URBAN EXODUS OR RURAL SHRINKAGE?

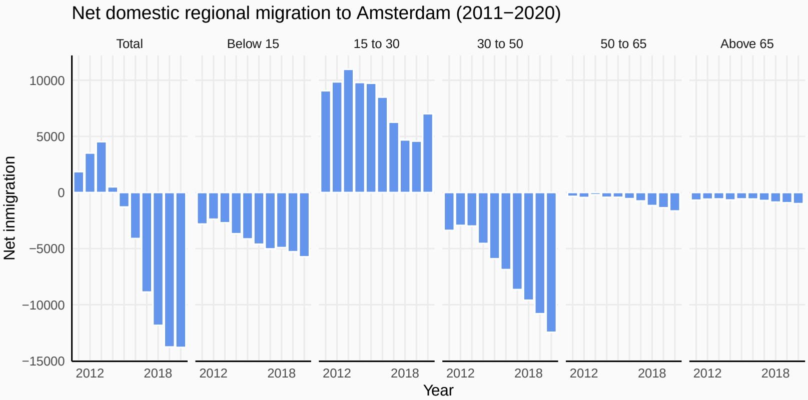
Regional migration and attractiviness in a tight Dutch housing market

Thomas de Graaff

September, 2021

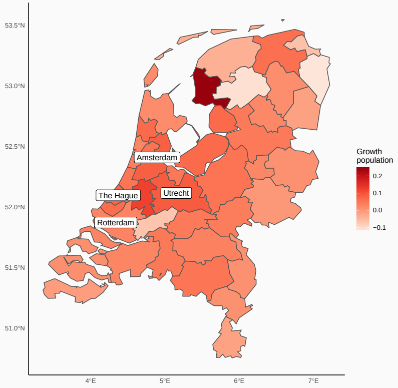
Vrije Universiteit Amsterdam
Tinbergen Institute Amsterdam

Urban Exodus?



Dutch population growth 2012–2020

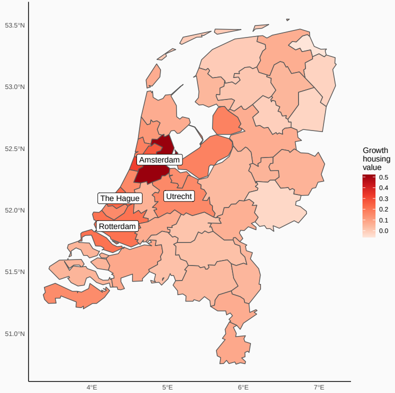
- NUTS-3 regions
  - originally (1970) labour market regions
- Last decade:
  - homogeneous population growth
  - few peripheral regions decline
- Domestic migration
  - slightly more within regions than between
  - growth is the same



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



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, 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)

- Large literature on external effects of home-ownership (Dietz and Haurin, 2003)
  - negative: moving costs (Oswald, 1996, 1999)

My contributions to the literature

- Large empirical (economic) literature on impact home-ownership as drivers of interregional migration, but:
  - usually concerns marginal effect of home-ownership
  - less attention for the whole network (e.g., Amsterdam effect)

- 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)
- *Statistical Rethinking* from McElreath (2020)
- `ggplot2` code from Kurz (2020)

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

Why a multilevel approach for the gravity model?

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

Why a multilevel approach for the gravity model?

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 $= R^2 - R$)

Why a multilevel approach for the gravity model?

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 $= R$)

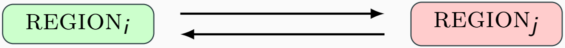
Why a multilevel approach for the gravity model?

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 $= R$)

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



Why a Bayesian multilevel approach?

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

Why a Bayesian multilevel approach?

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

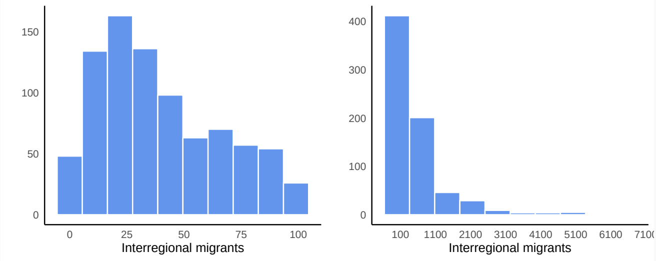
Why a Bayesian multilevel approach?

