NETHERLANDS INTERDISCIPLINARY DEMOGRAPHIC INSTITUTE

What do migration statistics users expect from statisticians?

Perspective from academics.

Frans Willekens, NIDI

UNECE-Eurostat Work Session on Migration Statistics,
Geneva, 29-31 October 2019

Evidence-Based Monitoring of International Migration Flows in Europe

Frans Willekens[1]

Journal of Official Statistics, Vol. 35, No. 1, 2019, pp. 231–277,

http://dx.doi.org/10.2478/JOS-2019-0011

NETHERLANDS INTERDISCIPLINARY DEMOGRAPHIC INSTITUTE

Content

1. Challenge: users want high-quality migration statistics
   "The best possible population statistics using all available data sources" (ONS)
2. Strategy:
   - Focus on relocations and their measurement
   - Simulate relocations and measurement methods
3. Conclusion

Challenge 1 Obtain 'true' migration flows

- Infer true migration flows from the limited observations on the flows.
- Observations need to be:
  - **Valid**: measure what one is supposed to measure
  - **Accurate**: observed value is close to true value (implies absence of bias)
  - **Reliable**: repeated measurements lead to same results
  - **Comparable**: permit comparison (requires that consistent concepts and methods are used)

Challenge 2. Harmonization

= Adjustment of differences and inconsistencies among different measurements, methods and procedures to make them uniform or mutually compatible

-> comparable data

= Procedures to produce comparable statistics from data that are diverse and inconsistent

-> ***comparable statistics***

Strategy

1. **Focus on the process generating relocations**
2. **Focus on different ways of measuring relocations**

NETHERLANDS INTERDISCIPLINARY DEMOGRAPHIC INSTITUTE

Strategy: simulate the relocation process

- **Relocation = change of usual residence** (address or place of residence)
  - No temporal constraint
  - No spatial constraint
- **Relocation occurs in continuous time**
  -> continuous-time process
- **All migrations are relocations but not all relocations are migrations**
  - temporal and spatial conditions
- **Continuous observation of individuals and their location is not feasible**

Observation of relocation

- **Direct observation**
  - **At time of event**
    - Method: self-reporting -> registration
  - **At later date**
    - Method: interview (census, survey)
  - **Before the event**
    - Method: interview (relocation intention)
    - With/without adjustment (eg IPS: switchers)

Observation of relocation

- **Indirect observation: comparison of locations at two points in time**
  - Prospective: panel data
  - Retrospective: current place of residence and place of residence at prior date (census, survey)
    - Interval fixed length (eg 5 years prior to census)
    - Interval variable length (eg lifetime (place of birth))

Each method has strengths and weaknesses

- **Registration**
  - timely
  - self)reporting => underreporting (under-registration), overreporting (double count) and misreporting

  -> emigration flows < immigration flows
- **Census and sample survey**
  - coverage -> representativeness
  - underreporting/nonresponse (skip a question) -> undercount

Coverage and undercount

- **Population undercount**: *presence vs observation*
  - Persons may not be counted (undercount) or may be counted twice (overcount)
- **Event undercount**: *occurrence vs observation*
  - Events (immigrations and emigrations) in a given period are not counted (undercount) or may be counted twice (overcount)

Metadata (documentation of concepts and methods of observation)

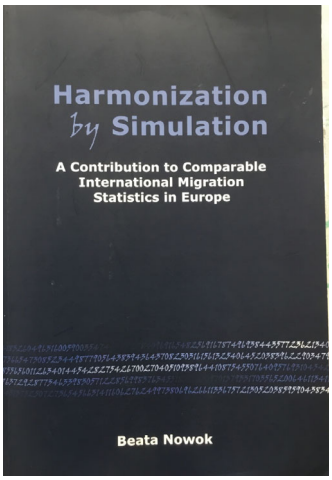
- **Producers of data (registration office; survey organisation; census bureau)**
- **Statistical offices**
- **Inventories**
  - THESIM
  - PROMINSTAT

Metadata should be complete and accurate.
and allow
- **interpretation of data**
- **quality assessment of data**
- **modelling of underlying relocation process**

Simulation

A method for the harmonization of statistics

Simulation of relocation process and its measurement



POPULATION, SPACE AND PLACE
Popul. Space Place 17, 521–533 (2011)
Published online 28 July 2010 in Wiley Online Library
(wileyonlinelibrary.com) DOI: 10.1002/ psp.624

