

Exploring Well-Being and Social Integration in the Context of Liquid Migration

Survey methodology

Māris Goldmanis & Inta Mieriņa

2021 (May)

The quantitative data in the project were collected in a large-scale web survey conducted between 24 September and 11 November 2019.

For recruiting respondents, a large variety of recruitment channels and sources were used – both general and diaspora-specific, – accompanied by a media campaign. Information about the survey was disseminated through the diaspora media (e.g., baltic-ireland.ie, latviansonline.com), placing information in Internet groups and newsgroups of Latvian citizens living abroad (16 groups draugiem.lv, three LinkedIn groups and more than 150 Facebook groups), with the help of Latvian institutions working with the diaspora (Ministry of Foreign Affairs, Ministry of Environment and Regional Development, Ministry of Education and Science, Latvian Language Agency, Latvian Investment and Development Agency), as well as sending out information and asking to share it 390 previously identified Latvian diaspora organizations, groups, societies, congregations, etc. (including the European Latvian Association and the World Association of Free Latvians). Informative banners about the survey were placed on Facebook, on the most popular portal in Latvia DELFI, and press.lv, draugiem.lv (Latvian social media portal) invitations to its users, as well as active communication on the websites and social networks of the researchers (including migracija.lv, fsi.lu.lv, and other), as well as sending of a press release. Special attention was paid to reaching Russian-speaking Latvian emigrants, taking into account that in previous studies the response rate of this group was relatively low. We also took advantage of the large database of emigrants' e-mail addresses (n≈6000) collected in previous projects that applied a similar methodology, involving respondents who had agreed to take part in future studies (Mieriņa, 2019). Of all 6242 emigrants who took part in the 2019 survey, 1073 were panel participants.

The questionnaire was available in Latvian, Russian and English. After excluding duplicates and low quality or suspicious questionnaires, in total, 6242 emigrants aged 15 and over and living abroad took part in the survey, including 334 transnationals living simultaneously in Latvia and abroad. The large number of respondents allow us to explore separately the integration outcomes and well-being of different types of migrants. In addition to demographic variables, we are able to focus on the experiences of the most typical representatives of 'liquid migration' mobility patterns, shedding light on the specific situation in this group.

The largest risks associated with web surveys are the potential bias caused by self-selection and the difficulties of reaching certain socio-demographic groups via the Internet (Bethlehem 2010; Askitas & Zimmermann 2015). To correct the deviations caused by differences in socio-demographic composition of the respondents as opposed to the target population, sampling weights have been applied. The sampling weight of each respondent is inversely proportional to that respondent's estimated conditional sample inclusion probability, given the following characteristics (control variables): (1) geographical area of residence (8 individual countries and 7 broader regions), (2) sex (male or female), (3) age (15-24, 25-34, 35-54, or 55+), (4) level of education (low (less than secondary), medium (secondary), or high(tertiary)), and (5) main language (Latvian or Russian). If these control variables capture most of the variation in inclusion probabilities, then the weighted data yield (approximately) unbiased and consistent estimators (Horvitz & Thompson 1952). Separate weights have been computed for four emigrant groups: (a) Latvian nationals, (b) Latvian-born individuals, (c) recent emigrants (emigrated since 2000), and (d) the entire Latvian diaspora. Reference data have been compiled from multiple sources, including Eurostat, OECD's International Migration Database (IMDB) and Database on Immigrants in OECD countries (DIOC), data from Latvia's Office of Citizenship and Migration Affairs, and data from the main recipient countries' national statistics offices. The distributions of the socio-demographic characteristics (2)–(5) have been computed separately for each geographical area. Due to only marginal distributions being available within each area, a raking algorithm has been applied to produce a joint distribution matching these marginals (as in Battaglia et al. 2004). These weighting procedures have already been tested and successfully used in the first wave of the study of Latvian migrants (Mieriņa, 2019).

Special attention was paid to issues of ethics and data protection. In all cases, the project sought informed consent from research participants. In line with the national and EU legal framework on data protection, respondents were informed as to how their data is processed and stored, and how their identity is protected. e of the questions that may be asked of children, so that children are not distressed during the interview process. We also sought and received clearance on research instruments and procedures from the Ethics Committee for Humanities and Social Sciences at the University of Latvia. The data processing solutions employ the best practices of industry standard protection measures for data in transit and data at rest. At the data collection stage, communication with respondents was carried out through encrypted end-to-end channels (https protocol) while taking reasonable effort to minimise digital footprints. The raw dataset was carefully screened for any potential identifiers and pseudonymised for internal project use. All contact information was detached from the main survey datafile at the first stage of data processing. A range of measures were enforced to ensure the highest level of data protection, including: the UL access rights management regulation, declarations of compliance to confidentiality requirements by research and technical personnel, data access control, encrypted storage, strict separation of anonymised/pseudonymised and identifiers data.

Acknowledgements

This research was funded by the Latvian Council of Science, project *Exploring Well-Being and Social Integration in the Context of Liquid Migration: A Longitudinal Approach*, under Grant lzp-2018/1-0042.

References

Askitas, N., & Zimmermann, K. F. (2015). The internet as a data source for advancement in social sciences. *International Journal of Manpower*, 36(1), 2–12.

Battaglia, M. P., Izrael, D., Hoaglin, D. C., & Frankel, M. R. (2004). Tips and tricks for raking survey data (aka sample balancing). *Abt Associates*, 1, 4740–4744.

Bethlehem, J. (2010). Selection bias in web surveys. *International Statistical Review*, 78(2), 161–188.

Horvitz, D. G., & Thompson, D. J. (1952).

A generalization of sampling without replacement from a finite universe. *Journal of the American Statistical Association*, 47, 663–685.

Mieriņa, I. (2019). An Integrated Approach to Surveying Emigrants Worldwide. In *The Emigrant Communities of Latvia* (pp. 13-33). Springer, Cham.