URBAN EXODUS OR RURAL SHRINKAGE?

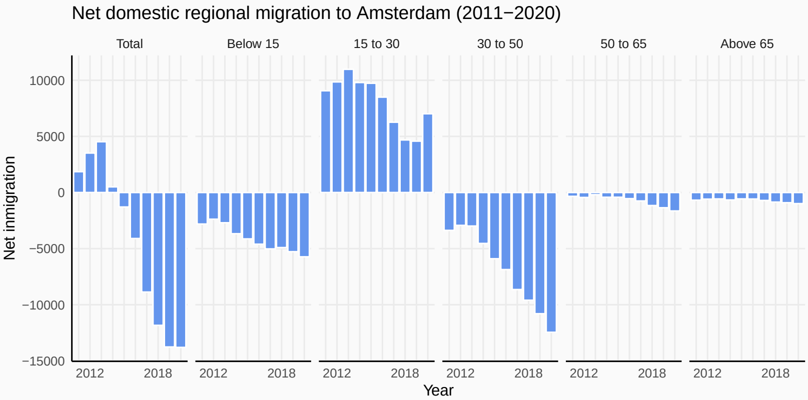
Regional migration and attractiveness in a tight Dutch housing market

Thomas de Graaff

March 17, 2022

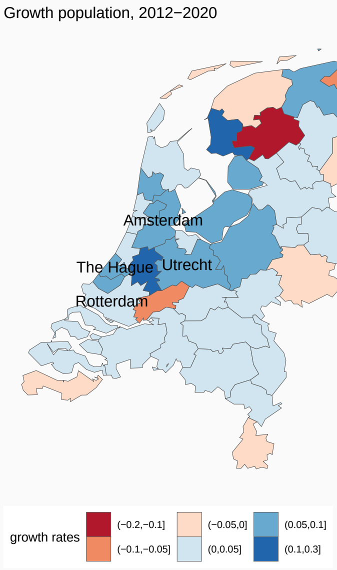
Vrije Universiteit Amsterdam
Tinbergen Institute Amsterdam

Urban Exodus?



Dutch population growth 2012–2020

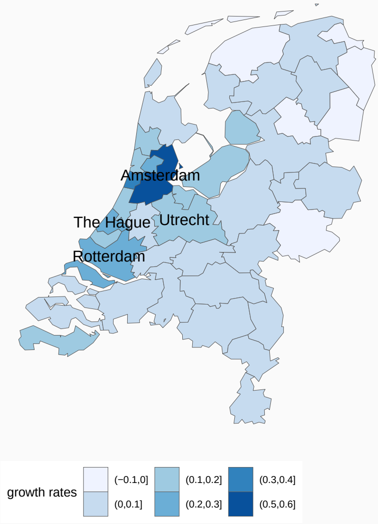
- NUTS-3 regions
  - originally (1970) labour market regions
- Last decade:
  - homogeneous population growth
  - few peripheral regions decline
- Domestic migration
  - slightly more within regions than between
  - growth is the same



Tight Dutch housing market

- Average housing price: €410,000
- Change last year +20%
- Waiting list social renting Amsterdam: 13 years
- Large shortage of housing
- Decrease in housing transactions

Growth property tax value, 2012−2020



Housing market, urban regions and interregional migration: why bother?

Possible drivers of urban out-migration?

- suburbanisation of poverty (Hochstenbach and Musterd, 2018)
- crowding-out of the housing market by short-term rentals (Koster et al., 2021)
- Influx of high-skilled migrants (Beckers and Boschman, 2019)
- Housing market structure (external effects of home-ownership Dietz and Haurin (2003))
  - negative: moving costs of home-ownership (and social renting) (Oswald, 1996, 1999)

My contributions to the literature

- Large empirical (economic) literature on impact housing market structure as driver of interregional migration, but:
  - usually focuses on marginal effect of home-ownership
  - less attention for (asymetric) network effects (e.g., push vs. pull effects of larger cities)
- Literature on impact of social renting on migration flows is scarce (De Graaff et al., 2009)
  - In the Netherlands social renting is a large phenomenon ($\approx$ 24% of total housing stock)
  - Social renting rights only valid within city/region
  - Social renting is an urban phenomenon (e.g. $\approx$ 30–40% in Amsterdam)

So, this paper

Does what? Estimates the impact of housing market structure on Dutch interregional migration flows using a multilevel gravity model

- UK context by Congdon (2010)
- social relations model *cf.* Koster and Leckie (2014) and Zhang et al. (2020)

Aim To simultaneously assess the impact of housing market structure and region specific effects on domestic migration flows

- home-ownership and social renting
- household size
- percentage western immigrants

Why a multilevel approach for modeling (migration) flows?

