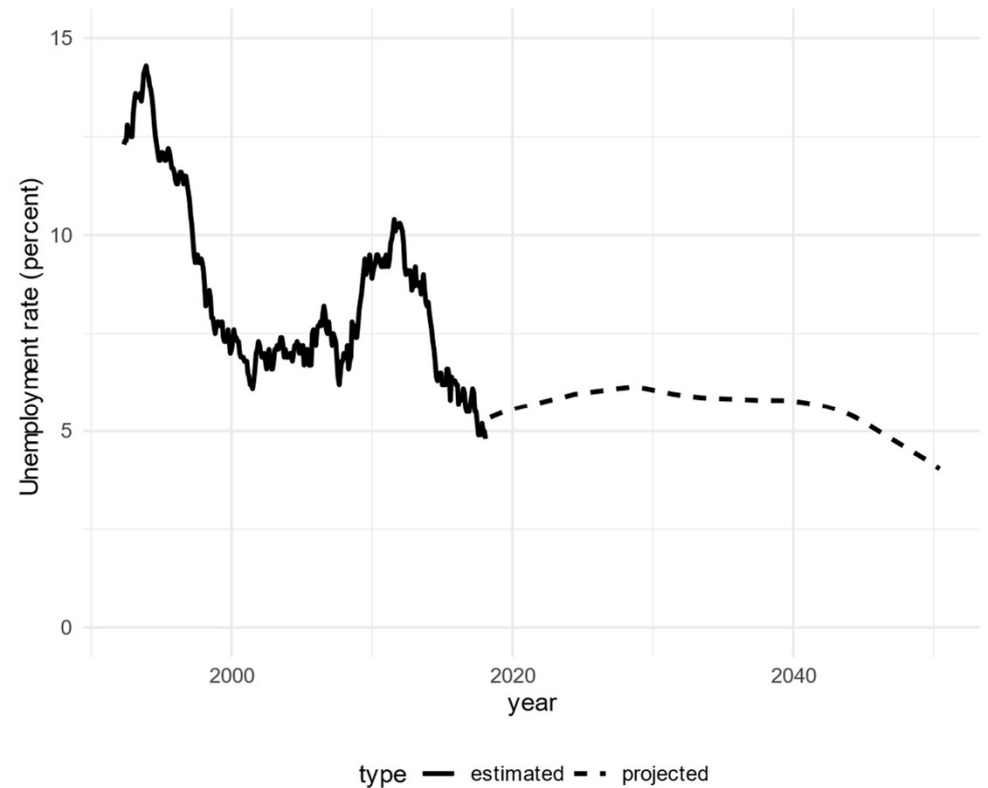
- u = unemployment rate
- y = year
- e = economic activity rate
- p = population
- l = London residents who work in London
- o = London residents who work outside London
- s = scenario

