POST-PRINT

Molina, J. L., Lubbers, M. J., Hâncean, M.-G., & Fradejas-García, I. (2022). Short Take: Sampling from Transnational Social Fields. **Field Methods**, 34(3), 256-264. https://doi.org/10.1177/1525822X221105920

Short Take: Sampling from Transnational Social Fields

José Luis Molina*[a], Miranda J Lubbers[a], Marian-Gabriel Hâncean[b] and Ignacio Fradejas-García[a]

[a] GRAFO-Department of Social and Cultural Anthropology, Universitat Autònoma de Barcelona.

[b] Department of Sociology, University of Bucharest.

*Corresponding author (joseluis.molina@uab.cat)

Abstract

Thanks to the latest developments in network-oriented sampling, it is now possible to measure "transnational social fields", or emergent social structures that connect places or regions in different countries. These structures are instrumental in explaining socio-cultural phenomena like the emergence of ethnic or demographic enclaves, social and economic remittances, and ethnic identifications. Nevertheless, they have only been mentioned metaphorically so far.

Introduction

International migration flows are not evenly distributed but polarized across a complex web of "migrant corridors" (Carling and Jolivet 2016) connecting sending and destination countries, and within them, sending and destination places or regions. While international and national statistical agencies provide periodic information about migration flows at the country level, there is no systematic account of such flows at the place or region levels, although some progress has recently been made using digital traces of mobility from mobile phone records and social media (Mazzoli et al. 2020).

Among the factors explaining such clustering is "chain migration " (MacDonald and MacDonald 1964), which states that the costs of (international) migration decrease over time because former migrants usually provide information and support to new ones, which helps them successfully overcome the pitfalls of the process. This network mechanism explains emergent local phenomena like ethnic or demographic "enclaves" (Molina et al. 2018) and "transnational social fields" (TSFs), defined as "interlocking egocentric networks that extend across the borders of two or more nation-states and that incorporate its participants in the day-to-day activities of social reproduction in these various locations"

(Fouron and Glick Schiller 2001). TSFs include not only migrants but also "persons born in the country of origin who never migrated, and persons born in the country of settlement of many different ethnic backgrounds" (Glick Schiller and Fouron 1999: 544). Thus, TSFs are defined on the basis of migrants who move between places of origin and destination, and also include their families and acquaintances in both places, regardless of their nationality (Lubbers, Verdery, and Molina 2020). The unbounded nature of TSFs complicates sampling from them and, thus, a better understanding of migrant adaptation, ethnic entrepreneurship, social and economic remittances, and ethnic identity (Levitt 1998; Vertovec 2009).

The issue is then: how to measure these TSFs (Molina, Petermann, and Herz 2015)? Drawing on the growing literature about this issue and the lessons learned during the research project ORBITS[1], which aimed to measure two TSFs connecting places in Romania and Spain, this paper explains a step-by-step method to sample from such structures. We first briefly review the literature on network-oriented sampling methods and their application to migration (see for further information about network approaches to operationalizing TSFs Lubbers, Verdery & Molina 2020).

Literature review

Network-oriented sampling methods were developed for sampling hidden or hard-to-reach populations, such as populations at high risk for HIV and AIDS (Bernard et al. 2010). These methods use chains of referrals or link-tracing designs, i.e., where participants recruit other participants.

Snowball sampling

In snowball sampling, a first set of non-randomly selected individuals of a subpopulation (the "seeds") is interviewed and asked to nominate other potential respondents of the same subpopulation. These persons are then interviewed, and asked to nominate other potential respondents. The process is repeated until the desired sample size is reached.

Respondent driven sampling (RDS)

Respondent-driven sampling (RDS) aims to achieve a quasi-probability sample overcoming the statistical biases inherent in snowball sampling (e.g. self-selection of highly cooperative participants, homophily in recruitment processes, oversampling of individuals with larger networks). To do so, and if various assumptions are met (Heckathorn 1997, 2002; Salganik & Heckathorn, 2004), researchers carefully select a small number of seeds who represent key subpopulations. Each "seed" recruits a fixed, small number of other members of the population, and researchers keep track of the chains of references to

[1] https://pagines.uab.cat/orbits/en [visited: 6-11-2020]. See Hâncean, Lubbers, & Molina, 2020 for an overview.

control for the dependence structure during analyses. At least four stages should be completed to ensure a good penetration in the target population. Often, a dual-rewarding system is employed: participants are rewarded separately for their participation and for successfully recruiting referees.