There are at least two levels in migration (I use three)

Why a multilevel approach for modeling (migration) flows?

There are at least two levels in migration (I use three)

Observed migration flows Migration between i and j with friction (e.g., distance) attributes (obs $= R^2 - R$)

$$\text{REGION}_i \xrightarrow{\mathbf{x}_{ij}} \text{REGION}_j$$

Why a multilevel approach for modeling (migration) flows?

There are at least two levels in migration (I use three)

Observed migration flows Migration between i and j with friction (e.g., distance) attributes (obs $= R^2 - R$)

Observed push & pull factors Attributes of i and j (obs $= R$)

Why a multilevel approach for modeling (migration) flows?

There are at least two levels in migration (I use three)

Observed migration flows Migration between i and j with friction (e.g., distance) attributes (obs $= R^2 - R$)

$$\text{REGION}_i \xrightarrow{\ \mathbf{X}_{ij}\ } \text{REGION}_j$$

Observed push & pull factors Attributes of i and j (obs $= R$)

$$\mathbf{X}_i \rightarrow \text{REGION}_i \qquad \mathbf{X}_j \rightarrow \text{REGION}_j$$

Observed flows within regional dyads migration from $i \rightarrow j$ is correlated with migration from $j \rightarrow i$ (obs $= \frac{R^2-R}{2}$)

$$\text{REGION}_i \rightleftarrows \text{REGION}_j$$

Why a Bayesian multilevel approach?

- Multilevel, hierarchical, mixed effects, varying intercept/parameter, shrinkage, partial pooling models

Why a Bayesian multilevel approach?

- Multilevel, hierarchical, mixed effects, varying intercept/parameter, shrinkage, partial pooling models
- Increasingly used for model performance and flexibility

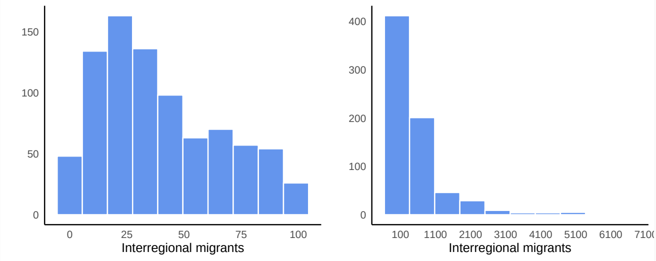
Why a Bayesian multilevel approach?

- Multilevel, hierarchical, mixed effects, varying intercept/parameter, shrinkage, partial pooling models
- Increasingly used for model performance and flexibility
- Simultaneous modeling at various levels (e.g., cities, regions, flows, individuals)
  - no two-stage models anymore
  - precision (standard errors) is correct at all levels

Why a Bayesian multilevel approach?

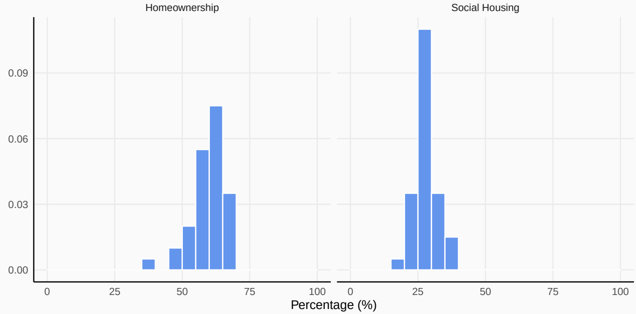
- Multilevel, hierarchical, mixed effects, varying intercept/parameter, shrinkage, partial pooling models
- Increasingly used for model performance and flexibility
- Simultaneous modeling at various levels (e.g., cities, regions, flows, individuals)
  - no two-stage models anymore
  - precision (standard errors) is correct at all levels
- Partial pooling: For example, origin specific effects are drawn from a distribution: $o_i \sim \mathcal{N}(0, \sigma)$
  - $\sigma \longrightarrow 0$: complete pooling
  - $\sigma \longrightarrow \infty$: no pooling (fixed effects)