- 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?

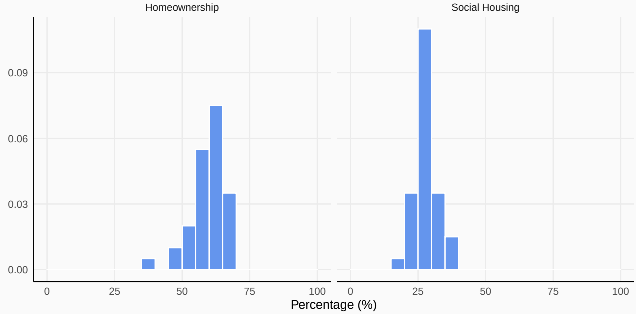
- Hierarchical, mixed effects, varying intercept/parameter, shrinkage, partial pooling models
- Increasingly used for model performance and flexibility
- Simultaneous modeling at various levels (e.g., cities, regions, flows, individuals)
  - no two-stage models anymore
  - precision (standard errors)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

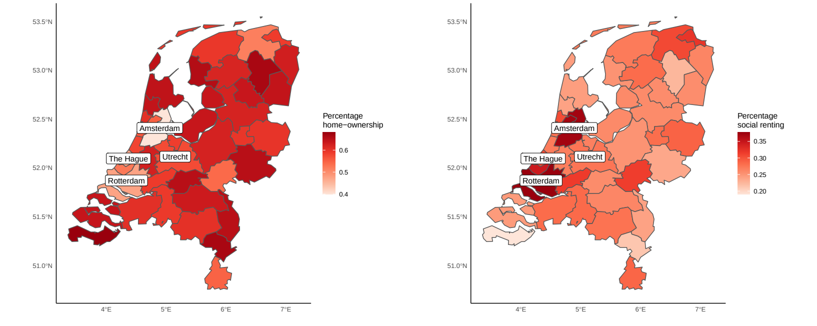


Figure 1: Share of home-ownership (left) and social renting (right)

Modeling framework: traditional gravity modeling

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

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

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

Modeling framework: multilevel gravity modeling

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

$$\begin{aligned}\log(\lambda_{ijt}) = & \alpha + o_i + d_j + t_t + \text{dyad}_{ij} + \\ & \beta_1 \ln(\text{pop}_{it}) + \beta_2 \ln(\text{pop}_{jt}) + \gamma \ln(\text{dist}_{ijt}) + \\ & \beta_3 \ln(\text{home}_{it}) + \beta_4 \ln(\text{home}_{jt}) + \beta_5 \ln(\text{soc}_{it}) + \beta_6 \ln(\text{soc}_{jt})\end{aligned} \qquad \text{(linear model)}$$

$$\begin{pmatrix} o_i \\ d_j \end{pmatrix} \sim \mathcal{N} \left\{ \begin{pmatrix} 0 \\ 0 \end{pmatrix}, \begin{pmatrix} \sigma_i^2 & \rho_{ij} \\ \rho_{ij} & \sigma_j^2 \end{pmatrix} \right\} \qquad \text{(regional varying effects)}$$

Modeling framework: multilevel gravity modeling

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

$$\begin{aligned}\log(\lambda_{ijt}) =& \alpha + o_i + d_j + t_t + \text{dyad}_{ij} + \\ & \beta_1 \ln(\text{pop}_{it}) + \beta_2 \ln(\text{pop}_{jt}) + \gamma \ln(\text{dist}_{ijt}) + \\ & \beta_3 \ln(\text{home}_{it}) + \beta_4 \ln(\text{home}_{jt}) + \beta_5 \ln(\text{soc}_{it}) + \beta_6 \ln(\text{soc}_{jt})\end{aligned}$$

(linear model)

$$\begin{pmatrix} o_i \\ d_j \end{pmatrix} \sim \mathcal{N}\left\{ \begin{pmatrix} 0 \\ 0 \end{pmatrix}, \begin{pmatrix} \sigma_i^2 & \rho_{ij} \\ \rho_{ij} & \sigma_j^2 \end{pmatrix} \right\} \qquad \text{(regional varying effects)}$$

$$\begin{pmatrix} \text{dyad}_{ij} \\ \text{dyad}_{ji} \end{pmatrix} \sim \mathcal{N}\left\{ \begin{pmatrix} 0 \\ 0 \end{pmatrix}, \begin{pmatrix} \sigma_{\text{dyad}}^2 & \rho \\ \rho & \sigma_{\text{dyad}}^2 \end{pmatrix} \right\}$$

