Improving the understanding of mobility in cross-border areas using a pooled travel survey

Justin Delloye[1] · Sylvain Klein[2] · Olivier Klein[2] · Guillaume Drevon[3] ·
Mathieu Jacquot[4] · Viviane Rapp[4] · Eric Cornelis[5] · Joanna Rousseaux[5] ·
Julien Schiebel[6] · Guy Besch[6] · Stéphane Godefroy[7] · Philippe Gerber[2]

Abstract

Understanding mobility behaviour in cross-border regions is a major challenge for transport policies and the wellbeing of people living near national borders. However, the available relevant information is often limited to a collection of unharmonized national sources. To overcome this limitation, we propose a methodology for building a pooled travel survey, by merging microdata from travel surveys conducted independently in different cross-border areas of the same cross-border region. Anchored in the Luxembourg functional cross-border region as a study area, this paper provides a general, ready-to-use harmonization methodology for pooling travel surveys. It critically discusses the resulting harmonized values of standard daily mobility indicators in regard to the specificities of the study area. Lastly, it shows that the values for the proportions of trip purposes computed with the pooled survey exhibit substantial differences from values obtained using only local travel surveys (up to 6 percentage points, and up to 33 per cent in the total number of trips). This illustrates the added value of a pooled travel survey compared with an unharmonized collection of local surveys.

✉ Philippe Gerber
philippe.gerber@liser.lu

1   Université catholique de Louvain – UCLouvain, École de Géographie, Louvain-La-Neuve, Belgium

2   Department of Urban Development and Mobility, Luxembourg Institute of Socio-Economic Research - LISER, Esch-Sur-Alzette, Luxembourg

3   Laboratoire de Sociologie Urbaine, Ecole Polytechnique de Lausanne, Lausanne, Switzerland

4   Centre d'études et d'expertise sur les risques, l'environnement, la mobilité et l'aménagement - CEREMA, DTerEst/DTMI/MT, Metz, France

5   Université de Namur, naXys, Transportation Research Group (GRT), Namur, Belgium

6   Ministère de la Mobilité et des Travaux publics, Département de la mobilité et des transports, Luxembourg, Luxembourg

7   Agence d'urbanisme et de développement durable Lorraine Nord - AGAPE, Equipe Mobilité, Longwy, France

Introduction and objectives

In an increasingly integrated European labour market, cross-border regions raise specific challenges in managing passenger mobility, especially commuting issues (Chilla and Heugel 2019). The Schengen agreement and the single European market have also encouraged the growth of cross-border commuting: while around 390,000 cross-border workers were recorded in the European Union at the start of the 1990s, the increase in their numbers remained relatively slow until the beginning of the 2000s, rising to 490,000. The acceleration is apparent from then on, reflecting in particular the consequent enlargement of the EU and its labour force: from 800,000 cross-border workers in 2006, according to the sources, the number reached 1.7 million in 2015 — equivalent to nearly 1 per cent of the total EU-28 workforce (Fries-Tersch et al. 2017). The magnitude of these flows not only depends on economic opportunities and wage differentials (Gerber 2020), but also on other non-economic factors, such as language, transferability of skills and the existence of social networks that can help reduce the risks associated with mobility (Pires and Nunes 2018). The cost of travel, transport infrastructures and the distance travelled could also bring emancipation and resources for the cross-border labour market (Sohn 2014).

In cross-border regions, the national borders and local administrative limits usually do not capture the functional extension of an employment area (Brunet-Jailly 2005) and in general, transportation jurisdictions are narrower than the employment area that shapes passenger mobility. Thus, cross-border regions are heterogeneous in many aspects, including the political means and goals with regard to transport and mobility (Dörry and Decoville 2016). This heterogeneity yields inefficiencies in the management of regional transport systems (Durand and Nelles 2014). They typically fail to properly deal with large commuting flows, leaving the road network heavily congested in rush hours, whilst saturation of public transport causes discomfort and delays. To a lesser extent, non-commuting trips that people make, for example to attend social events or to go shopping, also contribute to regional congestion (Schiebel et al. 2015). Coordinated management, based on exhaustive information about cross-border passenger mobility, is required in order to address these problems. However, an integrated understanding of mobility patterns in cross-border regions is hard to grasp (Drevon et al. 2018).

To study passenger mobility in a defined area, it is necessary to take into account all the trips made there. However, they are not made only by local residents in a specific area, but also by people from outside, who come for specific purposes or are simply passing through. The proportion of trips made by outside travellers can nevertheless be particularly important in cross-border areas, where cross-border travellers, especially commuters, interact with local residents on the transport system. As cross-border travellers usually have specific travel habits (Schiebel et al. 2015), a comprehensive understanding of passenger mobility in cross-border areas requires a joint assessment of the mobility of residents and cross-border travellers. However, this joint assessment cannot be made using a mere collection of unharmonized national travel surveys from different countries. As noted in the next section, such surveys often comprise the only available information.

To overcome this limitation, the current paper proposes a methodology for building a pooled travel survey, by merging microdata from travel surveys conducted independently in different cross-border areas of the same cross-border region. More specifically, anchored in the Luxembourg functional cross-border region as a study area, this paper has three goals. The first is to provide a general ex-post harmonization methodology for building and updating a pooled travel survey. The second goal is to argue for the quality of the resulting

dataset, by computing and discussing a few standard indicators from travel behaviour research. The last goal is to demonstrate the added value of the pooled survey, by comparing two values of an areal-based indicator — the distribution of trip purposes for all trips ending in a given area — computed using the pooled survey and using only local surveys.

In doing so, we aim to sequentially answer the followings, specifically for our study area: (1) What is the conceptual proximity between local travel surveys in the Luxembourg cross-border region, and on the basis of which ontology can they be compared? (2) What are the harmonized and updated values of some standard indicators of travel behaviour in the Luxembourg cross-border region? (3) What are the distributions of travel purposes for all trips ending in three specific cross-border areas: namely, the urban neighbourhoods of Arlon (BE), Luxembourg (LU) and Thionville-Yutz (FR)? Also, are these indices substantially different from estimations using only local travel surveys?

The paper is organized as follows. In Sect. "Data sources on cross-border mobility", we present the problem concerning information about cross-border mobility, as well as the approach of the pooled travel survey, arguing that the latter provides a partial solution to the former. In Sect. "Detailed reporting of the method", we detail our study area, as well as reporting in depth the ad hoc implementation of the pooled survey approach. We strive to emphasize the challenges of its implementation in an experiment, showing ways that make it replicable in other cross-border regions. In Sect. "Results and Discussion", we provide the results supporting the achievements of the three aforementioned goals, as well as giving answers to the research questions specific to our study area. Section "Conclusion" concludes.

Data sources on cross-border mobility

Household travel surveys have long been the major source of data (Hubrich et al. 2018) to help understand passenger mobility and its transport implications. Local authorities and nation states often carry out such surveys — for example, National Travel Surveys (or NTS) — with a focus on the resident population and ignoring external travellers; especially cross-border travellers. Information on incoming travellers' trips is then traditionally assessed through vehicle counting data, which is costly and much less informative. The few exceptions of surveys that address cross-border mobility usually focus on specific aspects (see Sect. "Definition of the study area"), since carrying out an exhaustive travel survey in cross-border regions is very difficult for several reasons, for example, survey costs are higher due to their typically large area. Other potential difficulties can occur that involve additional costs: language (translations), sampling and contacting people to interview, and the need to access and to harmonize several international and interregional databases. Moreover, the difficulties in exploiting different statistical data and integrating different geographical and institutional contexts are challenging (Decoville et al. 2013).

Altogether, the need for interregional and international coordination is a serious barrier to overcome. There is a governance issue when different stakeholders and public authorities with different competencies need to manage cross-border territories where equivalent power may not exist, with different levels of competences related to different NUTS levels (e.g. state vs region vs municipality). There is currently no harmonized source of data for passenger mobility at the European level and no harmonized international methodology, as they are primarily designed to respond to national or regional information needs. There are only guidelines and recommendations to improve statistical ex-post harmonization

procedures (Armoogum et al. 2014; Eurostat 2018). Countries that already have a long tradition of NTS are reluctant to introduce major changes, for the sake of preserving their long time series.

Among alternative solutions to household travel surveys, new data sources — such as big data — have raised strong interest in recent years (Ben-Elia et al. 2018). With the increasing use of smartphones, it is possible to obtain detailed records of people and their activities over space and time (see the review by Wang et al. 2018). Mobile phone and GPS routing companies can now provide fine or more aggregated data to complement mobility survey data. However, several challenges remain before operational use, such as the control of sampling biases and the difficulty to integrate these new datasets with standard socioeconomic and demographic statistics. Moreover, the cross-border context brings new issues in the application of current big data methods (e.g. roaming problems when changing countries, or mobile phone operators that may be less representative of the population on one side of a border than on the other).

To build consistent statistics on passenger mobility in cross-border areas, we applied another solution by constructing a pooled travel survey. We created this pooled survey for a cross-border region by merging microdata from local or regional travel surveys conducted independently in different areas. The success of the method relies on the trade-off between exploitation of the conceptual proximity between independent local surveys, and the assumptions required to overcome their differences. In other words, the pooled travel survey is necessarily less complete than any of the local surveys, since only the information common to all the local surveys is retained. It nevertheless has the advantage of providing consistent information about a wider population, which is the union of all the local surveys' populations. As it is based on microdata, the pooled travel survey can be regarded as an operationalization of the first ex-post harmonization methodology considered in Eurostat (2018). Wittwer et al. (2019) developed a comparable ex-post harmonization method to provide cross-sectional analysis of mobility in several European cities.