Bi-national link tracing

RDS has not been developed for transnational populations. Mouw et al. (Mouw et al. 2014; Mouw and Verdery 2012) developed a method for studying TSFs connecting migrants from Guanajuato (Mexico) with two sites of the USA (North Carolina and Denver). After conducting exploratory ethnographic fieldwork, the research started from a small convenience sample in the two USA sites (5 and 12 people, respectively, all from Guanajuato). Researchers asked these "seeds" to recruit others in both destination and origin areas. The survey inquired about personal networks, with name generators eliciting a list of relatives and friends in both places, along with nominees' sociodemographic characteristics. Once the desired sample size was reached in the destination places, the team moved to Guanajuato, where 20 seeds were randomly selected from the pool of alters elicited previously. There, the team conducted three additional waves. Although no alter-alter information was collected, the cross-nominations between respondents render a rich transnational social network.

Adapted binational link-tracing: The Orbits research project

We have adapted the binational link-tracing design to measure TSFs of Romanians living in Spain and their local and transnational relationships. The following adaptations were made. First, instead of following two stages, fieldwork was conducted simultaneously in the two places, enabling a timely follow-up of non-migrants. Second, non-migrants were also asked to nominate people in both places, leading to new referrals in the receiving country and a wider diversity of non-migrants' exposure to migrants. Third, we added a name generator enquiring about people living in other places in Spain, Romania, or elsewhere, to not artificially limit transnationalism bilocally. Fourth, we measured alter-alter relationships for a random selection of alters per respondent to estimate network structure more adequately. Our method consisted of the following steps.

Step-by-step instructions

1. *Select the target population in the destination place or region*

First, a place is selected where the target population, migrants in this case, is disproportionately present in a place or region compared with other nationalities. The presence of "enclaves" may indicate the existence of one or several TSFs feeding them. "Ethnic" neighborhoods, touristic villages with one

minority occupying an occupational niche (Valenzuela García et al. 2014), areas next to (agro)-industrial districts (Fradejas-García, Molina, and Lubbers 2021) are natural candidates, but also the often-invisible collectivities of domestic care workers. Place-based official statistics on migration need to be examined with a high level of detail. Next, the sample size in each place needs to be defined depending on the size of the target population and the resources of the research project (in our case, we fixed the sample size at 150 individuals in each end, N=300 per surveyed TSF).

2. *Gain access to the field (institutions, associations, local experts, social media)*

Access to the field is gained by introducing the research project to local institutions, migrant associations of the targeted population, and religious leaders, among other potential key actors. It is also advisable to contact (local) researchers with expertise in the area or target population, and, if needed, achieve their endorsement. A public presentation of the project with media coverage can augment its visibility, inviting the relevant actors and stakeholders. In the field site in the sending country, similar efforts should be made to give the project visibility and public support. *Important*: In some populations, you may expect reactions of mistrust from prospective participants, especially on social media. Prepare for a quick reaction to questions and information requests.

3. *Conduct ethnographic fieldwork*

Participant observation provides unique insights about the population of interest, its internal diversity, the social contexts it attends, and the community structure, and helps build trustful relationships with individuals and groups. This knowledge is crucial for selecting and building rapport with appropriate "seeds" that can potentially launch chains of contacts in each identified subgroup, and for drafting and pilot-testing the survey and detecting misunderstandings (see Step 4). Similar exploratory research should be conducted in the field site in the sending country. *Important*: as gatekeepers who give access to part of the field can limit access to others, developing independent relationships during ethnographic fieldwork at each end of the TSF is key.

