URBAN EXODUS?

Housing market structure and interregional migration revisited
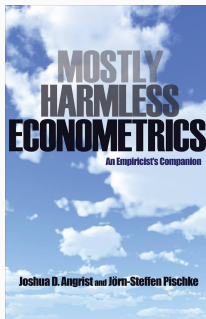
Thomas de Graaff

December 8, 2020

Vrije Universiteit Amsterdam
Tinbergen Institute Amsterdam

Background: two different cultures (Breiman, 2001)
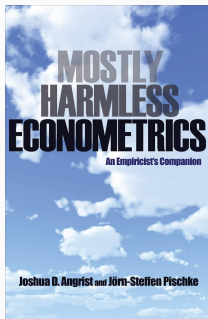
In economics:

- causal impact of x on y
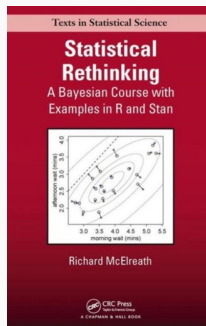- focuses on $\hat{\beta}$
- marginal effect
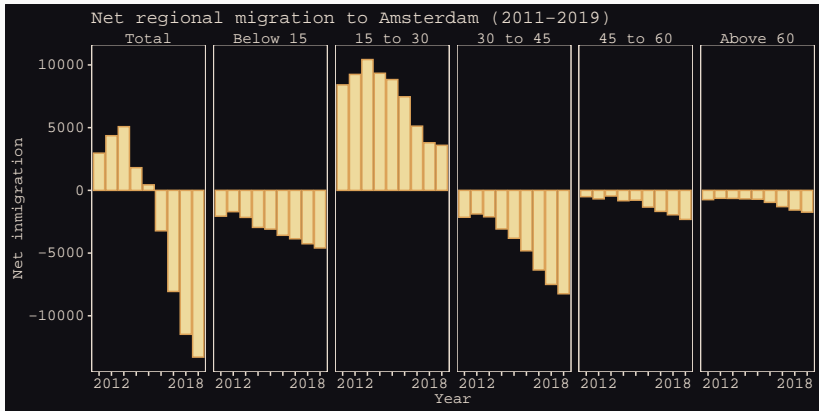
In economics:

- causal impact of x on y
- focuses on $\hat{\beta}$
- marginal effect

Outside economics:

- model performance
- focuses on $\hat{y}$
- prediction of total effect

Net regional migration to Amsterdam (2011–2019)

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., 2018)
  - 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., 2018)
  - Influx of high-skilled migrants (Beckers and Boschman, 2019)

- Dutch housing market: tight & regulated
  - large shortage of housing
  - decrease in housing transactions
  - large regional variation

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., 2018)
  - Influx of high-skilled migrants (Beckers and Boschman, 2019)

- Dutch housing market: tight & regulated
  - large shortage of housing
  - decrease in housing transactions
  - large regional variation

- 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 to predictions for the whole network

- 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
  - Social renting is an urban phenomenon (e.g. $\approx$ 40–50% 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 Solomon Kurz (2020)

Aim To predict regional migration flows using housing market structure, regional specific and regional-pair specific effects

Why a multilevel approach?

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

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 $= R^2 - R$)

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

Why a **multilevel** approach?

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?

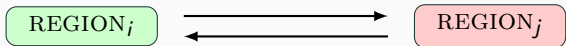There are at least two levels in migration (I use three)

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

$$\boxed{\text{REGION}_i} \quad \longrightarrow \boxed{\mathbf{x}_{ij}} \longrightarrow \quad \boxed{\text{REGION}_j}$$

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

$$\boxed{\text{REGION}_i} \qquad\qquad \boxed{\text{REGION}_j}$$
$$\uparrow \qquad\qquad\qquad \uparrow$$
$$\boxed{\mathbf{x}_i} \qquad\qquad\qquad \boxed{\mathbf{x}_j}$$