What are the advantages of the approaches using a pooled travel survey? Starting from microdata from local travel surveys, the pooled survey offers a high degree of flexibility to retain as much information as possible. It also stands out in terms of lower costs, compared with a dedicated cross-border travel survey. Thus, its potential impact is substantial given the possibility of gathering useful information at lower cost. Pooled surveys can accordingly be used to produce consistent information about a cross-border region, and can be integrated into multimodal transport models. Thus, they ultimately support the elaboration of coordinated cross-border policies regarding transport and mobility. On the other hand, Eurostat (2018) identified the implementation time for harmonization and the necessary agreements amongst the data owners as two potential difficulties. These issues are discussed later in the current paper.

Detailed reporting of the method

Definition of the study area

Given a cross-border region of interest with some administrative data about (cross-border) workers, the goal of the pooled travel survey (providing indicators of travel behaviours and supporting a transport model) will guide the choice of the study area. In practical terms, it is necessary to choose a year of reference and a basic spatial unit (BSU). From there, the

study area is built up from cross-border BSUs; that is, places where at least one resident worked in another country of the cross-border region during the year of reference. More precisely, it is the largest contiguous area made up by cross-border BSUs that have been covered by at least one travel survey in the 10 years before the year of reference.

The time threshold of 10 years is set arbitrarily and usually results from a trade-off between the age of the local travel surveys to be pooled and the size of the study area. However, caution is required, as a larger threshold brings some issues. First, while changes in population numbers and distributions of socioeconomic profiles can be addressed by reweighting the microdata, the initial survey still has to contain the desired profiles of residents. Second, the travel behaviour that will be associated with different socioeconomic profiles will, by necessity, be assumed not to have changed since the original survey. Yet the longer the time interval, the greater the likelihood that these behaviours will have changed, for example, following changes to the transport supply. A review of major changes to the transport system during the time interval can help in assessing the validity of this assumption. The remainder of this section details the definition of the study area in the building of our pooled travel survey.

This research is actually part of the MMUST project (*Modèle MUltimodal et Scénarios de mobilité Transfrontaliers* for modelling of multimodal mobility and scenarios of cross-border transport, see www.mmust.eu), co-funded by the European Regional Development Fund (ERDF) within the framework of the INTERREG Va Greater Region programme. Its goal is to develop a multimodal transport model to simulate especially passenger transport scenarios in a functional cross-border area around Luxembourg. Thus, the pooled travel survey, its study area and the methodological choices aim to support the transport model.

Luxembourg and its cross-border region offer a striking example of a strong increase in cross-border passenger flows resulting from European economic integration (Schiebel and Gerber 2015). In fact, the transnational cooperation entity called 'Grande Région' — which includes Luxembourg and territories in Belgium, France and Germany — has the largest number of cross-border workers in the European Union, at approximately 250,000 in 2019 (Eurostat 2021). Within the EU, the largest proportions of commuting to different European countries at the regional level were observed in Luxembourg province (Belgium), at 32 per cent of the labour force, followed by Trier (Germany) at 15 per cent and Lorraine (France) at 13 per cent. In this regard, Luxembourg is a strong attractor and employed approximately 200,000 cross-border workers at the end of 2020.

In our case, the transport modelling perspective of the project imposed fine-grained BSUs no larger than the LAU2 zones of the Eurostat NUTS nomenclature. From there, the study area was built up from LAU2 zones in which at least one resident worked in Luxembourg in 2017. More precisely, it comprises the largest contiguous area around Luxembourg made up by these LAU2 zones, and that have been covered by at least one travel survey since 2010. As a result, while the Luxembourg functional cross-border region naturally includes Luxembourg and parts of Belgium, France and Germany, the latter was actually excluded due to the lack of any survey coverage. More precisely, results of travel surveys were not available at the scale of the chosen BSUs. Thus, the final study area covers the whole country of Luxembourg, the Belgian 'Province de Luxembourg' and in France, the regions of influence of three institutional entities with spatial planning and transport competences, the SCoTs (in French *Schéma de Cohérence Territoriale*): SCoT *Nord Meurthe-et-Moselle*, SCoT *Agglomération de Thionville* and SCoT *Agglomération de Metz* (Fig. 1). The study area covers 10,905 km^2 and had 1,669,095 inhabitants in 2017, including 118,210 cross-border workers employed in Luxembourg (of whom 28% lived in Belgium and 72% in France).

The identification of the relevant surveys must be carried out simultaneously with the delineation of the study area. Any methodical reviewing is appropriate here, but we underline that in our case, transport-related public institutions turned out to be insightful sources of

Fig. 1 Study area

information. In the end, the input data for our study area came from eight regional travel surveys, which cover the entire study area with the most recent information (Fig. 2 and Table 1). The Belgian national mobility survey, BELDAM (2010) (Cornelis 2012) is used to cover the Belgian province of Luxembourg. Five travel surveys cover the French side of the study area and were independently carried out by the CEREMA (centre for studies and expertise on risks, the environment, mobility and development). Luxembourg is covered by the national mobility survey, LUXMOBIL (2017). The LUXMOBIL survey also sampled cross-border workers at their place of employment in Luxembourg, regardless of their place of residence. By comparison, the EMF survey (Enquête Mobilité des Frontaliers), also focusing on cross-border workers, sampled them at their place of residence (Schmitz et al. 2012; Enaux and Gerber 2014). In addition, the LUXMOBIL and EMF surveys focus on different aspects of cross-border mobility. While the EMF focuses on cross-border workers' mobility regardless of their trip purposes and without taking into account the whole population traveling on the territory, LUXMOBIL considers the mobility of Luxembourg residents and cross-border workers for any purposes, but does not consider other cross-border non-workers traveling to
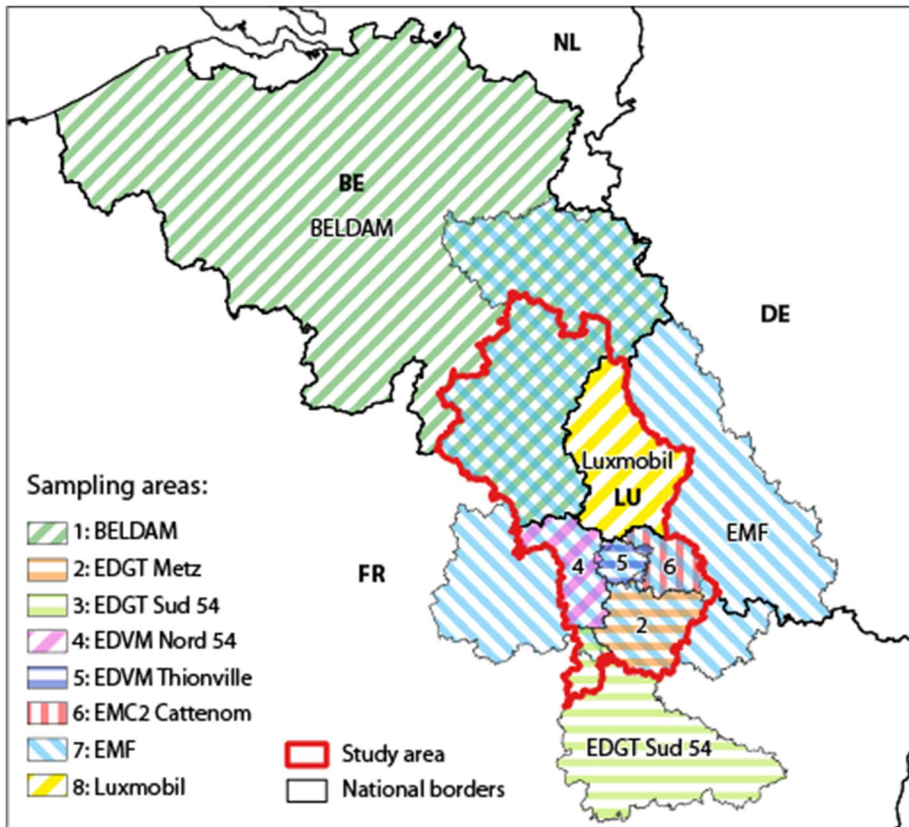
Fig. 2 Sampling areas of the eight surveys used in the pooling. *Notes: Please see* Table 1 *for an overview of the eight surveys. In addition to residents of Luxembourg, the Luxmobil survey also includes cross-border workers who work in the country, wherever their place of residence*

Luxembourg for purposes other than work. These two surveys therefore do contribute, but cannot substitute for a pooled survey analysis as presented in this paper. The eight local surveys were carried out during a 9-year period between 2010 and 2019. It should be noted each survey covers different time periods, thus comprising the first issue in the construction of the pooled composite travel survey, as discussed at the beginning of this section.