4. *Refine the research design: fieldwork phases, incentives, survey mode, questionnaire*

With the information gathered in Step 3, the survey mode (face-to-face, online, or mixed) must be chosen. The researchers should decide whether to conduct fieldwork in the two sites sequentially or simultaneously. Simultaneous fieldwork is logistically complex as two teams work in parallel, but it has the advantage that referrals can be followed up timely, increasing referrals' likelihood of participation. Furthermore, a suitable reward system needs to be established, given the population, interview duration, type of information collected, and the available resources. The questionnaire needs to be

designed, translated, and implemented in survey software for CAPI administration. Questionnaires may differ depending on respondents' status (e.g., migrant, non-migrant, or return migrant). An important issue is defining the number and wording of the name generators and referral questions. To understand how TSFs work, the name generators need to capture contacts in different places. Following Mouw et al. (2014), we specifically asked respondents to nominate people who reside in the community or origin, people who reside in migrants' community of residence, and return migrants (see supplemental materials). We added a name generator that enquired about contacts who live outside these places, for a more complete view on TSFs. We used a fixed maximum number of nominations for each name generator (see Supplemental Material) which capped the maximum total number of nominations at 40. We also asked respondents to refer to three potential Romanian respondents in the place of origin and three in migrants' community of residence. As multiple respondents may nominate or refer to the same person in the TSF, a robust system needs to be developed for uniquely identifying anonymized individuals, i.e., without collecting their personal data (e.g., following Mouw's recommendation, we used a centralized system that assigns to each participant, nominee, and referral the three first letters of both the first name and surname and the first four digits of the mobile phone number). Pilot-tests should be conducted in both places. *Recommendation*s: try to keep the survey short to avoid high refusal rates for participation or referral. In case of face-to-face interviewing, participants need to be asked to contact their referees during the interview by phone, social media or instant messaging tools to ask them for their consent with the researchers contacting them.

5. *Start the fieldwork. Prepare for a long fieldwork phase, fraught with pitfalls: team, coordination, and data management*

Fieldwork, especially in pandemic times, is a formidable challenge. Even after a thorough preparation, accomplishing the previous phases satisfactorily, you need to prepare for a long fieldwork phase. If fieldwork is conducted in two sites simultaneously, two fieldwork teams need to be formed. They are preferably mixed in gender (if the population is mixed-gender), and should be properly trained and supervised. Contextual fieldnotes, receipts of compensations, consent forms, and the timely uploading of data and referrals' contact information, are crucial for successful fieldwork (see Fieldwork Data Management section in the Supplemental material). During the fieldwork stage, reference chains should be monitored, to control the level of homophily in the referrals within each chain (e.g., in gender, race, religion, or occupational status) and to control that the different chains eventually reach each other, indicating saturation. If homophily is high in different chains, researchers may decide to add a new seed

that represents a part of the population not represented in the chains. Similarly, if a chain does not lead to any new contacts after only a few steps, a new seed may be added.

Warning: The large number of referees and network members, begets problems of unique identification that should be anticipated and managed properly.

6. *Prepare the data files*

First, identities need to be controlled to ensure that each individual has a single, unique ID. Problems with identities occur, for example, when individuals do not know the phone number or last name of nominees, or when two individuals have been erroneously given the same ID (especially likely when some information of the name or phone number is missing) or even or the same individual is erroneously given different IDs. In case of doubt, alter attributes (e.g., gender, place of living, occupation) can be used as control variables. Second, based on the referral chains, RDS-style survey weights need to be assigned . Third, the data need to be curated and anonymized, as in any survey. The supplemental material shows the organization of the data files. *Important*: network-oriented methods are especially sensitive to selection effects like gender homophily. Exploring the referral chains is therefore essential in this phase.

7. *Provide feedback and plan transfer activities with stakeholders*

Finally, it is important to share the main results with stakeholders as soon as possible, typically a few months after the project's end. These meetings provide new information and ideas that can be incorporated in further dissemination activities and academic publications.

Conclusion

Sampling TSFs is a hard but feasible task that makes hidden structures connecting places transnationally visible and that can provide local authorities with valuable information for developing coordinated policies (see, for instance, the video[2] of the twinning agreement between the cities of Castelló, Spain, and Târgoviște, Romania).

Acknowledgments

We want to thank Iulian Oană, Bianca-Elena Mihăilă, Adelina Alexandra Stoica, Adriana Suiu, Mihai Busuioc, Renáta Hosnedlova, Jürgen Lerner, Daniel Holgado, Isidro Maya-Jariego, Ashton Verdery, Ted Mouw and Giovanna Merli for their invaluable contribution to the design and data collection phases of the research project. The authors acknowledge a grant of the Ministry of Economy, Industry and

[2] https://www.youtube.com/watch?v=r-HdqYCug1A [Accessed 10-11-2020].