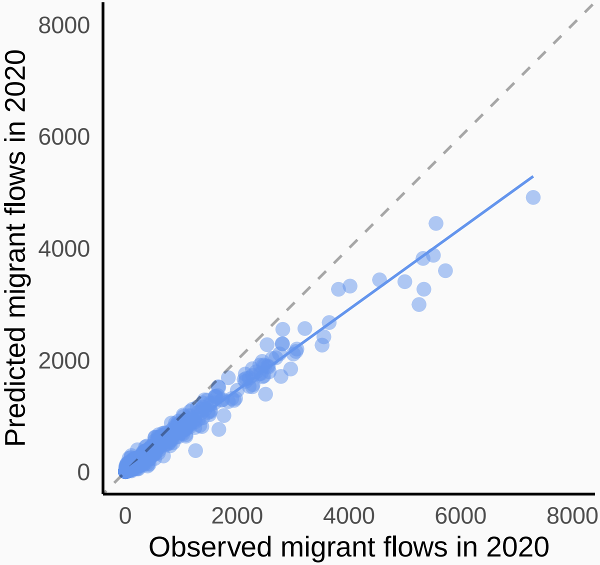
(dyad varying effects)

Main Estimation results

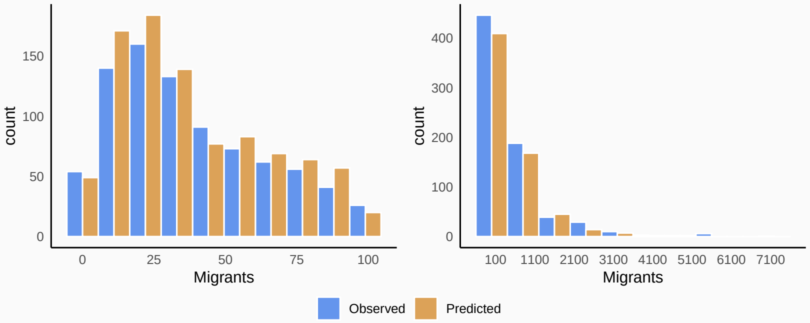
parameter	no varying effects	with varying effects
intercept	**4.48**	**4.49**
origin:		
log(population)	**0.77**	**0.32**
log(homeownership)	**−1.67**	**1.60**
log(social renting)	**−1.82**	**−0.26**
destination:		
log(population)	**0.84**	**0.55**
log(homeownership)	**-1.14**	**0.17**
log(social renting)	**-1.47**	**0.87**
migrants flow:		
log(distance)	**−1.39**	**−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 2018 ($R^2 = 0.98$)