Harmonization of the observations

The harmonization of observations is the selection of relevant rows in the microdata from the local surveys. However, no examination of the microdata is required to infer the selection rules, given that the local surveys have proper metadata (typically a technical report with the sampling plan and a copy of the questionnaires). The goal of this harmonization is to avoid exclusion biases generated by the pooling process. The first step consists of identifying the actual comparison basis of the local surveys; that is, their common conceptual objects, and their definitions, properties and relationships (in other words, their common ontology). Then, as Wittwer et al. (2019) explain, this harmonization actually deals with four fields: the surveys' coverage (e.g. populations of interest), the surveys' definitions (e.g. definition of trips), the surveys' method

Table 1 Travel surveys used to build the pooled one

ID	Acronym	Surveyed population (place of residence or workplace)	Sampling country	Year
1	BELDAM	Belgian residents	Belgium	2010
2	EMF	Cross-border workers working in Luxembourg	Belgium, France, Germany	2010
3	LUXMOBIL	Luxembourg residents and cross-border workers working in Luxembourg	Luxembourg	2017
4	SIMO/EDVM Thionville	Residents of the SMiTU (*Syndicat Mixte Transports Urbains*) around the city of Thionville (see Fig. 2)	France	2012
5	SIMO/EDGT Sud 54	Residents of the *Sud Meurthe-et-Moselle* (cf. Figure 2)	France	2012–2013
6	SIMO/EDVM Nord 54	Residents of the *Nord Meurthe-et-Moselle* and *Pays Haut Val d'Alzette* (cf. Figure 2)	France	2013–2014
7	SIMO/EDGT Metz	Residents of the 2016 SCOTAM (*Schéma de Cohérence Territoriale Agglomération Messine*) area (see Fig. 2)	France	2016–2017
8	SIMO/EMC2 Cattenom	Residents of an ad hoc area around Cattenom (see Fig. 2)	France	2019

The surveys in France were conducted following the methodology of Cerema, in which this methodology differs according to the extent of the surveyed area (CEREMA 2020). Thanks to the similarities of these surveys in France, a pooled survey was able to be performed separately by Cerema under the title SIMO

(e.g. sampling design) and defining the surveys' area (e.g. the granularity of the spatial resolution varies according to the definition of the statistical units used in each survey).

In the context of a pooled travel survey, the specific goal of the harmonization of the observations is to identify and control for methodological differences that may introduce exclusion biases and oversampling biases in the final pooled survey. First, pooling-generated exclusion bias is explained by the following example: A local survey may integrate teenagers in the population of interest, while another, targeting other goals, may not. Yet once considered together in a pooled survey aiming to produce indicators about teenagers, this would introduce an exclusion bias of teenagers in some parts of the study area. Since the information on missing teenagers cannot be found afterwards, the only way to avoid this bias is to discard the microdata for teenagers in all other local surveys. The same logic extends to any kind of sampling units in the complex sampling plans (households, trips, etc.), as illustrated in our implementation of the pooled travel survey.

Second, pooling-generated oversampling bias is explained by the following example: A local travel survey may include non-resident travellers in the population of interest, whereas another may not. Yet once considered together in a pooled survey, this could introduce oversampling of the residents in certain areas who would have been sampled twice (as a non-resident traveller in one area, and as a resident in the other). This bias can be limited by using a reweighting procedure, which is explained in Sect. "Reweighting the pooled microdata". The remainder of the current section details the results of the harmonization of observations from the local travel surveys in the building of our pooled travel survey.

There are six conceptual objects that are present in all or some of the eight local travel surveys used in this research: households, vehicles, individuals, public transit passes, opinions and trips. A household is a group of individuals living in the same dwelling (up to 10 members in the surveys used here). The place of residence, which may be a postal address or another basic spatial unit such as a municipality, is the location of the dwelling and hence a key characteristic of the household. It should be noted that individuals who divide their life between two or more regular residences are excluded by some surveys and hence excluded from the pooled survey (see Appendix Table 8). Individuals are human beings, who are grouped into households. Vehicles, which are mechanical and potentially motorized means of transport, are associated with households. Public transit passes and opinions on transport and mobility are linked to individuals, and both are reported in some of the local surveys. Lastly, trips are related to individuals, and describe movements over space characterized by an origin, a destination, departure and arrival times, a purpose and one or several transport modes.

We underline that this ontology seems general enough to provide an overall framework for the comparison of travel surveys and, accordingly, it provides valuable output for the generalization of our work. As mentioned above, comparing the local travel surveys in the light of these conceptual objects exposes differences in the coverage, definitions, method and area definition of surveys that may induce exclusion and oversampling biases. The encountered issues are now discussed.

First, with regard to coverage, some conceptual objects are not used in every local survey. Opinions on transport and mobility are included (in different ways) in all the surveys but one — LUXMOBIL. Besides, public transit passes are absent in the French surveys. Thus, information on opinions and public transport passes are discarded to avoid exclusion biases in the pooled travel survey. Only information about households, individuals and trips will remain in the pooled survey (see Table 1). This loss of information has no influence regarding our application to the Luxembourg area; however, depending on the content of local travel surveys, it may restrict the application of the pooled travel survey method to other purposes.

Second, for the method of surveys, local surveys have complex sampling plans based on various cluster sampling structures. The EMF survey, which examines the population of cross-border workers employed in Luxembourg, takes the individual as the primary sampling unit. In that survey, cross-border workers are the only individuals whose trips were surveyed, and they were asked about their household on behalf of the other household members. By contrast, in the other local surveys, the household is the primary sampling unit and all or some of its members were asked about their trips. Aside from the cluster structure, sampling plans also differ by the sub-population they use. In some surveys, such as the EMF, individuals surveyed on their trips were deterministically chosen on the basis of their individual characteristics. In other surveys, such as the EDVM Thionville and EDVM NORD54, they were randomly sub-sampled. The methodical examination of the sampling plans enables us to split the populations of interest of each local survey into a set of subpopulations with the associated gathered information. To avoid exclusion biases, only the subpopulations sampled by all local surveys are integrated into the pooled survey. In our case, facing the constraints presented in Annex 1, the microdata from our eventual pooled travel survey contains all the respondents from the local surveys who were living in the study area. While the values for the socioeconomic variables are available for all of them (variables 1 to 20 in Table 1), only a subpopulation of the individuals older than 11 were asked about the trips they took on a specific day (for each respondent in all the local surveys, the day of interest was randomly chosen among working days).

Third, also with regard to the methods, French local transport surveys were carried out by interviewers, in comparison with the Belgian and Luxembourg local transport surveys, which were self-administered. As a result, Belgian and Luxembourg respondents were more likely to have forgotten small trips, especially those made on foot. The resulting bias is clearly apparent in our local surveys, but it cannot easily be disentangled from the other effects of interest, for example, due to a specific age structure of the local population. Therefore, since omitted short trips could not be subsequently accounted for, we decided to exclude them from our results altogether to avoid local exclusion biases. We set a threshold at 5 km on the basis of the content-related pooled survey, following a natural break of distance distribution of the trips which avoid excluding too many cross-border trips.

Fourth, for the defining of survey areas, all local surveys except EMF and LUXMOBIL surveyed residents in mutually exclusive areas (see Fig. 1). In fact, the LUXMOBIL survey also sampled cross-border workers at their workplace in Luxembourg, regardless of their residence place the EMF survey, also focusing on cross-border workers, sampled them at their residence place (Schmitz et al. 2012; Enaux and Gerber 2014). To limit the potential oversampling bias, we use a reweighting procedure, which is detailed in Sect. "Reweighting the pooled microdata".

Fifth and last, with regard to definitions in the surveys, local surveys also exhibit differences in their variables of interest. Thus, similar variables may have different codifications. For example, the education level of individuals is included in every local survey, but it follows the different national systems of educational qualifications, and hence the variables are not immediately comparable. In our Luxembourg case, the same problem occurred both for current activity status and trip purposes. Identifying variations in the codification of similar variables, and building new variables to merge them, is equivalent to defining the variables dictionary of the pooled travel survey presented in the next section.

Harmonization of the variables

The harmonization of variables involves the selection of relevant columns in the microdata from the local surveys to avoid exclusion biases generated by the pooling process. The difference from the harmonization of observations is that dropping variables is not the only solution: the definition of a new categorical variable that retains the information available in all local surveys often solves the problem.

The pooling-generated exclusion bias related to surveys' definition of variables is explained in the following example. A local survey may ask respondents about their highest level of formal education in three categories with 'higher education' as the maximal one. By comparison, another local survey may ask the same question with four answer categories, 'Doctorate' being the new maximal one. If considered together in a pooled survey aiming to produce indicators about PhD holders, this would introduce an exclusion bias concerning those with doctoral degrees in some parts of the study area where they are not separated from people with a Master's degree, for example. One way to avoid this bias would be to define a new variable for the highest level of formal education — a variable that enables the recoding of all the different categorizations of educational qualifications used in the local surveys. However, if there is at least one local survey that does not ask respondents about their highest level of formal education at all, then discarding this variable is the only solution. The remainder of this section details the results of the harmonization of variables from the local travel surveys when building our pooled travel survey.

First, the age of the respondents was directly asked in all the surveys except the EMF, which asked for their birth year. As a result, the pooled age variable is the age reached. This is approximated by the age in completed years for respondents of the EMF survey.

Second, the item for level of education differs between surveys in accordance with their population of interest and national education system. An ad hoc correspondence table can be drawn up by associating education levels that correspond to the same age group of a regular cursus (see Appendix Table 9). For example, the Belgian level *secondaire général* and the French level *secondaire, titulaire du bac* are part of the same pooled category *secondaire*, as they both designate a qualification achievable in at least 12 consecutive successful years of schooling. The resulting pooled variable aggregates from seven to ten categories of the initial variables into five categories (not including missing values).

Third, in contrast to education level, the variable for activity status could not always be generated by recoding a single variable. Instead, decision rules implying several variables were used (see Appendix Table 10). For the BELDAM and EMF surveys, decision rules were elaborated on the basis of, respectively, four and two variables specifying whether the individual was at work, a full-time or part-time worker, and whether she/he was a student. For the LUXMOBIL and French surveys, a simple correspondence table was used. The aggregation yields six categories (not including missing values) in the pooled variable, against nine categories in the LUXMOBIL and French surveys.

Fourth, trip purposes are categorized differently in the local surveys. The French surveys in particular, consider 36 different purposes whilst the others include 11 to 18 purposes. Nevertheless, the additional modalities of French surveys actually indicate a finer level of detail rather than conceptual differences. As a result, a 12-category trip purposes variable could be built by correspondence (Table 2).