Competitiveness (MINECO-FEDER; CSO2015- 68687-P, Spain), and of The Executive Agency for Higher Education, Research, Development and Innovation Funding (PN-III-P4-ID-PCE-2020-2828, Romania).

Declaration of Conflicting Interests

The authors declared no potential conflicts of interest with respect to the research, authorship, and/or publication of this article.

Supplemental Material

Supplemental material for this article is available online.

References

Bajardi, Paolo, Matteo Delfino, André Panisson, Giovanni Petri, and Michele Tizzoni. 2015. "Unveiling Patterns of International Communities in a Global City Using Mobile Phone Data." *EPJ Data Science* 4(1):3.

Bernard, H. Russell, Tim Hallett, Alexandrina Iovita, Eugene C. Johnsen, Rob Lyerla, Christopher McCarty, Mary Mahy, Matthew J. Salganik, Tetiana Saliuk, Otilia Scutelniciuc, Gene A. Shelley, Petchsri Sirinirund, Sharon Weir, and Donna F. Stroup. 2010. "Counting Hard-to-Count Populations: The Network Scale-up Method for Public Health." *Sexually Transmitted Infections* 86 Suppl 2:ii11-i15.

Carling, Jørgen and Dominique Jolivet. 2016. "Exploring 12 Migration Corridors: Rationale, Methodology and Overview." Pp. 18–46 in *Beyond Networks. Feedback in international migration*, edited by O. Bakewell, G. Engbersen, M. L. Fonseca, and C. Horst. Basingstoke, Hampshire: Palgrave-Macmillan.

Fouron, Georges and Nina Glick Schiller. 2001. "All in the Family: Gender, Transnational Migration, and the Nation-State." *Identities* 7(4):539–82.

Fradejas-García, Ignacio, José Luis Molina, and Miranda J. Lubbers. 2021. "Migrant Entrepreneurs in the 'Farm of Europe': The Role of Transnational Structures." Under review.

Glick Schiller, Nina and Georges E. Fouron. 1999. "Terrains of Blood and Nation: Haitian Transnational Social Fields." *Ethnic and Racial Studies* 22(2):340–66.

Hâncean, Marian-Gabriel, Miranda J. Lubbers, and José Luis Molina. 2020. "Measuring Transnational Social Fields Through Binational Link-Tracing Sampling." Under review.

Heckathorn, Douglas D. 1997. "Respondent-Driven Sampling: A New Approach to the Study of Hidden Populations." *Social Problems* 44(2):174–99.

Heckathorn, Douglas D. 2002. "Respondent-Driven Sampling II: Deriving Valid Population Estimates from Chain-Referral Samples of Hidden Populations." *Social Problems* 49(1):11–34.

Levitt, Peggy. 1998. "Social Remittances: Migration Driven Local-Level Forms of Cultural Diffusion." *International Migration Review* 32(4):926–48.

Lubbers, Miranda J., Ashton M. Verdery, and José Luis Molina. 2020. "Social Networks and Transnational Social Fields: A Review of Quantitative and Mixed-Methods Approaches." *International Migration Review* 54(1): 177-204.

MacDonald, John S. and Leatrice D. MacDonald. 1964. "Chain Migration, Ethnic Neighborhood Formation and Social Networks." *The Milbank Memorial Fund Quarterly* 42(1):82.

Mazzoli, Mattia, Boris Diechtiareff, Antònia Tugores, Willian Wives, Natalia Adler, Pere Colet, and José J. Ramasco. 2020. "Migrant Mobility Flows Characterized with Digital Data." *PLoS One* 15(3):e0230264.

Molina, José Luis, Luis Martínez-Cháfer, Francesc Xavier Molina-Morales, Miranda J. Lubbers. 2018.

"Industrial Districts and Migrant Enclaves: A Model of Interaction." *European Planning Studies* 26(6): 1160-1180.

Molina, José Luis, Sören Petermann, and Andreas Herz. 2015. "Defining and Measuring Transnational Social Structures." *Field Methods* 27(3):223–43.