Employment-led projections

Projected workforce jobs

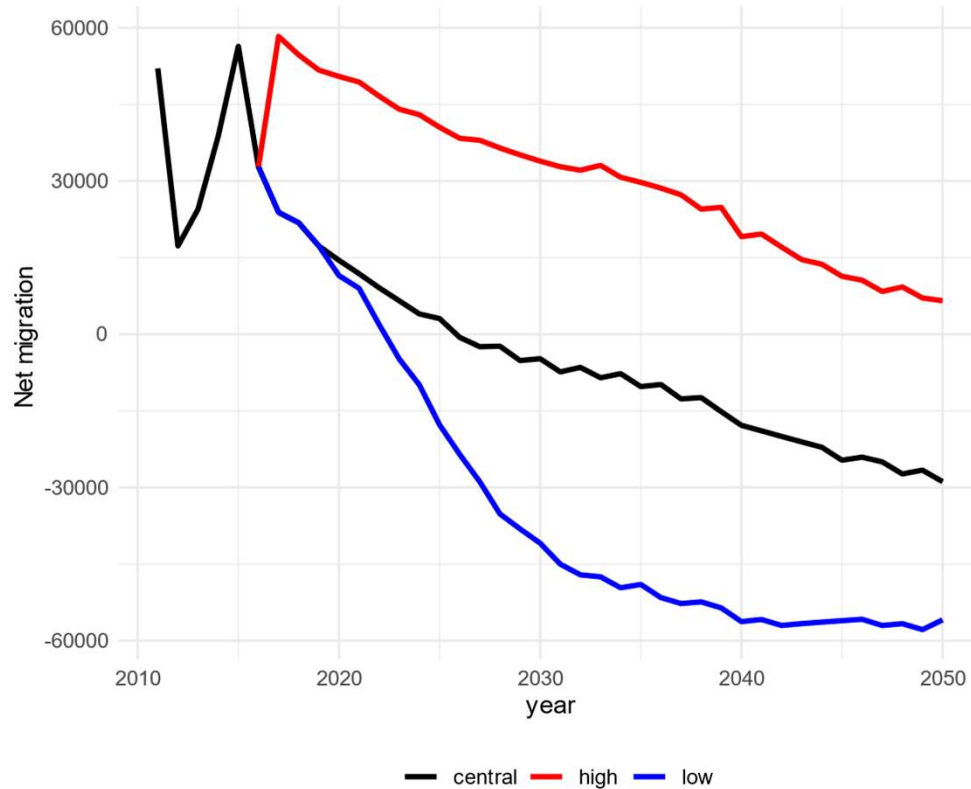


Assumed unemployment rate

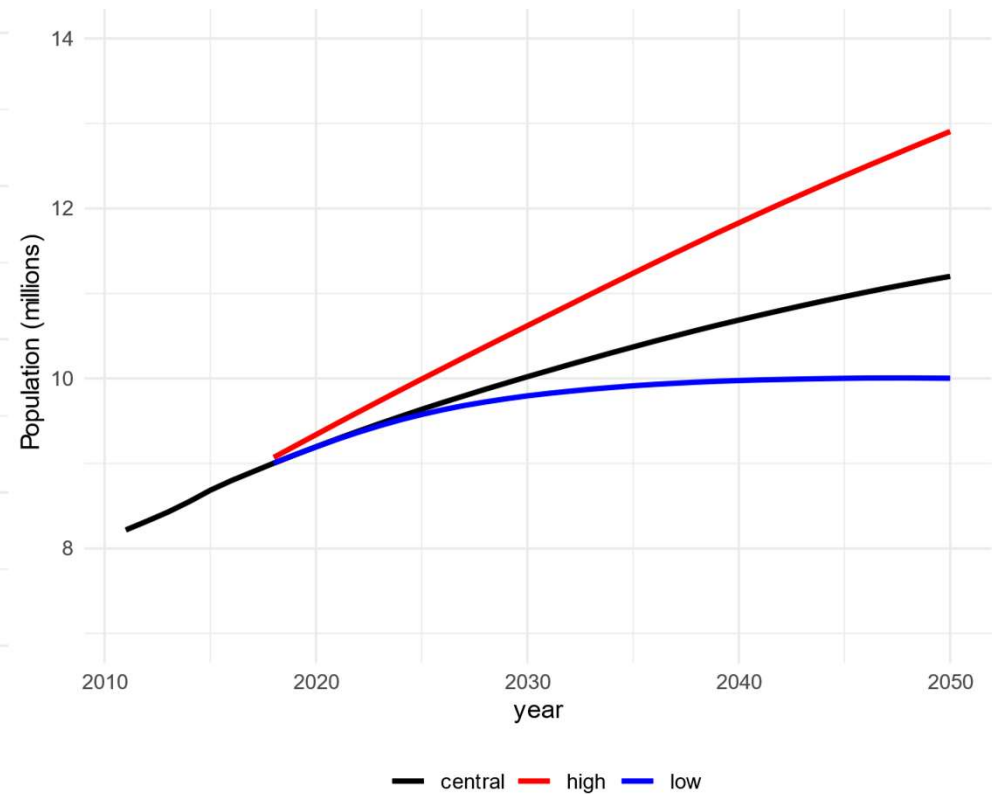


Employment-led projections

Projected net migration