Sixth, the departure and arrival times of trips are recorded in a comparable way in all the surveys, but the collection period differs. While it covers a period from 4 a.m. on the reference day to 4 a.m. the following day, LUXMOBIL shows trips taking place from 0

Table 2 Categorization of the trip purpose

Id	Category in the pooled survey	Corresponding categories in the local travel surveys			
		BELDAM	EMF	LUXMOBIL	SIMO
1	**Drop off or pick up someone**	1. Drop off / pick up someone	1. Drop off / pick up someone	5. Drop off / pick up someone 13 (Pre-)school 14 Children day-care centre 15 Accompanying adults	61 Accompanying someone (who is present) 64 Pick up someone (who is not present yet) 71 Drop off someone (who is present) to a transport node (e.g. bus stop) 74 Pick up someone (who is not present yet) from a transport node (e.g. bus stop)
2	**Go home**	2 Go home	2 Go home	18 Home 19 Returning from previous trip	1 Home (depart from or go to)
3	**Go to work**	3 Go to work	10 Go to work	1 Reaching the workplace	11 Work at the usual workplace 12 Telework 13 Work somewhere else (not teleworking) 14 Work somewhere else (no distinction)
4	**Travel for work duties**	4 Travel for work duties	3 Travel for work duties	16 Official or business duties	81 Professional travel (e.g. deliveries)

Table 2 (continued)

Id	Category in the pooled survey	Corresponding categories in the local travel surveys			
		BELDAM	EMF	LUXMOBIL	SIMO
5	**Taking a course**	5 Taking a course, (school, etc.)		2 Going to school	21 Be looked after (childminder, nursery school, etc.) 22 Studying at the usual preschool or elementary school 23 Studying at the usual middle school 24 Studying at the usual high school 25 Studying at the usual university or college 26 Studying at another place (pre-school or elementary school) 23 Studying at another place (middle school) 24 Studying at another place (high school) 25 Studying at another place (university or college)
6	**Have a meal out**	6 Have a meal out	4 Have a meal out	9 Have a meal out	53 Have a meal out
7	**Grocery and shopping**	7 Grocery and shopping	5 Grocery and shopping	3 Shopping	30 Visiting a store, mall or open-air market, without purchases 31 Do several things in a shopping mall 32 Shopping in department stores, supermarkets and their shopping malls 33 Shopping in small and medium-sized shops 34 Shopping in a market (covered/open air) 35 Shopping in a drive-through store

Table 2 (continued)

Id	Category in the pooled survey	Corresponding categories in the local travel surveys			
		BELDAM	EMF	LUXMOBIL	SIMO
8	**Personal services (doctor, bank, etc.)**	8 Personal services (doctor, bank, etc.)	6 Personal services (doctor, bank, training, etc.)	4 Personal services	41 Receiving care (health) 42 Doing something other than looking for a job 43 Looking for a job
9	**Visit family or friends**	9 Visit family or friends	7 visit family or friends	7 Visit family or friends 10 Care or grocery shopping for someone else	54 Visit family or friends
10	**Go for a walk, go for a ride**	10 Taking a walk, going for a ride	8 Taking a walk, going for a ride		52 Go for a walk, window shopping, take a driving lesson 82 Shopping trip without purchase
11	**leisure, sports, culture**	11 Leisure, sports, culture	9 Leisure, sports, culture	6 Leisure activity (sports, recreation)	51 leisure, sports, cultural or associative activities 55 'POKÉMON GO' walk
12	**Other**	12 Other	11 Other (no details)	8 Refuelling 11 vacation at least 2 days 12 other	2 Second home, occasional accommodation, hotel, other home (leaving from, going to) 91 Other motives (no details)
na	**No answer**	. No answer	12 No answer	98 I do not know 99 I do not want to answer	NA No answer

a.m. to midnight on the reference day. Thus, only trips that took place between 4 a.m. and midnight on the reference day were kept in the pooled survey.

Seventh, the last step is the equivalent of a categorization for the geographic variables (place of residence, and origin and destination of trips). To cover both the study area and the neighbouring ones, the BSUs are defined for the whole countries of Belgium, France, Germany and Luxembourg. In each country, the BSUs necessarily correspond to the coarser BSU used by any survey, and not only the local surveys. Thus, it is crucial to the success of the pooled survey that local surveys use fine-grained spatial units to describe origins and destinations that are located in the neighbourhood of their area of interest, and especially in border countries. For example, in this study, it is important that a survey carried out in France uses a fine-grained BSU to describe not only the French territory, but also the Belgian, German and Luxembourg ones. Fortunately, this is often the case in cross-border regions, and it is certainly true for the current study area. As a result, in our pooled survey the BSUs are LAU2 zones in each country of the study area, and national territories for other countries.

After all these steps, we are left with 40 complete variables that are divided into the three categories of households, individuals and trips (with 8, 12 and 20 variables, respectively). These variables are listed in Table 3. We emphasize that, in the case of a multi-modal trip, the sequence and locations of intermediate stops to change transport mode are unknown. Hence the use of binary variables for transport modes reflects the possible multiple modes of transportation characteristics (see Table 3, rows 35 to 40). The last step in elaborating the pooled travel survey is to weight the harmonized microdata.

Reweighting the pooled microdata

As with a traditional survey, the weighting step consists of computing estimation weights that are used to derive the population parameters of interest from the pooled survey. The first step is to compute the harmonized margins — that is, the margins of the harmonized variables — using socioeconomic data. Since cross-border regions rarely have ready-to-use harmonized statistics, this work requires contacting the official statistical institutes of all the countries in the study area. The second step is of course to perform the weighting of the microdata based on these margins. In the context of a pooled travel survey, the specific goal of *re*-weighting the microdata (which was already weighted once in the original local surveys) is to address the oversampling bias in areas where residents were sampled by several local surveys (see the example in Sect. "Harmonization of the observations"). The principle of reweighting is as follows.

First, we assumed the original weights to be best suited to compensate for sampling biases in the original surveys. This assumption is not only pragmatic (the original surveys comprise the input in our work and cannot be altered in retrospect), but also reasonable, since all the original surveys used complex sampling plans and relevant weighting procedures in order to tackle sample biases such as under-sampling, oversampling and non-response.

Second, we performed reweighting of the individuals present in the pooled survey, using the harmonized margins over the entire cross-border area and taking the original weights as the starting ones for the weighting process. That is, starting from the vector of weights households or individuals had in their local travel survey, the weighting algorithm aims to give the closest vector of weight that matches the harmonized margins. In this way, we compensate for any pooling-induced oversampling bias while departing as little as possible from the original weights. We should underline that addressing this bias is the reason why accessing the microdata with the original weighting is crucial for the reproducibility of the methodology.

Table 3 Variables dictionary of the pooled travel survey

	Level	Variable
1	household	Household's ID code
2	household	Original local travel survey
3	household	Household's residence location
4	household	Country of household's residence
5	household	Number of household's members
6	household	Number of four-wheeled vehicles available in the household
7	household	Household's weight in the original local travel survey
8	household	Household's weight in the pooled survey
9	individual	Individual's ID code
10	individual	Age
11	individual	Sex
12	individual	Highest educational qualification
13	individual	Activity status
14	individual	Was the individual surveyed about his or her movements?
15	individual	Has the individual declared at least one trip?
16	individual	Works in Luxembourg and lives elsewhere
17	individual	Does the individual own a car driving licence?
18	individual	Is a private parking spot available at work or at school?
19	individual	Individual's weight in the original local travel survey
20	individual	Individual's weight in the pooled survey
21	Trip	Trip's ID code
22	Trip	Trip's number in the trip chain
23	Trip	Trip's origin
24	Trip	Trip's origin country
25	Trip	Flag. Is the trip's origin within the study area?
26	Trip	Trip's destination
27	Trip	Trip's destination country
28	Trip	Flag. Is the trip's destination within the study area?
29	Trip	Trip's purpose at destination
30	Trip	Departure hour
31	Trip	Departure minutes
32	Trip	Arrival hour
33	Trip	Arrival minutes
34	Trip	Trip's length (km)
35	Trip	Trip by feet only
36	Trip	Use bicycle or not
37	Trip	Use motorcycle or not
38	Trip	Use car or not
39	Trip	Use bus or not
40	Trip	Use train or not

In our pooled survey, estimation weights are computed using a generalized ranking procedure, implemented in the ICARUS package of the R statistical software (Rebecq 2019). The final weighting margins contain age, sex, work status (working, retired or inactive) and

cross-border worker status (yes, no). Most of the statistics used in the weighting procedure came from the national statistical offices of Belgium (Statbel), France (INSEE) and Luxembourg (Statec). The numbers of cross-border workers employed in Luxembourg were provided by the Luxembourg social security office (IGSS).

Results and discussion

This section presents and discusses some of the main daily mobility indices estimated using the pooled travel survey. The harmonized mobility indices computed for all residents in the study area are discussed first. Then, to emphasize the potential of the pooled travel survey, the proportions of purposes for all trips ending in a selected area, whether made by local residents or cross-border travellers, are presented for three urban locations in the study area.

Generalization of the method

The methodology for building a pooled travel survey for a given year of reference in any cross-border region of interest can be roughly summarized by the following eight steps:

1. Define the goals of the pooled survey.
2. Gather information about the cross-border labour market and the travel surveys carried out in the region over previous years.
3. Define the study area according to the goal, by trading-off between (1) its size, (2) the resolution of the basic spatial units and (3) the time threshold beyond which local travel surveys are not considered.
4. Analyse differences in local surveys with regard to the coverage, definitions, methods and area definitions in the surveys in order to identify potential pooling-induced sampling biases.
5. Harmonize the observations by discarding those that raise pooling-induced exclusion biases.
6. Harmonize the variables by recoding into new ones those that raise pooling-induced exclusion biases, or discarding them if necessary.
7. Compute harmonized margins for the reweighting procedure, using the new harmonized variables.
8. Reweight the microdata to limit any pooling-induced oversampling bias.