Mouw, Ted, Sergio Chavez, Heather Edelblute, and Ashton Verdery. 2014. "Binational Social Networks and Assimilation: A Test of the Importance of Transnationalism." *Social Problems* 61(3):329–59.

Mouw, Ted and Ashton M. Verdery. 2012. "Network Sampling with Memory: A Proposal for More Efficient Sampling from Social Networks." *Sociological Methodology* 42(1):206–56.

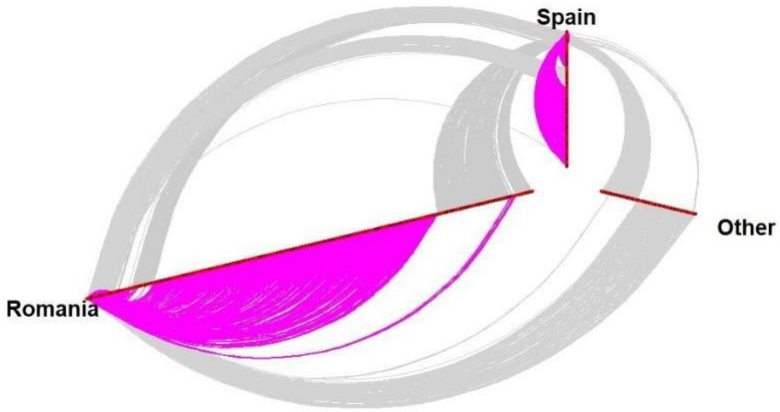
Salganik, Matthew J. and Douglas D. Heckathorn. 2004. "Sampling and Estimation in Hidden Populations Using Respondent-Driven Sampling." *Sociological Methodology* 34:193–239.

Valenzuela García, Hugo, José Luis Molina, Alejandro García Macías, Miranda J. Lubbers, Judith Pampalona, and Jürgen Lerner. 2014. "On Heterogeneous and Homogenous Networks in a Multilayered Reality. Clashing Interests in the Ethnic Enclave of Lloret de Mar." *Societies* 4:85–104.

Vertovec, Steven. 2009. *Transnationalism*. New York: Routledge.

Supplemental Material

Hiveplot of the Castelló (Spain) – Targoviste (Romania) transnational social field



Legend: this hiveplot illustrates how the 4,855 nodes and the 5,477 ties collected are partitioned on residence (countries where people live). Magenta ties indicate within-country social connections (i.e. 1,524 ties connect people living in Spain, and 2,237, people living in Romania), and gray ties between-country connections (i.e., 1,133 ties connect Spain and Romania, 223 ties connect Romanians living in Spain to Romanians living in other countries, and 360 ties connect Romanians living in Romania to Romanians living in other countries). The hiveplot was built using the hiveR package.

Name generators

The questionnaire, which was adapted from Mouw et al. (2014), had three versions for (1) Romanian migrants in Spain, (2) non-migrants, and (3) returned migrants. Name generators were common to all versions but the place names and countries were changed accordingly. The interview lasted about 1 hour on average.

The name generators were introduced as follows "I am now going to ask you basic information about your social connections. This information will allow us to observe how people remain connected to their communities. I am going to ask you for the name (acronym of the first three letters of the first name and the first three letters of the surname), the last four digits of the phone number, occupation and a few other characteristics to avoid confusing persons who have similar names. Let's start with friends and acquaintances".

The name generator was then as formulated as follows, for example for migrants, "Could you tell me which [friends and acquaintances] over the age of 18 live in [current place of residency]? They do not have to be very close. They can be Romanians, Spaniards, or of another origin. You can include any

person whom you know. However, if you could mention more than 10 people, please start with the people you feel closest with." Friends and acquaintances in the current place of residency formed the first name generator of five; the variations are listed in the table below (instead of the descriptions, the actual place names were given).