A Probabilistic Framework for Harmonisation of Migration Statistics

Beata Nowok[1]* and Frans Willekens[1]
[1]Netherlands Interdisciplinary Demographic Institute, The Hague and Population Research Centre, Faculty of Spatial Sciences, University of Groningen, The Netherlands

Journal of the American Statistical Association
ISSN: 0162-1459 (Print) 1537-274X (Online) Journal homepage: http://amstat.tandfonline.com/loi/uasa20

Integrated Modeling of European Migration

James Raymer , Arkadiusz Wiśniowski , Jonathan J. Forster , Peter W. F. Smith & Jakub Bijak

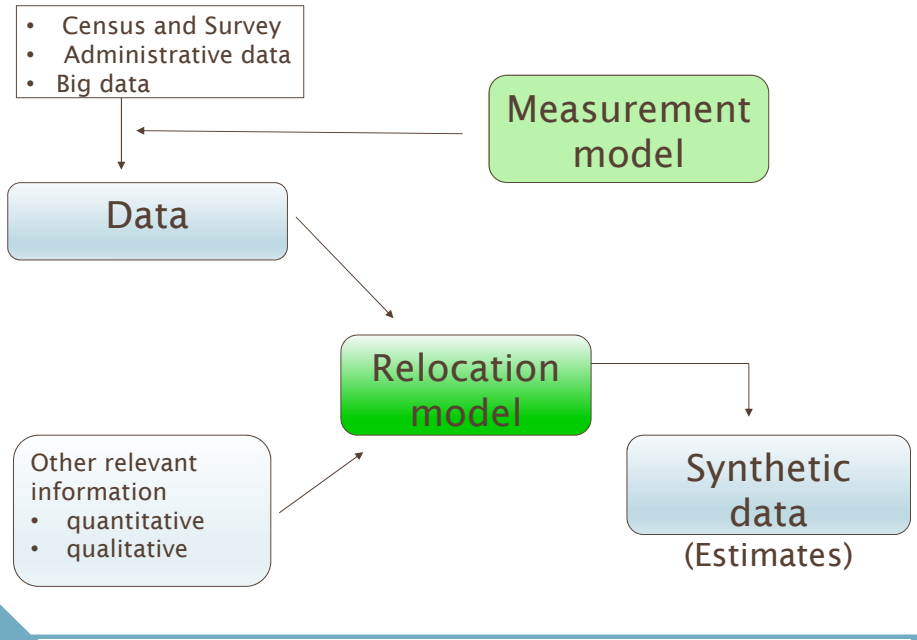
To cite this article: James Raymer , Arkadiusz Wiśniowski , Jonathan J. Forster , Peter W. F. Smith & Jakub Bijak (2013) Integrated Modeling of European Migration, Journal of the American Statistical Association, 108:503, 801-819, DOI: 10.1080/01621459.2013.789435
To link to this article: http://dx.doi.org/10.1080/01621459.2013.789435

NETHERLANDS INTERDISCIPLINARY DEMOGRAPHIC INSTITUTE

Simulation of relocation process and its measurement

- **Step 1. Describe the relocation process**
  - Relocation process is stochastic process -> stochastic process model -> **Relocation model**
- **Step 2.**
  - **Relate the parameters of the relocation model to observable quantities**
    - Account for different duration thresholds
  - **Account for uncertainties due to issues of coverage, undercount, and accuracy**

  -> **Measurement model**



Adapted from Raymer, J., Wiśniowski, A., Forster, J.J., Smith, P.W.F., and Bijak, J. (2013). Integrated modeling of European migration. *Journal of the American Statistical Association*, 108(503):801-819

The relocation process

- Stochastic process {N(t), t ≥ 0}, with N(t) the number of relocations in interval (0, t) is a counting process
- Simplest counting process: Poisson process
- **One parameter: relocation rate**