Data: migrations flows in 2018



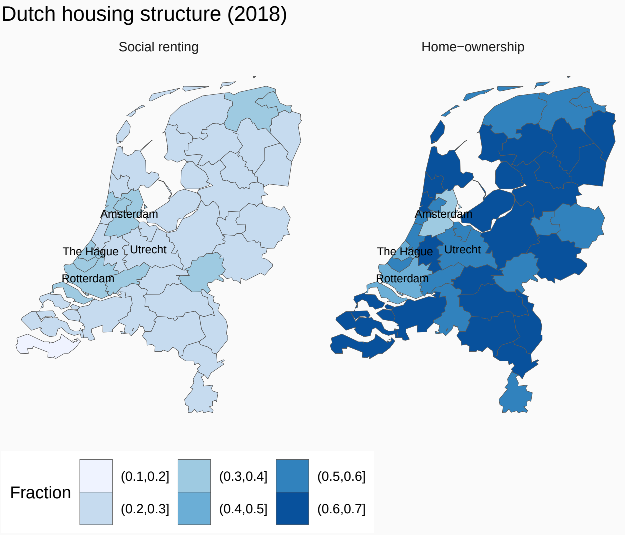
- Panel for the period 2012–2020
  - estimation: 2012–2019
  - out-of-sample prediction: 2020
- Migration flows between 40 Dutch regions
- Variance $\gg$ mean: over-dispersion

Data: regional housing structure in 2018

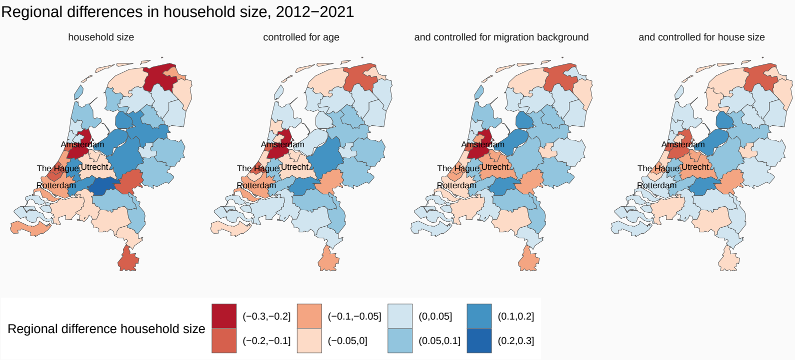


- Positive correlation between regional population and share social renting (0.46)
- Negative correlation between regional share social renting and share home-ownership (-0.88)

Data: regional housing structure in 2018 (cont.)



Data: regional household size



Modeling framework: traditional gravity modeling

$$\ln(\text{Migrants}_{ij}) = o_i + d_j + \gamma \ln(\text{dist}_{ij}) + \epsilon_{ij}$$

Origin and destination specific regional effects for multilateral resistance (Anderson and Van Wincoop, 2003), but:

- what about zeros in Migrants_{ij}?
- how to incorporate housing structure in the presence of o_i and d_j?
- over-dispersion and heteroskedasticity (Silva and Tenreyro, 2006)

Poisson versus negative binomial[1]

- Counts of migrants

- Constraints should hold

$$\sum_{j=1}^{R} \widehat{\text{Migrants}}_{ij} = O_i \qquad \sum_{i=1}^{R} \widehat{\text{Migrants}}_{ij} = D_j$$

  - poisson: ✓
  - negative binomial: ✗

- multilevel structure controls for overdispersion

[1] *We urge researchers to resist the siren song of the Negative Binomial (Head and Mayer, 2014)*

Modeling framework: multilevel gravity modeling

$\text{Migrants}_{ijt} \sim \text{Poisson}(\lambda_{ijt})$ (flow of migrants)

Modeling framework: multilevel gravity modeling

$$
\begin{aligned}
\text{Migrants}_{ijt} \sim & \text{Poisson}(\lambda_{ijt}) & \text{(flow of migrants)} \\
\ln(\lambda_{ijt}) = & \alpha + o_i + d_j + t_t + \text{dyad}_{ij} + \\
& \beta_1 \ln(\text{pop}_{it}) + \beta_2 \ln(\text{pop}_{jt}) + \gamma \ln(\text{dist}_{ijt}) + \\
& \beta_3 \ln(\text{home}_{it}) + \beta_4 \ln(\text{home}_{jt}) + \beta_5 \ln(\text{soc}_{it}) + \beta_6 \ln(\text{soc}_{jt}) + \\
& \beta_7 \ln(\text{hhsize}_{it}) + \beta_8 \ln(\text{hhsize}_{jt}) + \\
& \beta_9 \ln(\text{perc\_wi}_{it}) + \beta_{10} \ln(\text{perc\_wi}_{jt}) & \text{(linear model)}
\end{aligned}
$$