This procedure not only provides transport studies with a ready-to-use yet flexible methodology, but also offers an opportunity for future research aimed at establishing pooled surveys that can facilitate comparisons of cross-border regions. In addition, the ontology we presented in Sect. "Harmonization of the observations" appears sufficiently general to facilitate the realization of step 4 in future implementations of the pooled survey method. Nevertheless, the generalization of the method and its application to other cross-border regions face the potential issues raised in Sect. "Detailed reporting of the method". Obviously, the availability of previous local travel surveys, including complete (yet pseudonymized) microdata and exhaustive metadata is crucial.

In the perspective of generalization, it is worth noting that individual travel habits that have been surveyed through questions on the frequency of trips instead of being inferred from trip diaries are difficult to include in the pooled survey because the

formulations of these questions often exhibit a lot of differences among the original surveys. A good example from this work are questions about the general frequency of car use. Every local survey measured frequency in different and barely compatible ways. The BELDAM survey expresses frequency as a count of days with car use per week, the French surveys use a count of trips carried out by car per week and LUX-MOBIL uses a qualitative scale. In the EMF, the frequency of car use is approached through a distinct census of 'regular trips', including variables for the transport mode and frequency. In this latter survey, the frequency is not a continuous variable but a categorical one (less than once a week, once a week or more than once a week). Thus, an important limitation of our work is that the pooled travel survey method may not be as effective for pooling surveys based on frequency-questions as it is for diary-based travel surveys.

Beyond the pooled travel survey, we should underline that in addition to the prospects for mobile data that could compensate for certain data deficiencies — particularly in terms of updating already established traffic models — emphasis should also be placed on research governance issues and concerted action with regard to the implementation of transport surveys. Even if a pooled survey provides reliable and robust statistics, it would be appropriate for mobility stakeholders to work together upstream between regions in adjacent countries to ensure that comparable survey protocols are put in place. These should include questionnaires that respect certain existing and proven European and/or academic standards. For example, we could note the efforts made by the European Commission's DG Move to set up a standard European questionnaire (Armoogum et al. 2014), and the various research initiatives aimed at improving questionnaires in line with shortened, comparable versions that address both subjective and objective dimensions (e.g. von Behren et al. 2018).

Harmonized values of standard indicators

In order to critically discuss the quality of the pooled travel survey, the harmonized modal proportions and distribution of trip purposes are presented hereafter. We should emphasize that in this subsection, the figures per country refer to the place of residence of people travelling. Whilst the population of interest comprises approximately 1,723,000 individuals and 727,000 households, the final pooled survey pro&vides a sample of 57,154 individuals (3.3%) distributed in 30,905 households (4.3%). Some 134,304 trips were declared, yielding an estimated 4,663,170 total daily number of trips achieved by residents over the age of 11 in the study area. Residents in the study area make on average 3.8 trips a day. French residents are the most mobile (4.4 trips per day), ahead of those in Belgium (3.4 trips/day) and Luxembourg (3.2 trips/day). However, the estimated proportion of non-travelling residents is roughly equal in France and Luxembourg (12.3 and 12.1 per cent, respectively), far behind Belgium (34%). As already mentioned, it is not possible to disentangle the effects of the self-administration of questionnaires from the specificities of the Belgian Province of Luxembourg, which is more rural than the rest of the study area.

In order to compensate for differences in the way questionnaires are administered (e.g. Kalfs 1995) in different countries/regions, we had to exclude trips shorter than 5 km. This is not an issue in our application, which is a four-step model of longer cross-border trips. With regard to the modal proportions (for trips longer than 5 km,

see Table 4), 84% of them are made by car. This proportion peaks at 88% in France, ahead of 86% in Belgium and 77% in Luxembourg. By contrast, the proportion of public transport trips is highest in Luxembourg (15%), ahead of Belgium and France (both at 7%).

These figures can be understood in light of the geography of the study area. The population density in Luxembourg was 233 inhabitants per km² in 2018; far higher than the values in the French and Belgian sectors of the study area (respectively, 155 and 64.3 inhabitants per km²). Thus, places of interest are closer to each other and better connected to the public transport network. Apart from geographical concerns, Luxembourg also has a history of public transport policies, and these reached a peak in 2020 with the free public transport policy (Diederich 1998). Moreover, the afore-mentioned congestion of the road network in Luxembourg adds another incentive to use alternative means of transport.

With regard to the distribution of trip purposes, for those of more than 5 km (Table 5), we note that travelling for work is very apparent on all sides. It is shopping purposes that differ between France and the other areas; for the rest of the activities, the results are quite similar.

As in the previous results, in Table 6 we consider residents of the MMUST perimeter who travel on a working day. On average, with trips of more than 5 km, the residents considered make 2.6 daily trips.

European trends for daily trips are around an average of 3.54 per traveller, with all trips (short or long) combined (Armoogum et al. 2014, p. 88). This difference can easily be explained by the relatively long travel times and distances for all travellers within the MMUST perimeter. Among them, those from Belgium cover the most distance on average (almost 70 km per day), with an average total travel time of around 1.5 h per day. Travellers from Luxembourg logically cover the smallest distances (45 km), close to the European average (regardless of distance, see Armoogum et al. 2014), but with a relatively long travel time (82 min).

Proportions of trip purposes for travel ending in urban districts

We lastly illustrate the potential of the pooled travel survey to produce new mobility indices. In contrast to the harmonized distribution of trip purposes discussed in Sect. "Harmonized values of standard indicators", the figures per urban area for the

Table 4 Modal shares per region of residence for traveller trips of more than 5 km (in %)	Countries (MMUST perimeter)	Active modes	Car	Public transport	Motorized multimodal
	Belgium	2	86	8	4
	France	1	88	8	3
	Luxembourg	2	77	15	5
	Total	2	84	10	4

Table 5 Trip purposes per region of residence (MMUST perimeter) for trips of more than 5 km (in %)

Main activity	Belgium	France	Luxembourg
Go to work	32	24	36
Go to school	8	8	9
Accompanying	14	15	11
Grocery and shopping	15	21	8
Leisure, visit family-friends	21	24	19
Personal services, health, other	10	8	18
Total	100	100	100

Table 6 Distance, time and number of trips: average per day and region of residence for trips of more than 5 km

Countries (MMUST perimeter)	Average distance (km)	Average time (minutes)	Number of trips (average)
Belgium	67,85	88,95	2,57
France	65,90	87,52	2,68
Luxembourg	45,13	82,48	2,44
Total	58,99	86,01	2,58

proportions of travel purposes presented in this section refer to the urban area where the trip actually took place, regardless of who made them (residents or not). More precisely, we computed the proportions of trip purposes for all trips ending in a selected area — regardless of whether conducted by local residents or cross-border travellers — for three areas of the Luxembourg cross-border region. The district of Arlon (in Belgium), the city of Luxembourg (in Luxembourg) and the Thionville-Yutz district (in France) are three main urban areas with different profiles. The district of Arlon ranks second in terms of population, with 61,899 residents and an employment area of 17,819 jobs. Luxembourg City has a larger population, with 114,084 residents, and is a major employment area with 153,611 jobs. Lastly, Thionville-Yutz has 61,439 residents and 17,999 jobs. Geographically, Arlon lies at the Belgian-Luxembourg border, while Thionville-Yutz is roughly 15 km away from the French-Luxembourg border. Luxembourg City lies approximately equidistant to the Belgian and French borders, roughly 15 km from each.

The total number of trips and shares of trip purposes that end in a selected area were computed using both local surveys and the pooled survey, and are presented in Table 7.

First, the number of trips made in the three areas are substantially higher in the pooled survey (by between 10 and 33%). This confirms and quantifies the presence of cross-border travellers in these areas. The increase is especially large in Arlon (33%). The higher proportion of trips made by cross-border travellers is unsurprising, as Arlon is located on the border. Two patterns appear with regard to the changes in proportions of trip purposes, which reflect specificities in the reasons for trips by cross-border travellers. On the one hand, in Arlon and Thionville-Yutz, the proportion of trips to work has the highest increases: respectively, $+2.89$ and $+4.98$ percentage points (pp). This confirms that the trips made by cross-border travellers in these areas are essentially to go to work, and quantifies the areal intensity of this cross-border interaction. In Arlon, the second-highest increase is observed in trips made for personal services ($+1.10$ pp), whilst in Thionville-Yutz, the second-highest increase is

Table 7 Shares of trip purposes and total number of trips for all travel ending in the selected areas

	Arlon			Luxembourg			Thionville-Yutz		
	local survey	pooled survey	diff	local survey	pooled survey	diff	local survey	pooled survey	diff
Drop off or pick up someone	14.97	12.61	-2.36	5.51	6.30	0.79	10.49	8.99	-1.50
Go home	38.04	37.99	-0.04	29.72	30.14	0.41	38.48	32.44	-6.04
Go to work	8.27	11.16	2.89	32.62	29.09	-3.52	7.77	12.75	4.98
Travel for work duties	0.59	0.93	0.34	3.50	3.40	-0.10	0.23	0.22	-0.01
Taking a course	8.96	7.39	-1.57	5.95	6.64	0.69	5.49	7.19	1.70
have a meal out	1.41	1.15	-0.27	4.15	4.27	0.12	2.30	2.49	0.20
Grocery and shopping	11.27	10.70	-0.57	2.57	3.02	0.45	17.96	17.77	-0.19
Personal services (doctor, bank, etc.)	3.13	4.23	1.10	3.89	4.49	0.60	4.23	4.88	0.65
visit family or friends	5.18	4.83	-0.35	2.31	2.56	0.25	4.03	4.18	0.15
Going for a walk, going for a ride	2.35	2.28	-0.07	0.00	0.03	0.03	5.53	5.01	-0.52
leisure, sports, culture	2.96	3.39	0.42	4.84	5.09	0.25	3.25	3.41	0.16
Other	2.86	3.33	0.48	4.95	4.97	0.03	0.26	0.67	0.41
Total number of trips	**85,842**	**114,537**	**28,694 (+33%)**	**429,436**	**472,906**	**43,470 (+10%)**	**210,775**	**233,684**	**22,909 (+11%)**

'Local' or 'pooled' activities surveys are shown in %, and 'diff'. stands for difference in percentage points. All trips ending in any of the three selected urban areas were considered. However, in the local surveys, these trips comprise those made by local residents only, while in the pooled survey, trips of cross-border travellers are also included

observed in travel for taking courses (+ 1.70 pp), which is consistent with the presence of cross-border students. In both areas, the major decreases in proportions of travel purposes are found in trips made to drop off or pick up someone (-2.36 pp in Arlon, -1.5 pp in Thionville-Yutz) or to go home (-0.04 pp in Arlon, -6.04 pp in Thionville-Yutz). This result seems natural, as those two purposes imply that travellers either have a house in the area — which by definition is only the case for residents — or have acquaintances living in the area, which is more likely to occur for residents due to the spatial proximity of their living places.