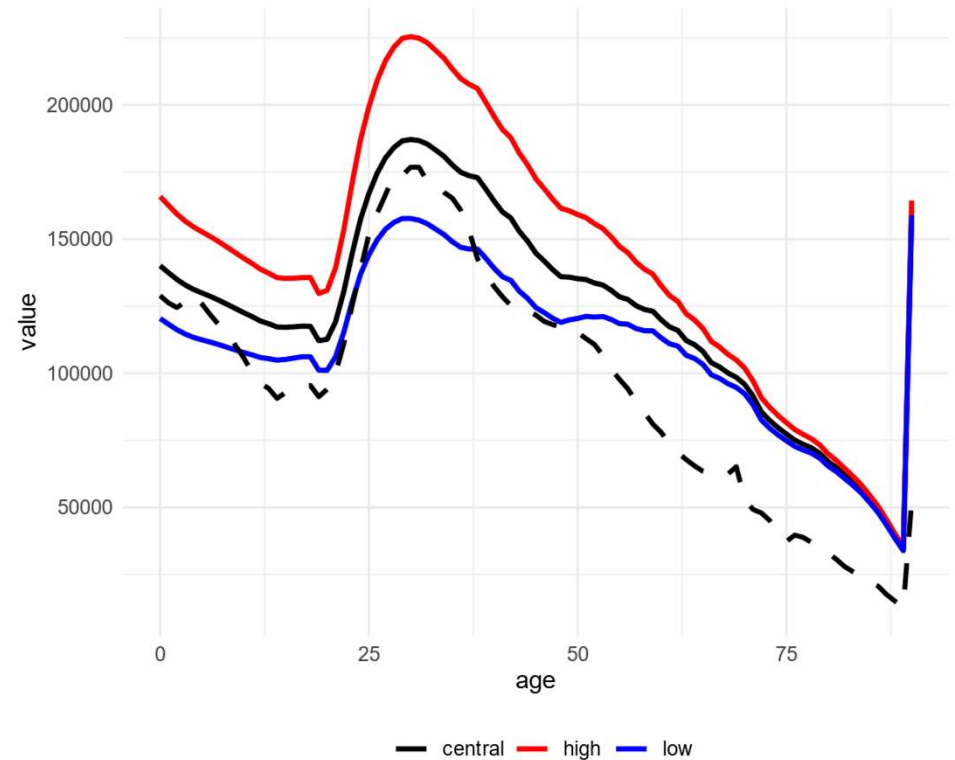
Projected total population



Employment-led projections

- Output of model can be used as input to household model
- Create consistent projections of GVA, jobs, population, housing need...