Modeling framework: multilevel gravity modeling

$$
\begin{aligned}
\text{Migrants}_{ijt} \sim & \text{Poisson}(\lambda_{ijt}) && \text{(flow of migrants)} \\
\ln(\lambda_{ijt}) = & \alpha + o_i + d_j + t_t + \text{dyad}_{ij} + \\
& \beta_1 \ln(\text{pop}_{it}) + \beta_2 \ln(\text{pop}_{jt}) + \gamma \ln(\text{dist}_{ijt}) + \\
& \beta_3 \ln(\text{home}_{it}) + \beta_4 \ln(\text{home}_{jt}) + \beta_5 \ln(\text{soc}_{it}) + \beta_6 \ln(\text{soc}_{jt}) + \\
& \beta_7 \ln(\text{hhsize}_{it}) + \beta_8 \ln(\text{hhsize}_{jt}) + \\
& \beta_9 \ln(\text{perc\_wi}_{it}) + \beta_{10} \ln(\text{perc\_wi}_{jt}) && \text{(linear model)} \\
\begin{pmatrix} o_i \\ d_j \end{pmatrix} \sim & \mathcal{N}\left\{ \begin{pmatrix} 0 \\ 0 \end{pmatrix}, \begin{pmatrix} \sigma_i^2 & \rho_{ij} \\ \rho_{ij} & \sigma_j^2 \end{pmatrix} \right\} && \text{(regional varying effects)}
\end{aligned}
$$

Modeling framework: multilevel gravity modeling

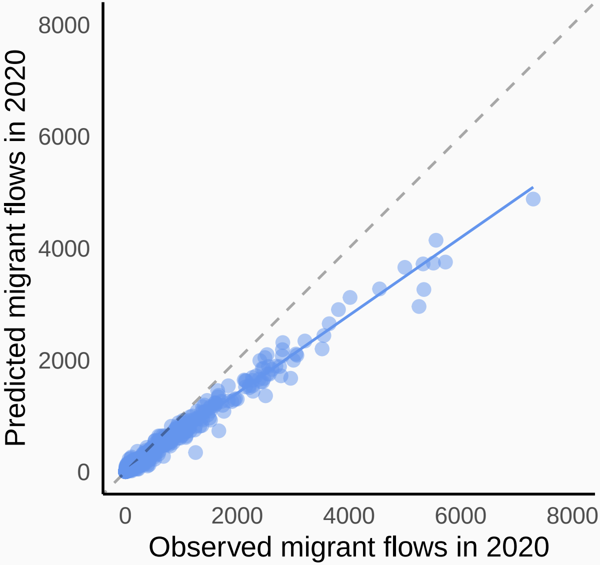
$$
\begin{aligned}
\text{Migrants}_{ijt} \sim & \text{Poisson}(\lambda_{ijt}) && \text{(flow of migrants)} \\
\ln(\lambda_{ijt}) = & \alpha + o_i + d_j + t_t + \text{dyad}_{ij} + \\
& \beta_1 \ln(\text{pop}_{it}) + \beta_2 \ln(\text{pop}_{jt}) + \gamma \ln(\text{dist}_{ijt}) + \\
& \beta_3 \ln(\text{home}_{it}) + \beta_4 \ln(\text{home}_{jt}) + \beta_5 \ln(\text{soc}_{it}) + \beta_6 \ln(\text{soc}_{jt}) + \\
& \beta_7 \ln(\text{hhsize}_{it}) + \beta_8 \ln(\text{hhsize}_{jt}) + \\
& \beta_9 \ln(\text{perc\_wi}_{it}) + \beta_{10} \ln(\text{perc\_wi}_{jt}) && \text{(linear model)} \\
\begin{pmatrix} o_i \\ d_j \end{pmatrix} \sim & \mathcal{N}\left\{ \begin{pmatrix} 0 \\ 0 \end{pmatrix}, \begin{pmatrix} \sigma_i^2 & \rho_{ij} \\ \rho_{ij} & \sigma_j^2 \end{pmatrix} \right\} && \text{(regional varying effects)} \\
\begin{pmatrix} \text{dyad}_{ij} \\ \text{dyad}_{ji} \end{pmatrix} \sim & \mathcal{N}\left\{ \begin{pmatrix} 0 \\ 0 \end{pmatrix}, \begin{pmatrix} \sigma_{\text{dyad}}^2 & \rho \\ \rho & \sigma_{\text{dyad}}^2 \end{pmatrix} \right\} \\
& && \text{(dyad varying effects)}
\end{aligned}
$$