On the other hand, the pattern is different in Luxembourg City: the shares of trips made to go to work and for work duties decrease substantially (respectively, -3.52 pp and -0.10 pp) whereas all the other proportions increase. The reason for this striking pattern is to be found in the specificities of the local survey. The Luxmobil survey not only targeted Luxembourg residents, but also cross-border workers. As a result, the cross-border trips made by cross-border workers were captured, particularly their 'go to work' travel, but not such trips made by people who were not cross-border workers, and hence can only have other purposes than going to work. In that area, the pooled survey essentially includes cross-border trips made for non-professional purposes, which is different from the two other areas studied here.

Discussion

In a European context involving three different countries, this article presents the possibilities offered by combining several national or regional mobility surveys, subsequently making it possible to feed, among other things, a traffic model in the general framework of an Interreg project. As a result of a very large amount of supporting secondary data, it has been possible to drastically limit the costs of fieldwork, while giving five undeniable aspects of added value to the daily travel survey data, despite nevertheless presenting certain limitations.

The first added value is the harmonization of the databases, thus allowing them to be compared on three national sides. This remains rare at a European or even global scale, with a particular difficulty being the existence of comparable data between the sides (Chilla and Heugel 2019). The second added value is the improved estimation of the number of trips at the individual level, both in spatial and temporal aspects, representing undeniable support for the development of sustainable policies (Liang et al. 2020). First, at the spatial level of the statistical areas considered, the confidence intervals narrow in most regional cells. At the temporal level, adjustments were able to be made according to a common recent period (2017) for all data. The third added value is the sharing of knowledge of the field in relation to surveys on both sides of borders, whether referring to the design of the samples, the survey modules or the exploitation of the results themselves. The fourth added value is the possibility of delivering to stakeholders and political decision-makers results that are common to all three sides. These results are more recent, more extensive in terms of spatial coverage and at lower cost due to the use of secondary data. This indicates a certain willingness to engage in cross-border governance at the European level in order to facilitate its territorial integration (Decoville and Durand 2021). The fifth added value lies in the future use of the results at the individual or household level in different European countries, particularly through the introduction of open-access data sharing. This research perspective

would make it possible to strengthen our knowledge of border areas and, if necessary, compare it with that of other researchers and other study areas, thereby broadening the scientific community. The methodological aspects underpinning the pooled travel survey will also be shared, so that some of these aspects can be emphasized to other actors involved in survey designs. For example, German stakeholders (e.g. The Federal Ministry for Digital and Transport; BMDV) may choose to complete their national travel questionnaire, at least in the border regions of Saarland and Rhineland-Palatinate, to cover the same standards linked to the 40 variables selected in our pooled survey. This will not only enable us to start from a broader common base, but also to analyse survey results that are more comparable between the border regions, and ultimately to develop an integrated traffic model with the German counterparts.

Nevertheless, in the age of digitalization, we may consider new collection methods that would make it easier to cross borders while ensuring homogenization of the protocols and the recent and promising information that can be obtained. With information and communication technologies, new perspectives are emerging for the study of people's daily mobility, and have the potential to offer real alternatives to traditional transport surveys (Bonnel et al. 2017). Whatever their purpose, these technologies usually automatically and passively produce large masses of data. Because of the size, this data has the advantage of being able to incorporate larger samples (close to exhaustive), with greater spatial coverage (not limited to the residents of a given area), and longer temporal coverage (well beyond the traditionally used weekday). The most interesting and promising solution is based on the mobile phone that most people carry with them during the day. The relevant data offers the possibility of tracking door-to-door movements and can be used to satisfactorily reconstruct the origin destination flows (Raballand and Le Corre 2018; Dypvik Landmark et al. 2021). However, there may be limitations in less densely populated areas, where there are fewer flows and therefore fewer mobile phone antennas, and where short trips are not well covered. This means that whatever the type of data collection procedure — Computer-Assisted Telephone Interview (CATI), Paper and Pencil Personal Interview (PAPI), Computer Assisted Web Interview (CAWI) or Computer Assisted Self Interview (CASI) — there is no real best method in terms of the validity of the results (Kalfs 1995). It may be more important to improve each type of procedure in order to combine them more effectively in their application (De Leeuv 2005). This could bring together the socio-demographic characteristics of the people interviewed, while associating them with a map-based questionnaire and/or GPS location via mobile phones (with active or passive use). These studies are more research oriented, with high costs; however, they remain a source of inspiration for future work (e.g. Chaix et al. 2019; Cuignet et al. 2020).

These approaches would be even more complex in cross-border territories, with the need to collaborate with different national mobile phone operators. Some recent experiences already exist in the cross-border region of Luxembourg or in Northern Europe in combining a big data approach, cross-border mobility patterns and social media data (e.g. Järv et al. 2022; Aagesen et al. 2022). The design of a common and interoperable cross-border mobile data management and analysis platform could be a major step forwards, fed with data from the main national operators. However, they need to be willing to join such an initiative and share comparable datasets, capable of

being harmonized. This would ensure a continuous or pseudo-continuous update of the traffic models.

Conclusion

This paper has presented a step-by-step methodology for constructing a pooled travel survey to support transport research and spatial planning in cross-border regions. Based on an applied research project aimed at building a transport model for the cross-border region of Luxembourg, we have demonstrated the operationalization of the pooled survey in this specific context. First, related to research question 1 (see Sect. "Introduction and objectives"), we have shown that the eight initial surveys we used, share enough concepts to identify a common ontology — including households, individuals and reported trips — that is rich enough to underpin a pooled survey. In addition, a pooled survey was successfully built, including information retained for 40 variables and 57,154 individuals. This is enough to feed a transport model and demonstrates the feasibility of the pooled travel survey, while providing a ready-to-use yet flexible methodology. Second, with regard to research question 2, we have shown that the harmonized indicators provide values that are meaningful with respect to the other cross-border regions, as well as to the specificities of the Luxembourg region. This critical discussion supports the validity of the pooled travel survey as a source of indicators of mobility and travel behaviours. Third, for research question 3, we have shown that the values for the proportions of purposes for all trips ending in selected urban districts computed with the pooled survey exhibit substantial differences from the values obtained using only local travel surveys (up to 6 percentage points, and up to 33 per cent in the total number of trips). This illustrates the added value of a pooled travel survey compared with an unharmonized collection of local travel surveys.

Ethical approval

All participants gave their informed consent to take part, worded in line with national legal authorities' guidelines.

Conflict of interest

The authors declare that they have no conflict of interest.

Appendix

Table 8 Comparison of the sub-populations used in the different local travel surveys. In each local travel survey, different information has been collected about individuals depending on the subpopulation they belong to

Survey	Sub-populations	Gathered Information
BELDAM	*All the people who live, even partially, in the same dwelling and share the same budget*	Census of individual characteristics
	…below 6 years old	Census of individual characteristics Travel survey
	…from 6 years old	Census of individual characteristics Travel survey Opinion survey
	…from age 16 years old	Census of individual characteristics Travel survey Opinion survey
EMF	*Cross-border workers working in Luxembourg*	Census of individual characteristics Travel survey
	Other persons living in the household, except children who live only occasionally in the household (weekend, vacations)	Census of individual characteristics
	Subsample of cross-border workers working in Luxembourg	Census of individual characteristics Travel survey Opinion Survey

Table 8 (continued)

Survey	Sub-populations	Gathered Information
EDGT SUD54 and EDGT Metz	*Individuals under 5 years old*	Type of individual, gender and relationship
	1) Individuals who live in this dwelling most of the year, including:	Census of individual characteristics
	-people who are temporarily away (vacation, business trip, hospitalization of less than one month, etc.), Travel survey	
	-infants, even if they are still in the maternity ward	
	-Subtenants and roommates occupying a part of the dwelling	
	OR	
	2) Persons of legal age who are living in this dwelling for their studies	
	OR	
	3) Domestic workers, employees and au pairs who live in this dwelling	
	Either1) Minor children housed elsewhere for their education and for whom this dwelling is the family IF the person is present:	
	residence	Census of individual characteristics
	2) Spouses who have another home for work reasons and who return to live in this dwelling for the Travel survey	
	weekend	
	Either	Census of individual characteristics
	1) Minor children who live in this dwelling for school and whose parents reside elsewhere	Travel survey
	2) Spouses who live in this dwelling for work reasons and return to their family dwelling on weekends	
	1 individual drawn per household, among individuals ≥ 16 years old	Opinion Survey
EDVM Thionville and EDVM Nord 54	*All individuals usually living in the dwelling at least 3 nights per week from Monday to Friday*	Census of individual characteristics
	1 or 2 individuals drawn per household, from individuals ≥ 11 years old	Census of individual characteristics
		Travel survey