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

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

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

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
- Partial pooling: For example, origin specific effects are drawn from a distribution: $o_i \sim$ Normal(α, σ)
  - $\sigma \longrightarrow 0$: complete pooling
  - $\sigma \longrightarrow \infty$: no pooling (fixed effects)
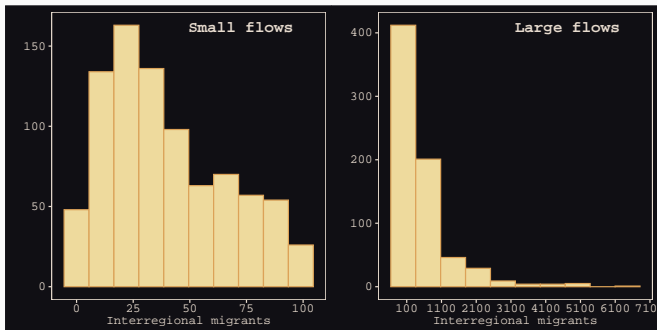
- Panel for the period 2012–2018
    - estimation: 2012–2017
    - out-of-sample prediction: 2018
- Migration flows between 40 Dutch regions
- Variance ≫ mean: over-dispersion

Data: regional housing structure in 2018
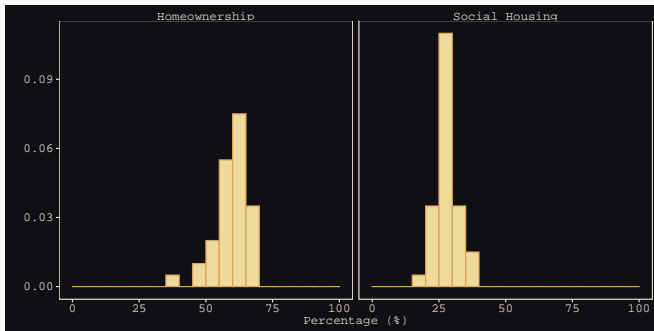
- Positive correlation between population and share social renting (0.46)
- Negative correlation between share social renting and share home-ownership (−0.88)
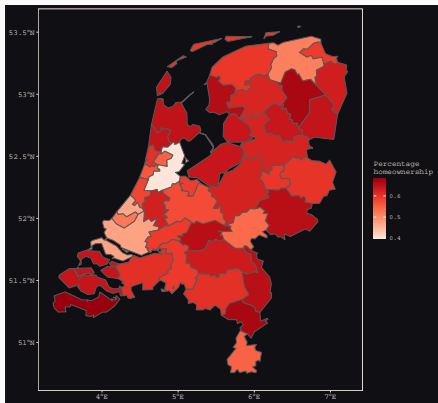
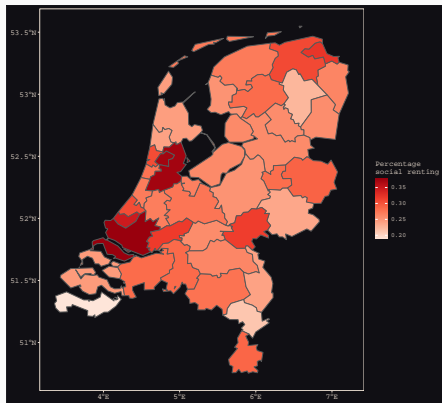**Figure 1:** Share of homeownership

Figure 2: Share of social renting

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 already 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

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

Modeling framework: multilevel gravity modeling

$$\text{Migrants}_{ij} \sim \text{Poisson}(\lambda_{ij})$$
$$\text{Migrants}_{ji} \sim \text{Poisson}(\lambda_{ji}) \qquad \text{(flow of migrants)}$$
$$\log(\lambda_{ij}) = \alpha + o_i + d_j + \text{dyad}_{ij}+ \qquad \text{(varying effects)}$$