Main Estimation results

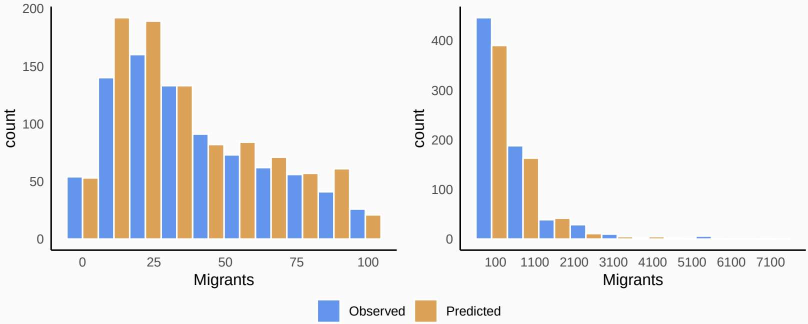
parameter	origin (push)	destination (pull)
ln(population)	**−0.10**	**0.70**
ln(% homeownership)	**−1.73**	**1.37**
ln(% social renting)	**−0.40**	**0.99**
ln(household size)	**5.46**	**−2.35**
ln(% western immigrants)	**−0.14**	−0.01
intercept	**3.89**	
migrants flow:		
ln(distance)	**−1.63**	
standard deviations:		
origin	**0.67**	
destination	**0.44**	
dyad	**0.39**	
correlation		
origin-destination	**0.78**	
dyad	**0.80**	

Bold: 89% credible intervals do not include zero
Samples are drawn using the NUTS sampler from STAN using 4 chains, each with 4,000 iterations and 1,000 warm-up samples

Out-of-sample prediction for 2020 ($R^2 = 0.98$)



Out-of-sample prediction for 2020 (cntd.)



Correlation patterns

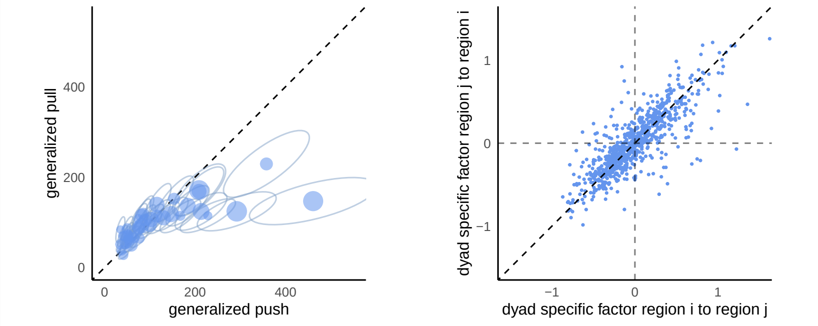
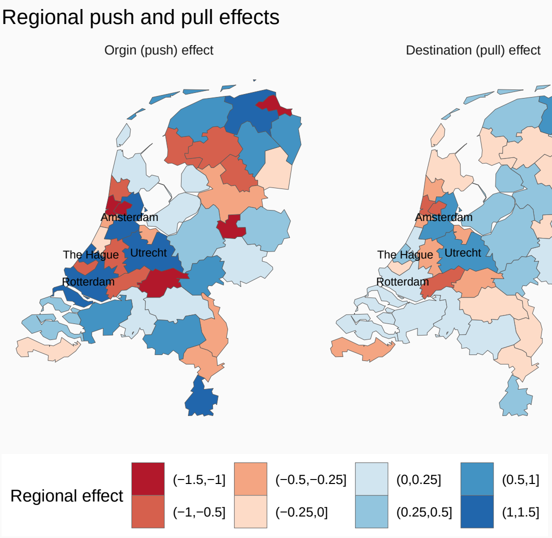