Table 9 Categorization of the level of education. Original categories have been translated by the authors. CAP stands for the French « certificat d'aptitude préfessionnelle ». BEP stands for the French « Brevet d'étude professionnelle »

Id	Pooled survey items	Corresponding categories in the local travel surveys			
		BELDAM	EMF	LUXMOBIL	SIMO
1	**No degree**	1 Not relevant (<12 years old) 2 No degree	*	0 <6 years old 1 In school (pupil) 6 No degree	00 In school (pupil) 09 No degree
2	**Elementary**	3 Elementary	1 Elementary	2 8 years old	01 Elementary
3	**Secondary**	4 General secondary 5 Technical secondary 6 Professional secondary 7 Special education	21 General or technical secondary 22 Professional secondary (CAP/BEP) 31 Professional *baccalauréat* 32 General or technical *baccalauréat*	3 University entrance qualification	02 Secondary (from 6th to 3rd, CAP) 03 Secondary (from 2nd to last year, BEP), without *baccalauréat* 04 Secondary, with *baccalauréat* 07 Apprenticeship (elementary or secondary schools only)
4	**Higher education**	8 College (2–3 years) 9 College (4–5 years) 10 University	4 Short program in higher education 5 Long program in higher education	4 Master	05 Higher education, until 2 years 06 Higher education, at least 3 years 08 Apprenticeship (higher education)
5	**Other**		6 Other training or degree (without precision)	5 Other	
na	**No answer**		7 No answer	7 No answer	

* Not included in the EMF's population of interest

Table 10 Categorization of activity status

Id	Name	Corresponding categories in the local travel surveys			
		BELDAM	EMF	LUXMOBIL	SIMO
1	**Work at least one full time**	Work AND do not study AND work at least 40 h a week OR work full-time	Work AND work full-time	1 Work full-time	1 Work full-time
2	**Work less than a full time**	Work AND do not study AND work less than 40 h a week OR do not work full-time	Work AND do not work full-time	2 Work part-time	2 Work part-time
3	**study**	Do not work AND study	*	4 Student below 18	4 Student with *baccalauréat* 5 Student without *baccalauréat*
4	**Work and study**	Work AND study	Continuous training	5 Students above 18 and apprenticeship	3 Work-linked training, internship
5	**Unemployed**	Do not work nor study AND looking for a job	*	8 currently unemployed	6 Unemployed
6	**Other inactives**	Do not work nor study AND not looking for a job	*	0 less than 6 years old 3 Pensioner(s)/retiree(s) 6 On parental leave 7 Housewife/househusband	7 Retired 8 Housewife/househusband 9 Other
na	**No answer**	No answer OR all other potential cases	All other potential cases	9 Other	

* Not included in the EMF's population of interest

Table 11 Shares of trip purposes and total number of trips for all travel starting in the selected areas

	Arlon			Luxembourg			Thionville-Yutz		
	local survey	pooled survey	diff	local survey	pooled survey	diff	local survey	pooled survey	diff
Drop off or pick up someone	13.82	12.22	*-1.60*	5.11	5.96	*0.84*	10.23	8.47	*-1.76*
Go home	35.77	36.68	*0.92*	51.00	50.22	*-0.78*	38.48	38.78	*0.30*
Go to work	13.92	18.02	*4.10*	16.32	13.97	*-2.35*	12.88	14.35	*1.47*
Travel for work duties	0.56	0.98	*0.42*	3.00	2.91	*-0.10*	0.26	0.28	*0.02*
Taking a course	7.65	4.98	*-2.67*	3.35	3.94	*0.60*	4.48	4.77	*0.30*
have a meal out	0.88	0.98	*0.09*	4.00	4.09	*0.09*	1.90	2.17	*0.27*
Grocery and shopping	9.85	9.12	*-0.73*	2.86	3.24	*0.38*	15.21	14.87	*-0.34*
Personal services (doctor, bank, etc.)	3.15	3.15	*0.00*	3.60	4.27	*0.67*	3.80	3.69	*-0.11*
visit family or friends	5.26	4.57	*-0.69*	2.29	2.60	*0.32*	4.65	4.82	*0.17*
Going for a walk, going for a ride	2.72	2.64	*-0.09*	0.00	0.00	*0.00*	4.78	4.29	*-0.49*
leisure, sports, culture	3.02	3.23	*0.21*	4.52	4.71	*0.19*	3.04	2.90	*-0.14*
Other	3.39	3.44	*0.04*	3.95	4.08	*0.13*	0.29	0.61	*0.32*
Total number of trips	**85,802**	**114,487**	**28,685 (+33%)**	**429,414**	**472,817**	**43,403 (+10%)**	**210,095**	**232,911**	**22,816 (+11%)**

'Local' or 'pooled' activities surveys are shown in %, and 'diff.' stands for difference in percentage points. All trips starting in any of the three selected urban areas were considered. However, in the local surveys, these trips comprise those made by local residents only, while in the pooled survey, trips of cross-border travellers are also included

🖄 Springer

Data availability The micro-data falls under the pseudonymized personal data category in the European General Data Protection Regulation. Therefore, the micro-data cannot be released as open data. However, the data is available at a certain aggregated level from the authors upon reasonable request and with the permission of the MMUST project consortium.

Declarations

Competing interests The authors declare no competing interests.

References

Aagesen Wallin, H., Järv, O., Gerber, P.: The effect of COVID-19 on cross-border mobilities of people and functional border regions: the Nordic case study from Twitter data. Geogr Ann Ser B Hum Geogr **105**(4), 356–378 (2022). https://doi.org/10.1080/04353684.2022.2101135

Armoogum, J., Bonsall, P., Browne, M., Christensen, L., Cools, M., Cornélis, E., Diana, M., Guilloux, T., Harder, H., Reinau, K.H. (eds.): Survey Harmonisation with New Technologies Improvement (SHANTI). COST Office, Brussels, Belgium (2014)

Aschauer, F., Hössinger, R., Axhausen, K.W., Schmid, B., Gerike, R.: Implications of survey methods on travel and non-travel activities: A comparison of the Austrian national travel survey and an innovative mobility-activity-expenditure diary (MAED). Eur. J. Transp. Infrastruct. Res. **18**(1), 4–35 (2018). https://doi.org/10.18757/ejtir.2018.18.1.3217

Behren, S.V., Minster, C., Esch, J., Hunecke, M., Vortisch, P., Chlond, B.: Assessing car dependence: Development of a comprehensive survey approach based on the concept of a travel skeleton. Transp. Res. Procedia **32**, 607–616 (2018). https://doi.org/10.1016/j.trpro.2018.10.015

Ben-Elia, E., Lyons, G., Mokhtarian, P.L.: Epilogue: The new frontiers of behavioral research on the interrelationships between ICT, activities. Time Use Mob. Transp. **45**(2), 479–497 (2018). https://doi.org/10.1007/s11116-018-9871-x

Bonnel, P., Hombourger, É., Olteanu-Raimond, A., Smoreda, Z.: Apports et limites des données passives de la téléphonie mobile pour la construction de matrices origine-destination. Rev. Econ. Rég. Urbaine **4**, 647–672 (2017). https://doi.org/10.3917/reru.174.0647

Brunet-Jailly, E.: Theorizing borders: An interdisciplinary perspective. Geopolitics **10**(4), 633–649 (2005)

CEREMA.: Les enquêtes mobilité certifiées Cerema (EMC2) - Principes méthodologiques. CEREMA, Bron (France) (2020) https://www.cerema.fr/fr/centre-ressources/boutique/enquetes-mobilite-certifiees-cerema-emc2

Chaix, B., Benmarhnia, T., Kestens, Y., Brondeel, R., Perchoux, C., Gerber, P., Duncan, D.T.: Combining sensor tracking with a GPS-based mobility survey to better measure physical activity in trips:

public transport generates walking. Int. J. Behav. Nutr. Phys. Act. **16**(84), 1–13 (2019). https://doi.org/10.1186/s12966-019-0841-2

Cuignet, T., Perchoux, C., Caruso, G., Klein, O., Klein, S., Chaix, B., Kestens, Y., Gerber, P.: Mobility among older adults: Deconstructing the effects of motility and movement on wellbeing. Urban Stud. **57**(2), 383–401 (2020). https://doi.org/10.1177/0042098019852033

Chilla, T., Heugel, A.: Cross-border commuting dynamics: Patterns and driving forces in the Alpine Macro-region. J. Borderlands Stud. (2019). https://doi.org/10.1080/08865655.2019.1700822

Cornelis E. (eds): La mobilité en Belgique en 2010 : résultats de l'enquête BELDAM, Belspo et SPF Mobilité & Transports, Bruxelles, Belgium. Available (2024.05.02) at: https://mobilit.belgium.be/sites/default/files/downloads/Rapport_final_beldamfr.pdf

Decoville, A., Durand, F., Sohn, C., Walther, O.: Comparing cross-border metropolitan integration in Europe: towards a functional typology. J. Borderl. Stud. **28**(2), 221–237 (2013). https://doi.org/10.1080/08865655.2013.854654

Decoville, A., Durand, F.: An empirical approach to cross-border spatial planning initiatives in Europe. Reg. Stud. **55**(8), 1417–1428 (2021). https://doi.org/10.1080/00343404.2021.1902492

Diederich, R. : Accessibilité et flux à Luxembourg-Ville; simulation plurimodale de la circulation motorisée, PhD Thesis, Strasbourg (France), 620 p. (1998)

Dörry, S., Decoville, A.: Governance and transportation policy networks in the cross-border Metropolitan Region of Luxembourg: A social network analysis. Eur. Urban Reg. Stud. **23**(1), 69–85 (2016). https://doi.org/10.1177/0969776413490528

Drevon, G., Gerber, P. Klein, O. & Enaux, C.: Measuring functional integration by identifying the trip chains and the profiles of cross-border workers: Empirical evidences from Luxembourg. J. Borderlands Stud. **33**(4), 549–68 (2018). https://www.tandfonline.com/action/journalInformation?journalCode=rjbs20

Durand, F., & Nelles J. Binding cross-border Regions: An analysis of cross-border governance in Lille-Kortrijk-Tournai Eurometropolis. Tijdschr. Econ. Soc. Geogr. **105**(5), 573–90 (2014). http://doi.wiley.com/https://doi.org/10.1111/tesg.12063.