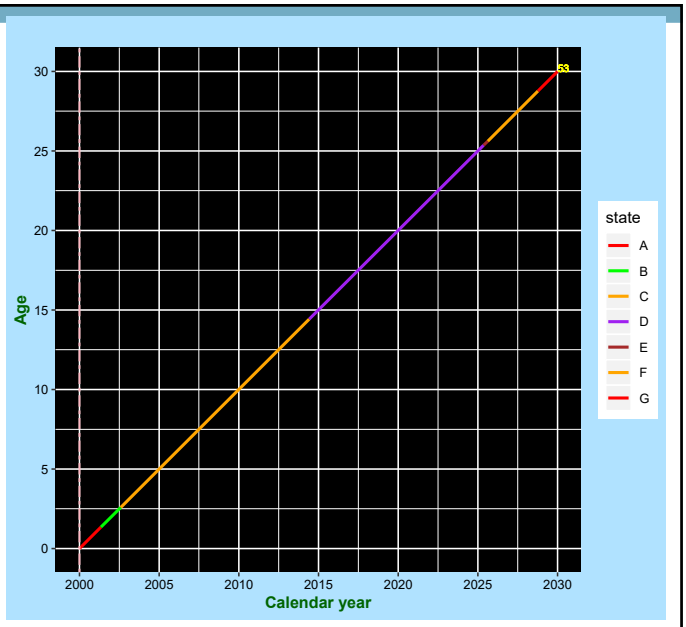
$$\Pr\{N(t) = n|\lambda\} = \frac{\lambda^n}{n!} exp[-\lambda]$$

E(N) = λ
var(N) = λ

λ = μt

- **Parameter μ may vary by migrant category and contextual variables**
- **Unobserved heterogeneity: mover-stayer model**
- **Direction of migration: Origin - destination**

Rate of relocation: 0.2



NETHERLANDS INTERDISCIPLINARY DEMOGRAPHIC INSTITUTE

Duration-of-stay criteria

Criterion	Percentage	Mig_rate
stay=0	100.00	0.200
stay=0.5	90.18	0.180
stay=1	81.43	0.162
stay=2	66.36	0.131
stay=5	35.95	0.069
stay=10	12.88	0.021

Percentage: number of relocations measured as percentage of total number of relocations

Proportion of relocations measured and relocation rate

Criterion	r_obs	r_exp	rates_obs	rates_exp
stay=0	1.0000	1.0000	0.200	0.200
stay=0.5	0.8994	0.9048	0.180	0.181
stay=1	0.8096	0.8187	0.162	0.164
stay=2	0.6552	0.6703	0.131	0.134
stay=5	0.3435	0.3678	0.069	0.074
stay=10	0.1070	0.1353	0.021	0.027

Indirect estimation of relocation: location at different points in time

		Residence					
Year	-	A	B	C	D	E	Total
0	0	10000	0	0	0	0	10000
15	0	517	1541	2172	2207	3563	10000
20	0	185	779	1426	1933	5677	10000
50	0	1	3	15	101	9880	10000

Residence	t1	t2
A	517	185
B	1541	779
C	2172	1426
D	2207	1933
E	1690	1924
F	1016	1534
G	525	1080
H	222	621
I	71	298
J	27	135
K	7	59
L	3	16
M	1	8
N	1	1
O	0	1
Sum	10000	10000

Number of persons, by location at t1 and t2

Transitions at t2 and t2-5

Residence at t1	Residence at t2									
	A	B	C	D	E	F	G	H	I	J
A	185	203	77	44	5	3	0	0	0	0
B	0	576	558	289	97	15	5	1	0	0
C	0	0	791	796	398	153	31	3	0	0
D	0	0	0	804	788	410	157	41	6	1
E	0	0	0	0	636	595	324	103	24	8
F	0	0	0	0	0	358	384	194	59	18
G	0	0	0	0	0	0	179	200	93	36
H	0	0	0	0	0	0	0	79	90	35
I	0	0	0	0	0	0	0	0	26	28
J	0	0	0	0	0	0	0	0	0	9

Transition probabilities during period (t1,t2)

Res at t1	Residence at t2								
	A	B	C	D	E	F	G	H	I
A	0.36	0.39	0.15	0.09	0.01	0.01	0.00	0.00	0.00
B	0.00	0.37	0.36	0.19	0.06	0.01	0.00	0.00	0.00
C	0.00	0.00	0.36	0.37	0.18	0.07	0.01	0.00	0.00
D	0.00	0.00	0.00	0.36	0.36	0.19	0.07	0.02	0.00
E	0.00	0.00	0.00	0.00	0.38	0.35	0.19	0.06	0.01
F	0.00	0.00	0.00	0.00	0.00	0.35	0.38	0.19	0.06
G	0.00	0.00	0.00	0.00	0.00	0.00	0.34	0.38	0.18
H	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.36	0.41
I	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.37