Modeling framework: multilevel gravity modeling

$$\text{Migrants}_{ij} \sim \text{Poisson}(\lambda_{ij})$$
$$\text{Migrants}_{ji} \sim \text{Poisson}(\lambda_{ji}) \qquad \text{(flow of migrants)}$$
$$\log(\lambda_{ij}) = \alpha + o_i + d_j + \text{dyad}_{ij} + \qquad \text{(varying effects)}$$
$$\mathbf{X}_i\beta_i + \mathbf{X}_j\beta_j \qquad \text{(explanatory variables)}$$

Modeling framework: multilevel gravity modeling

$$\text{Migrants}_{ij} \sim \text{Poisson}(\lambda_{ij})$$
$$\text{Migrants}_{ji} \sim \text{Poisson}(\lambda_{ji}) \qquad \text{(flow of migrants)}$$
$$\log(\lambda_{ij}) = \alpha + o_i + d_j + \text{dyad}_{ij} + \qquad \text{(varying effects)}$$
$$\mathbf{X}_i \beta_i + \mathbf{X}_j \beta_j \qquad \text{(explanatory variables)}$$
$$\log(\lambda_{ji}) = \alpha + o_j + d_i + \text{dyad}_{ji} +$$
$$\mathbf{X}_j \beta_i + \mathbf{X}_i \beta_j$$

Modeling framework: multilevel gravity modeling

$$\text{Migrants}_{ij} \sim \text{Poisson}(\lambda_{ij})$$

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

$$\log(\lambda_{ij}) = \alpha + o_i + d_j + \text{dyad}_{ij} + \qquad \text{(varying effects)}$$

$$\mathbf{X}_i\beta_i + \mathbf{X}_j\beta_j \qquad \text{(explanatory variables)}$$

$$\log(\lambda_{ji}) = \alpha + o_j + d_i + \text{dyad}_{ji} +$$

$$\mathbf{X}_j\beta_i + \mathbf{X}_i\beta_j$$

$$\begin{pmatrix} o_i \\ d_i \end{pmatrix} \sim \text{MVNormal}\left(\begin{pmatrix} 0 \\ 0 \end{pmatrix}, \begin{pmatrix} \sigma_o^2 & \sigma_o\sigma_d\rho \\ \sigma_o\sigma_d\rho & \sigma_d^2 \end{pmatrix} \right)$$

$$\text{(varying region effects)}$$

Modeling framework: multilevel gravity modeling

$$\text{Migrants}_{ij} \sim \text{Poisson}(\lambda_{ij})$$

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

$$\log(\lambda_{ij}) = \alpha + o_i + d_j + \text{dyad}_{ij}+ \qquad \text{(varying effects)}$$

$$\mathbf{X}_i\beta_i + \mathbf{X}_j\beta_j \qquad \text{(explanatory variables)}$$

$$\log(\lambda_{ji}) = \alpha + o_j + d_i + \text{dyad}_{ji}+$$
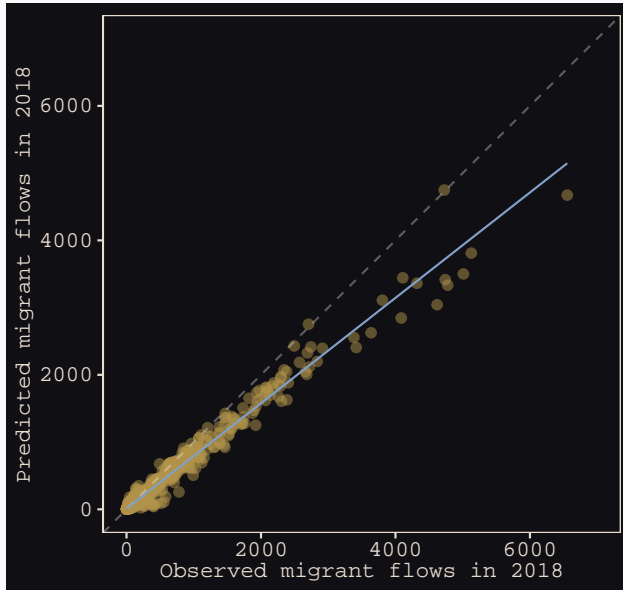
$$\mathbf{X}_j\beta_i + \mathbf{X}_i\beta_j$$