	Name generators	Max. number of friends and acquaintances	Max. number of relatives	Number (maximum)
1	Friends and acquaintances living in the respondent's current place of residency (max. 10)	10		10
2	Relatives living in the current place of residency (max. 5)		5	5
3	Relatives, friends and acquaintances (max. 5) who lived in the fieldsite in Spain but have returned to Romania	5		5
4	Relatives (max. 5) and friends or acquaintances (max. 5) who live in the other end of the migrant corridor	5	5	10
5	Relatives (max. 5) and friends and acquaintances (max. 5) living outside the migrant corridor (anywhere)	5	5	10
	Total	20-25	15-20	40

For each alter, the *name interpreters* collected were sex, occupation, religion, closeness, and frequency of interaction. Furthermore, for family members the precise relationship was collected and for friends and acquaintances the duration of the relationship. Next, from each name generator we systematically selected one or two alters (9 in total), and asked for each of the 36 alter-alter pairs, whether these persons know each other and would contact each other independently of them (response categories yes, no, don´t know).. Finally, respondents were asked to nominate three referrals in one end of the migrant corridor and three referrals on the other for further interviewing, either repeating someone from the name generators or new names. All referrals collected were invited to participate in the survey.

Fieldwork Data Management

All data collected during fieldwork was uploaded to a secure server as soon as possible: photos, videos, ethnographic reports, meeting reports, questionnaires, brief fieldnotes, interview summaries, and consent forms.

Questionnaire data were split into three Excel files (ego responses including ego-alter ties, alter-alter ties reported by ego, and ego referrals) whereas fieldnotes were recorded in text files.

The structure of the ego responses file is as follows (the ego-IDs are fictive).

EgoName	RESP_TYPE	…	AlterName	AlterCountry	Alter/Ref_Sex	…

MIHILI6221	MS (Migrant Spain)	IULVAS9788	SPAIN	2
MIHILI6221	MS	CRIVAS5793	SPAIN	1
…	MS	AMPASE0755	SPAIN	2
IONGRI6099	MS	MANGAB3653	SPAIN	1
IONGRI6099	MS	ALEION1463	SPAIN	2
…	MS	ALEAND9905	SPAIN	2

The alter-alter ties were coded as follows:

EgoID	Alter 1 ID	Alter 2 ID	Relationship
IFGCAS002MS	IULVAS9788	CRIVAS5793	1
IFGCAS002MS	IULVAS9788	AMPASE0755	0
…	CRIVAS5793	AMPASE0755	99
IONILI6099	MANGAB3653	ALEION1463	0
IONILI6099	MANGAB3653	ALEAND9905	0
…	…	…	…

The referral database structure is as follows.

ID Referee (e.g., JLMCAS001MS)	First name of the referee (first 3 letters)	Surname of referee (first 3 letters)	Phone of referee (last 4 digits)	Date of interview of referee YY/MM/DD	ID Referrals	Full phone number / Facebook /WhatsApp	Place of residence of referral
IFGCAS002MS	ION	ILI	6099	10/11/17	ANRGHE2145	12345678	CASTELLÓ
IFGCAS002MS	ION	ILI	6099	10/11/17	IOIBUT8969	…	TARGOVISTE
IFGCAS002MS	ION	ILI	6099	10/11/17	LUBION3700	…	TARGOVISTE
IFGCAS003MS	DRA	DUM	5007	12/11/17	VALRIS2044	…	CASTELLÓ
IFGCAS003MS	DRA	DUM	5007	12/11/17	MIHCON7927	…	CASTELLÓ

(cont.)

Was referral contacted by text/voice by the referee? Yes/No	Has referral agreed to participate? Yes/No/other situation	Was referral contacted by text/voice by an interviewer? If yes: by whom?	Was referral interviewed? If so, ID of the referral	Wave
Yes	Yes	ADS	ADSCAS001MS	1
Yes	Yes	BEM	MGHDAM001NM	1
Yes	Yes	BEM	MGHDAM002NM	1
Yes	Yes	ADS	ADSCAS002MS	1
Yes	Yes	ADS	ADSCAS003MS	1

Finally, an instance of the header of interviews' brief fieldnotes is as follows.

Field note: *#26*
Interview code: *026AASDAM008NM*
Date: *01/02/2018*
Total time: *Meeting duration 01:04:00. Duration of the presentation of the project 00:03:00. From 14:43:00 until 14:46:00. The interview started at 14:46:00 and ended at 15:47:00.*
ID Referee: *AASDAM007RM*
ID Referral: *IULFUR8867*
Field note by *AAS*