≠ probability of moving from i to j during period (t1,t2)
Multiple moves during period possible

Transitions during period (t1,t2)

- **Underlying process: Poisson process with relocation rate 0.2**
- **Residence at t1 is A**
- **Residence at t2 (t1+5): depends on number of relocations**
- **Distribution: theoretical vs sample**

	A	B	C	D	E	F
Poisson	0.368	0.368	0.184	0.061	0.015	0.003
Simulation	0.371	0.360	0.188	0.063	0.015	0.003

Add country of birth

- Migrations by origin, destination AND country of birth
- Abel (2018) "Estimation of global bilateral migration flows by gender between 1960 and 2015" International Migration Review, https://doi.org/10.1111/imre.12327

- **Assumption: underlying Poisson process**
- **Abel uses log-linear models, which are Poisson regression models**

Bilateral stock data:

		Place of residence (t)				
		A	B	C	D	Sum
Birthplace	A	100	10	10	0	120
	B	20	55	25	10	110
	C	10	40	140	65	255
	D	20	25	20	200	265
	Sum	150	130	195	275	750

		Place of residence (t + 1)				
		A	B	C	D	Sum
Birthplace	A	70	30	10	10	120
	B	25	60	10	15	110
	C	10	55	140	50	255
	D	40	45	0	180	265
	Sum	145	190	160	255	750

Estimates of origin–destination–place of birth flow tables:

Birthplace = A

		Destination				
		A	B	C	D	Sum
Origin	A	*70*	20	0	10	**100**
	B	0	*10*	0	0	**10**
	C	0	0	*10*	0	**10**
	D	0	0	0	*0*	**0**
	Sum	**70**	**30**	**10**	**10**	**120**

Birthplace = B

		Destination				
		A	B	C	D	Sum
Origin	A	*20*	0	0	0	**20**
	B	0	*55*	0	0	**55**
	C	5	5	*10*	5	**25**
	D	0	0	0	*10*	**10**
	Sum	**25**	**60**	**10**	**15**	**110**

Conclusions

- **Large diversity in how migration is measured -> statistics not comparable**
- **Each method of observation is a different way of looking at the same underlying process: the relocation process**
- **The relocation process is a stochastic process, eg Poisson process**
- **The stochastic process generates individual relocation/migration histories (*statistical individuals)***

NETHERLANDS INTERDISCIPLINARY DEMOGRAPHIC INSTITUTE

Conclusions

- **Simulated relocation histories**
  - = allows the estimation of all migrations, including circular migration
  - = tool to make migration data comparable across observation plans: *Harmonization by simulation*
- **The approach requires data and metadata**
  - Meta data: documentation of migration concepts used, methods of data collection and quality assessment

thank you

Willlekens@nidi.nl

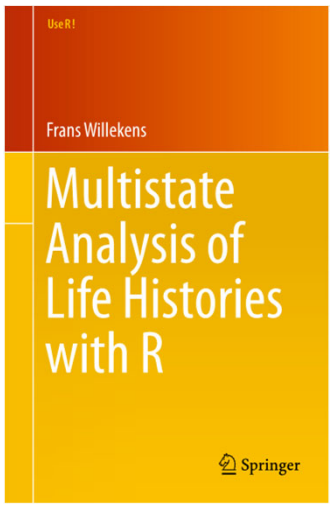
NIDI is an institute of the Royal Netherlands Academy of Arts and Sciences KNAW and is affiliated to the University of Groningen

KONINKLIJKE NEDERLANDSE AKADEMIE VAN WETENSCHAPPEN

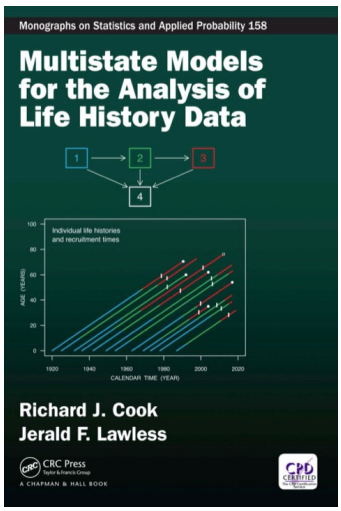
university of groningen

www.nidi.nl

NETHERLANDS INTERDISCIPLINARY DEMOGRAPHIC INSTITUTE



2014



2018

NETHERLANDS INTERDISCIPLINARY DEMOGRAPHIC INSTITUTE