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



Correlation patterns

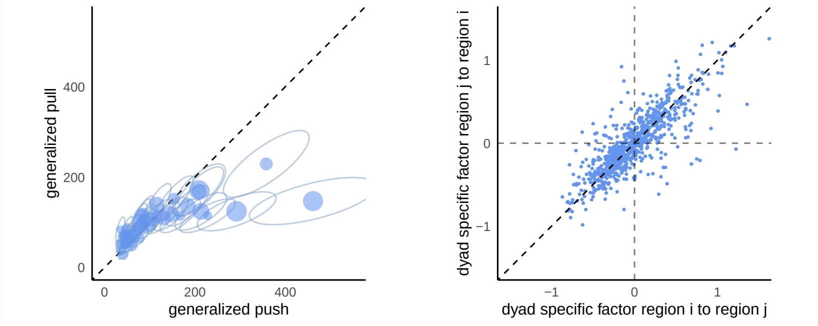
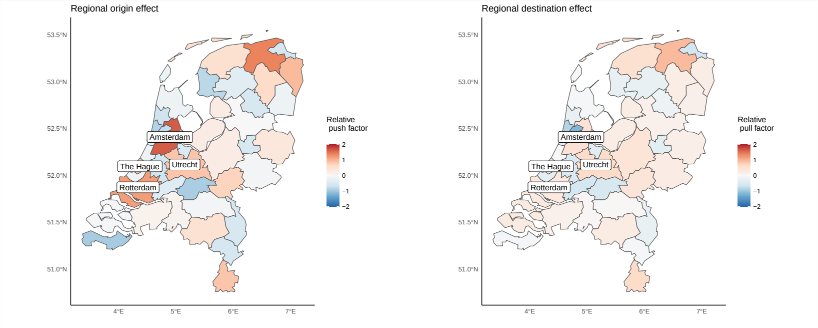
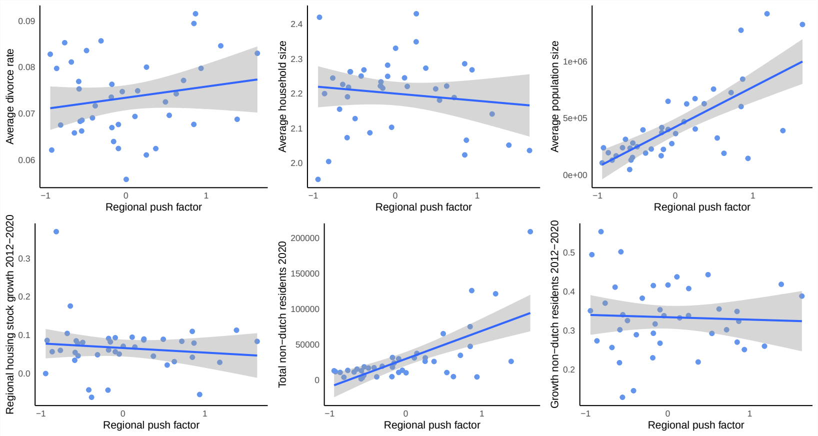


Figure 2: Correlation (0.78) between unobserved push and pull factors region (left) and flows (correlation = 0.8) within dyad pairs (right)

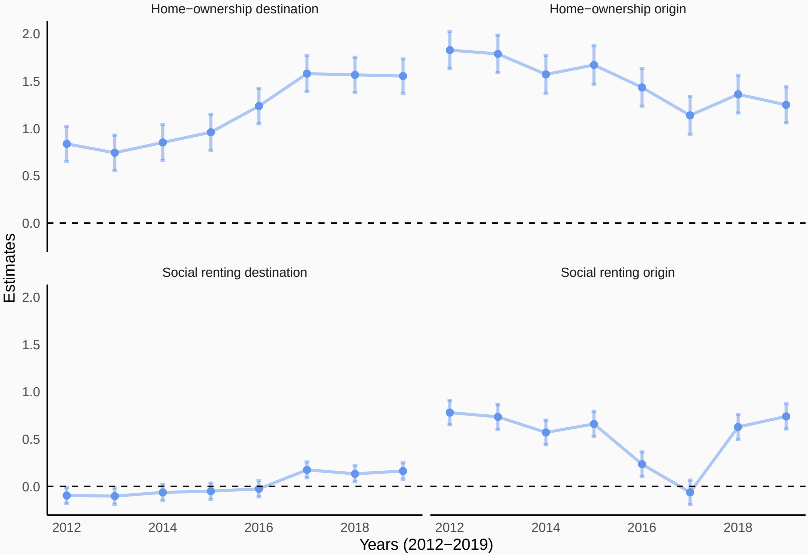
Asymmetric push and pull factors



Determinants of push factors?



Sensitivity check: temporal stability?



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

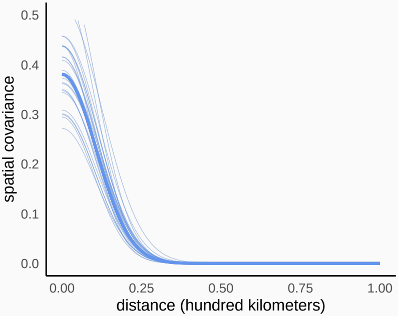
Sensitivity check: spatial autocorrelation

- 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 has a positive impact on regional domestic migration
  - social renting to a lesser extent
- large urban areas have large push effects
  - effect is different from housing market structure
  - similar to and larger than push effects in periphery

Speculation:

- home-ownership is a proxy for satelite communities close to major urban areas?
- tourism, 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.

Boyle, P (1998). "Migration and housing tenure in South East England". In: *Environment and Planning A* 30.5, pp. 855–866.

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.

References ii

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.

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.

McElreath, R. (2020). *Statistical rethinking: A Bayesian course with examples in R and Stan*. CRC press.

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

— (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.