$$\begin{pmatrix} o_i \\ d_i \end{pmatrix} \sim \text{MVNormal}\left(\begin{pmatrix} 0 \\ 0 \end{pmatrix}, \begin{pmatrix} \sigma_o^2 & \sigma_o\sigma_d\rho \\ \sigma_o\sigma_d\rho & \sigma_d^2 \end{pmatrix} \right)$$

$$\qquad \text{(varying region effects)}$$

$$\begin{pmatrix} \text{dyad}_{ij} \\ \text{dyad}_{ji} \end{pmatrix} \sim \text{MVNormal}\left(\begin{pmatrix} 0 \\ 0 \end{pmatrix}, \begin{pmatrix} \sigma_{\text{dyad}}^2 & \sigma_{\text{dyad}}^2\rho_{\text{dyad}} \\ \sigma_{\text{dyad}}^2\rho_{\text{dyad}} & \sigma_{\text{dyad}}^2 \end{pmatrix} \right)$$

$$\qquad \text{(varying dyad effects)}$$

Out-of-sample prediction for 2018 ($R^2 = 0.98$)

Estimation results

| parameter	mean	standard deviation
intercept	**4.48**	0.14
origin:		
log(population)	**1.08**	0.04
log(homeownership)	**1.85**	0.11
log(social renting)	**1.39**	0.08
destination:		
log(population)	**0.95**	0.03
log(homeownership)	**-0.67**	0.10
log(social renting)	-0.09	0.07
migrants flow:		
log(distance)	**-1.62**	0.02
standard deviations:		
origin	**0.51**	0.06
destination	**0.31**	0.04
dyad	**0.39**	0.01
correlation		
origin-destination	**0.84**	0.05
dyad	**0.80**	0.01

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
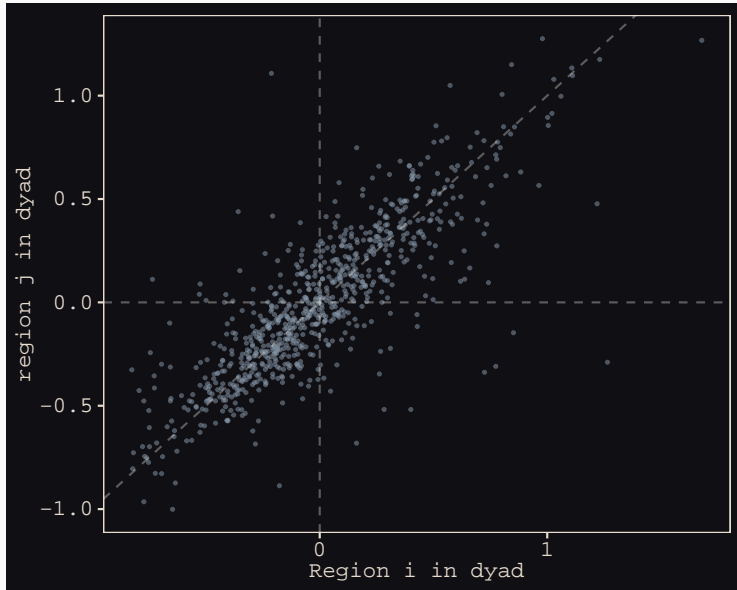
Correlation between origin and destination $\rho = 0.84$
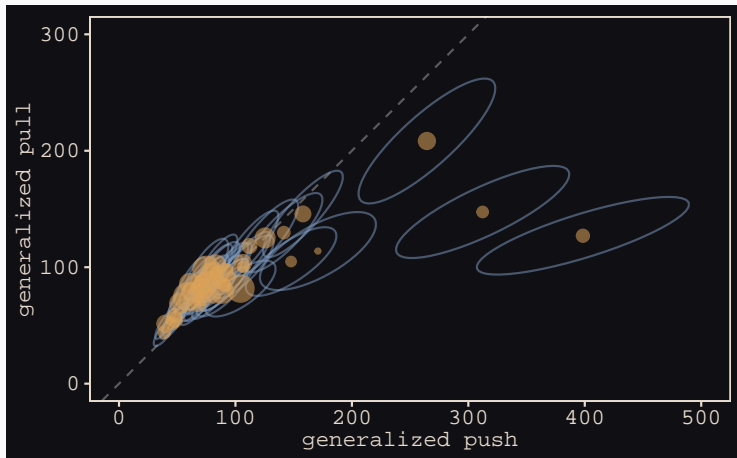
Asymmetric push and pull factors
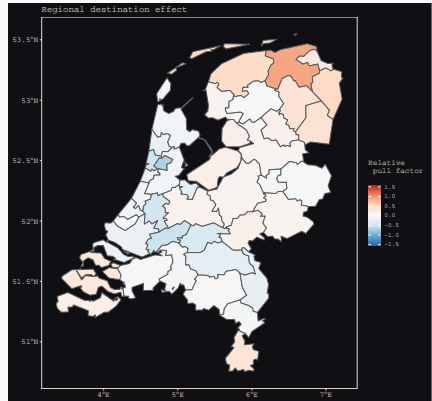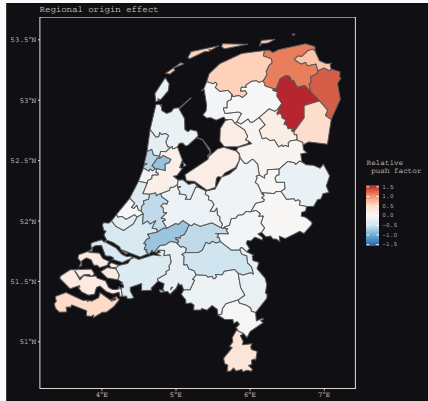
Dyad specific effects $\rho_{\mathbf{dyad}} = 0.80$