Projected population by age vs 2016



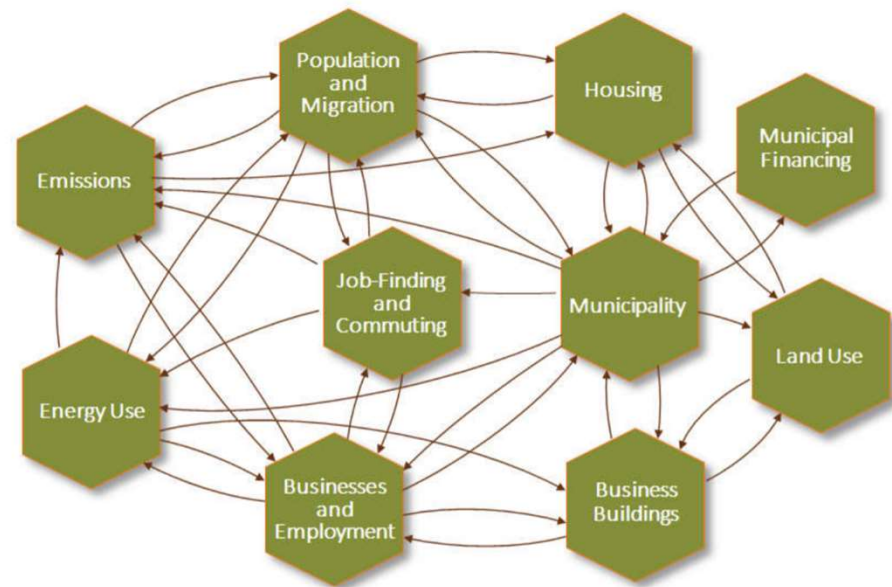
■ Diversifying models

Additional approaches being investigated

- Systems Dynamics
- Spatial interaction
- Agent based

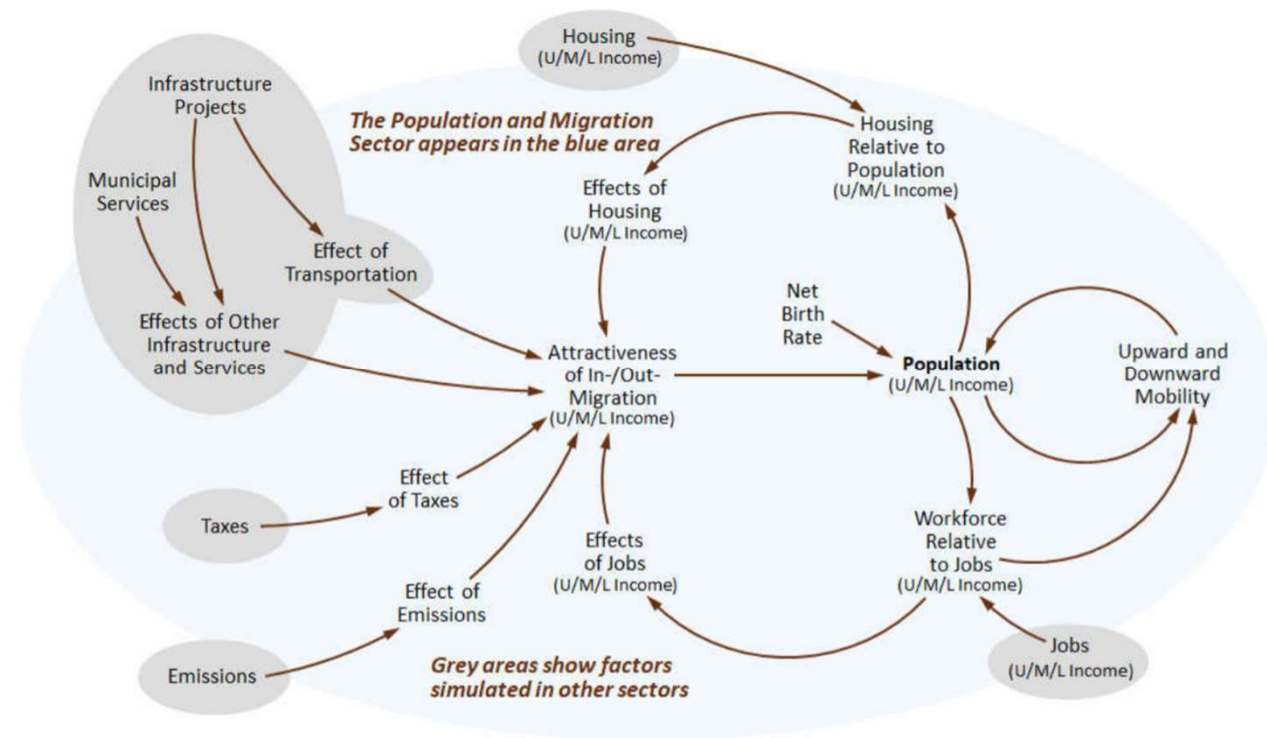
Diversifying models - Systems Dynamics

- SD models represent systems in terms of stocks, flows and links between them
- Good at capturing interrelationships and feedbacks between wide range of elements
- Challenging to incorporate high levels of spatial or categorical detail



Diversifying models – Systems Dynamics

- *The London Simulator* developed with external partners
- Primarily use assessing impacts of infrastructure investment
- Can explore potential for general population modelling use



■ Diversifying models – Agent Based Models

- Represent people/dwellings/institutions as distinct individuals
- Model interactions between individuals and with environment
- Applications in social science as explanatory or exploratory models
- Use as predictive models has been more limited

■ Diversifying models – Agent Based Models

Predictive ABMs:

- Have high input data requirements
- Are computationally intensive
- Can be challenging to validate

But - have the potential to reflect complex dynamic behaviour far better than standard approaches

■ Diversifying models – Agent Based Models

Potential to use ABMs has increased rapidly:

- Input data improving through new sources and use of microsimulation
- Cloud platforms helping to remove computational barriers
- Frameworks and protocols to design and structure ABMs have emerged (ODD+D and POM)

Effective communication

Understanding the audience(s)

- Uncertainty not an easy concept for many users of projections
- Aim to present the information in a way that will encourage them to make use of it
- Need to understand how users interpret information about uncertainty

Effective communication

- How much does it help, in practice, to present users with statistical measures?

“The projected 2040 population of 10.5 – 11.2 million is within one standard deviation of the mean output”

Effective communication

- How much does it help, in practice, to present users with statistical measures?
- Is something like this more useful?

*“The projected 2040 population of 11 million is consistent with:
a continuation of the last decade’s migration trends;
annual housing delivery of 35 to 50 thousand units;
annual GVA growth of 2.2 to 2.7%”*

Effective communication

- Develop set of coherent scenarios that users can relate to
 - Realistic
 - Challenging
 - Internally consistent
- Produce suite of projections that embody these
 - Demographic, economic, transport
- Use to stress-test plans and policies



Contact:

demography@london.gov.uk

ben.corr@london.gov.uk