Figure 1: Correlation (0.78) between unobserved push and pull factors region (left) and flows (correlation = 0.8) within dyad pairs (right)

Asymmetric push and pull factors



Sensitivity check: spatial autocorrelation

- spatial autocorrelation in regional effects:

$$
\begin{aligned}
o_i, d_j &\sim \text{MVNormal}(0, \mathbf{K}) \\
\mathbf{K}_{ij} &= \eta^2 \exp(-\rho^2 \mathbf{D}_{ij})
\end{aligned}
$$

- results remain robust

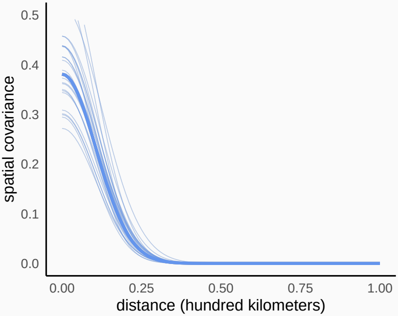
Sensitivity check: spatial autocorrelation

- spatial autocorrelation in regional effects:

$$
\begin{aligned}
o_i, d_j &\sim \text{MVNormal}(0, \mathbf{K}) \\
\mathbf{K}_{ij} &= \eta^2 \exp(-\rho^2 \mathbf{D}_{ij})
\end{aligned}
$$

- results remain robust

Modest spatial autocorrelation



Conclusions

Main results:

- home-ownership and social renting have negative effect on push and positive impact on pull factors
- household size have positive impact on push and negative on pull factors
- percentage western immigrants small effect
- still large urban areas have large push effects
    - effect is different from housing market structure
    - more dynamic than in periphery

Speculation:

- internationalisation: tourist, short stay (high-skilled), and large housing investment companies drive natives out?

Supplementary materials

Paper, presentation, data and code can be retrieved from the project's GitHub page:

https://github.com/Thdegraaff/migration_gravity

Thank you!

References i

Anderson, J. E. and E. Van Wincoop (2003). "Gravity with gravitas: a solution to the border puzzle". In: *American economic review* 93.1, pp. 170–192.

Beckers, P. and S. Boschman (2019). "Residential choices of foreign highly skilled workers in the Netherlands and the role of neighbourhood and urban regional characteristics". In: *Urban Studies* 56.4, pp. 760–777.

Congdon, P. (2010). "Random-effects models for migration attractivity and retentivity: a Bayesian methodology". In: *Journal of the Royal Statistical Society: Series A (Statistics in Society)* 173.4, pp. 755–774.

De Graaff, T., M. Van Leuvensteijn, and C. Van Ewijk (2009). "Homeownership, social renting and labor mobility across Europe". In: *Homeownership and the labour market in Europe*, pp. 53–81.

Dietz, R. D. and D. R. Haurin (2003). "The social and private micro-level consequences of homeownership". In: *Journal of urban Economics* 54.3, pp. 401–450.

References ii

Head, K. and T. Mayer (2014). "Gravity equations: Workhorse, toolkit, and cookbook". In: *Handbook of international economics*. Vol. 4. Elsevier, pp. 131–195.

Hochstenbach, C. and S. Musterd (2018). "Gentrification and the suburbanization of poverty: Changing urban geographies through boom and bust periods". In: *Urban Geography* 39.1, pp. 26–53.

Koster, H. R., J. van Ommeren, and N. Volkhausen (2021). "Short-term rentals and the housing market: Quasi-experimental evidence from Airbnb in Los Angeles". In: *Journal of Urban Economics* 124, p. 103356.

Koster, J. M. and G. Leckie (2014). "Food sharing networks in lowland Nicaragua: an application of the social relations model to count data". In: *Social Networks* 38, pp. 100–110.

Oswald, A. J. (1996). *A conjecture on the explanation for high unemployment in the industrialized nations: Part I*. Tech. rep.

References iii

Oswald, A. J. (1999). "The housing market and Europe's unemployment: a non-technical paper". In: *Homeownership and the labour Market in Europe*.

Silva, J. S. and S. Tenreyro (2006). "The log of gravity". In: *The Review of Economics and statistics* 88.4, pp. 641–658.

Zhang, X. N., W. W. Wang, R. Harris, and G. Leckie (2020). "Analysing Inter-Provincial Urban Migration Flows in China: A New Multilevel Gravity Model Approach". In: *Migration Studies* 8.1, pp. 19–42.