Sensitivity checks

Results are robust to

- year

Sensitivity checks

Results are robust to

- year
- interaction effects with population

Sensitivity checks

Results are robust to

- year
- interaction effects with population
- inclusion of household size

Sensitivity checks

Results are robust to

- year

- interaction effects with population

- inclusion of household size

- spatial autocorrelation in regional effects:

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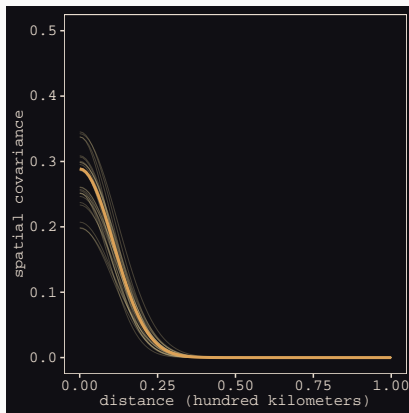\end{aligned}
$$

Sensitivity checks

Results are robust to

- year
- interaction effects with population
- inclusion of household size
- spatial autocorrelation in regional effects:

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

Modest spatial autocorrelation

Conclusions

Main results

- housing structure asymmetric impact on migration
    - positive on push/negative on pull
- No indication for specific urban exodus in period 2012–2017
    - time-trends?
- impact social renting smaller than homeowership (Boyle, 1998)
    - social housing is like a different ball game

Powerful Bayesian multilevel gravity model:

- predictive power—shrinkage
- flexibility

Supplementary materials

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

`https://github.com/Thdegraaff/migration_gravity`

Thank you!

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.

Breiman, L. et al. (2001). "Statistical modeling: The two cultures (with comments and a rejoinder by the author)". In: *Statistical science* 16.3, pp. 199–231.

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.

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., J. van Ommeren, and N. Volkhausen (2018). *Short-term rentals and the housing market: Quasi-experimental evidence from Airbnb in Los Angeles*. Tech. rep.

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 Europes 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.