Dypvik Landmark, A., Arnesen, P., Södersten, C.J., et al.: Mobile phone data in transportation research: methods for benchmarking against other data sources. Transportation **48**, 2883–2905 (2021). https://doi.org/10.1007/s11116-020-10151-7

Enaux, C., Gerber, P.: Beliefs about energy, a factor in daily ecological mobility? J. Transp. Geogr. **41**, 154–162 (2014). https://doi.org/10.1016/j.jtrangeo.2014.09.002

Eurostat Task Force on Passenger Mobility. EU transport statistics. Eurostat guidelines on passenger mobility statistics, Luxembourg. Eurostat, Luxembourg (2018)

Eurostat.: Commuting between regions. https://ec.europa.eu/eurostat/web/products-Storper-news/-/ddn-20210610-1 (2021). Accessed 22 September 2023

Fries-Tersch, E., Bradley, H., Tugran, T.: 2016 Annual report on intra-EU labour mobility. http://ec.europa.eu/social/main.jsp?catId=738&langId=en&pubId=7981&type=2&furtherPubs=yes. (2017)

Gerber, P.: Navetter à la frontière. In: Amilhat-Szary A.-L., Hamez G. (eds.) Frontières. Armand Colin, Paris (France) (2020)

Hamez, G.: Du transfrontalier au transnational: Approche géographique. L'exemple de La Frontière Franco-Belge. Université Paris 1 Panthéon Sorbonne. https://tel.archives-ouvertes.fr/tel-00007191 (2004)

Hubrich, S., Wittwer, R., Gerike, R.: Quality indicator set for household travel surveys. Transp. Res. Procedia **33**, 219–226 (2018). https://doi.org/10.1016/j.trpro.2018.10.098

Järv, O., Aagesen, H.W., Väisänen, T., Massinen, S.: Revealing mobilities of people to understand cross-border regions: Insights from Luxembourg using social media data. Eur. Plan. Stud. (2022). https://doi.org/10.1080/09654313.2022.2108312

Kalfs, N.: Effects of different data collection procedures in time use research. Transp. Res. Rec. **1493**, 110–117 (1995)

Leeuw, E.D.: To mix or not to mix data collection modes in surveys. Journal of Official Statistics **21**(2), 233–255 (2005)

Liang, L., Xu, M., Grant-Muller, S., Mussone, L.: Household travel mode choice estimation with large-scale data—an empirical analysis based on mobility data in Milan. Int. J. Sustain. Transp. **15**(1), 70–85 (2020). https://doi.org/10.1080/15568318.2019.1686782

Pires, I., Nunes, F.: Labour mobility in the Euroregion Galicia-Norte de Portugal: Constraints faced by cross-border commuters. Eur. Plan. Stud. **26**(2), 376–395 (2018). https://doi.org/10.1080/09654313.2017.1404968

Raballand, W., Le Corre, M.: Pertinence des données de téléphonie mobile pour les matrices origine-destination. 45e Congrès ATEC ITS France - Les Rencontres de la Mobilité Intelligente, Montrouge, France (2018). https://hal.archives-ouvertes.fr/hal-02284096

Rebecq, A.: Package 'icarus' (2019). https://cran.r-project.org/web/packages/icarus/icarus.pdf

Schiebel, J., Omrani H., Gerber P.: Border effects on the travel mode choice of resident and cross- border workers in Luxembourg. Eur. J. Transp. Infrastruct. Res. **15**(4), 570–96 (2015). https://doi.org/10.18757/ejtir.2015.15.4.3098

Schmitz, F., Drevon G. Gerber P. (eds.) La mobilité des frontaliers du Luxembourg : Dynamiques et perspectives. CEPS/INSTEAD, Esch-sur-Alzette, LU (2012)

Sohn, C.: Modelling cross-border integration: The role of borders as a resource. Geopolitics **19**(3), 587–608 (2014). https://doi.org/10.1080/14650045.2014.913029

Thomas, I., Jones, J., Caruso, G., Gerber, P.: City delineation in European applications of LUTI models: Review and tests. Transp. Rev. **38**(1), 6–32 (2018). https://doi.org/10.1080/01441647.2017.1295112

Wang, Z., He, S.Y., Leung, Y.: Applying mobile phone data to travel behaviour research: A literature review. Travel Behav. Soc. **11**, 141–155 (2018). https://doi.org/10.1016/j.tbs.2017.02.005

Wittwer, R., Gerike, R., Hubrich S.: Peak-car phenomenon revisited for urban areas: Microdata analysis of household travel surveys from five European capital cities. Transp. Res. Rec.: J. Transp.Res. Board **26**73(3), 686–99. http://journals.sagepub.com/doi/10.1177/0361198119835509

Justin Delloye Geographer by training, is a regional scientist studying the dynamic interplay between land use, migration, and transport systems, both at urban and regional scales. He is committed to a multiscale, individual-based approach to complex spatial systems that integrates mathematical, statistical, and stochastic modeling methods. He is currently working at Statbel, the Belgian Statistical Office, and he also teaches Economic Geography at the Université catholique de Louvain.

Sylvain Klein Senior Research Analyst at LISER. He trained as a computer scientist and has over 20 year's experience in processing geographical data applied to mobility issues in Luxembourg and its cross-border context. He is an active member of the MMUST cross border traffic model. Other research of him focuses on the relationship between people's mobility patterns, their health and environmental variables. He uses a variety of data sources such as questionnaires, map-based surveys and sensors.

Olivier Klein An urban geographer working as Research Scientist at Luxembourg Institute of Socio-Economic Research (LISER). His current research interests concern the study of daily mobility addressed through spatiotemporal approaches with a special interest for daily mobility of workers, particularly in cross-border contexts. He is also interested in issues related to GIS sciences, active mobility in link with physical activity and health, modelling and visualization.

Guillaume Drevon Scientist at Urban Sociology Lab of Swiss Federal Institute of Technology, Lausanne. Guillaume DREVON is also Research Associate at Luxembourg Institute of Socio-Economic Research. He is mainly interested in daily rhythms, in particular the strategies developed by individuals to cope with chronic-time pressures. He also has strong expertise in survey design and the application of related analytical methods. In doing so, he contributes to a better understanding of contemporary rhythms and social acceleration impacts by developing the concept of rhythmology.

Mathieu Jacquot Project manager for travel knowledge and modelling at Cerema (France). He has 20 year's experience in the field, during which time he has carried out a wide variety of assessments of public transport policies (road and public transport projects, regulation and taxation of HGV traffic, development of new commercial or business zones, etc.). Currently, his work focuses on two main subjects: analysing cross-border mobility within the "Grande Région" and the study of the effects of changes in cycling on mobility.

Eric Cornélis Mathematical sciences, is working in mobility researches since more than thirty years. His interests are both in modeling and in mobility surveys. He is senior research associate and lecturer at the University of Namur (UNamur). He is leading the Transport Research Group which is part of naXys, the

Namur Institute for Complex Systems. He is also vice-president of BIVEC-GIBET (Benelux Interuniversity Association of Transport Researchers), member of the Belgian high council for Statistics and member of the Belgian central council for Economy.

Joanna Rousseaux Holds a Master's degree in Transport and Logistics, as well as in Urban Planning and Real Estate Management. She has been involved in mobility research for over a decade. Currently, she serves as a Project Manager for Brussels Mobility. Previously, she worked as a Research Associate at the University of Namur, focusing on modeling mobility behaviors. She also worked at the Ministry of Infrastructure and Construction in Poland, contributing to the modernization of railway infrastructure.

Julien Schiebel A geographer with main interests in mobility strategy, especially with a focus on transportation modelling. He is working at the Ministry of Mobility and Public Works of Luxembourg since 2018 where he is in charge of both analyzing and disseminating personal and non-personal mobility data in the Mobility Digital Observatory and assessing the impact of different transportation and land use options to identify how the transport system in Luxembourg is likely to perform in the future.

Stéphane Godefroy A doctorate in geography and is mainly interested in mobility and its links with spatial planning. He has been working at the Agence d'urbanisme Lorraine Nord - AGAPE since 2011. As part of this role, he oversees and contributes to the drafting of various planning documents, carries out mobility studies and coordinates European projects.

Philippe Gerber A geographer with main interests in land use and transport interaction (LUTI), especially linking interactions between residential and daily mobility in a sustainable and wellbeing perspective. He is using spatial analysis, econometric tools and mixed methods dedicated to survey design and working at the Luxembourg Institute of Socio-Economic Research – LISER – since 2001. He is also a lecturer at the University of Luxembourg and was visiting researcher at McGill and Montreal Universities (Canada) in 2018-2019.