

Non-Probability vs. Address-Based Sampling during SARS-CoV-2

Findings from Two Travel Surveys by Smartphone App in Dresden, Germany

Johannes Weber^{*a}, Stefan Hubrich^a, Rico Wittwer^a, Regine Gerike^a

12th International Conference on Transport Survey Methods, March 2022, Porto Novo, Portugal

Same Survey Mode — Different Sampling Methods

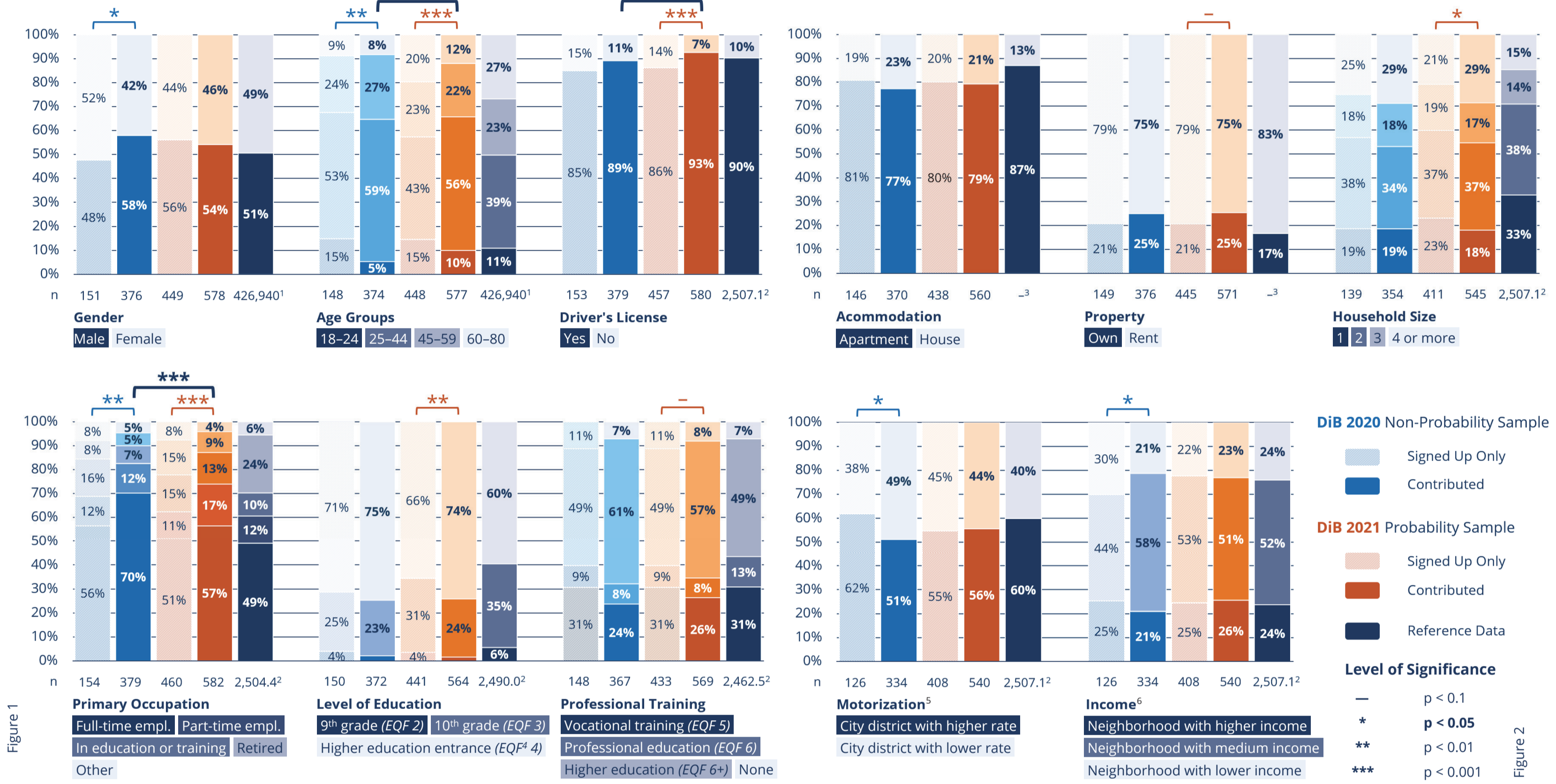
Non-probability sampling (NPS) for app-based travel surveys has potential compared to random address-based sampling (ABS): samples can be obtained more targeted, cost-effectively and with higher willingness to participate (KUHNNIMHOF et al., 2018).

On the other hand, they may be prone to systematic losses, selectivity problems and bias (see e.g., ERIKSSON et al., 2018; ZEGRAS et al., 2018; WEBER et al., 2020.; FAGHIH IMANI et al. 2020).

This study evaluates its effects in NPS and ABS in-depth by comparing two travel surveys in the city of Dresden, branded “Dresden in Bewegung” (DIB, in English: “Dresden in motion”). Using the same app, they were dedicated to travel behavior during pandemic restrictions and their reliefs — but differed in sampling method. Sample structures (incl. those signing up only vs. those contributing), impacts on travel estimates and contrasts to individuals’ stated affectedness by SARS-CoV-2 were analyzed:

Characteristics	Dresden in Bewegung 2020 Non-Probability Sample (NPS)	Dresden in Bewegung 2021 Probability Sample (ABS)
Sample Selection/Recruitment	Direct (private/workplace environments) and indirect recruitment (news, social media, e-mail circulars/newsletters)	Simple random sample from the city’s register of residents
Data Collection Period	4 May–30 June	17 May–18 July
Net Sample	379 persons	582 persons
Contributed Days (per Person)	7,013 (18.5)	8,807 (15.1)
Trips on Contributed Days	30,083	37,713

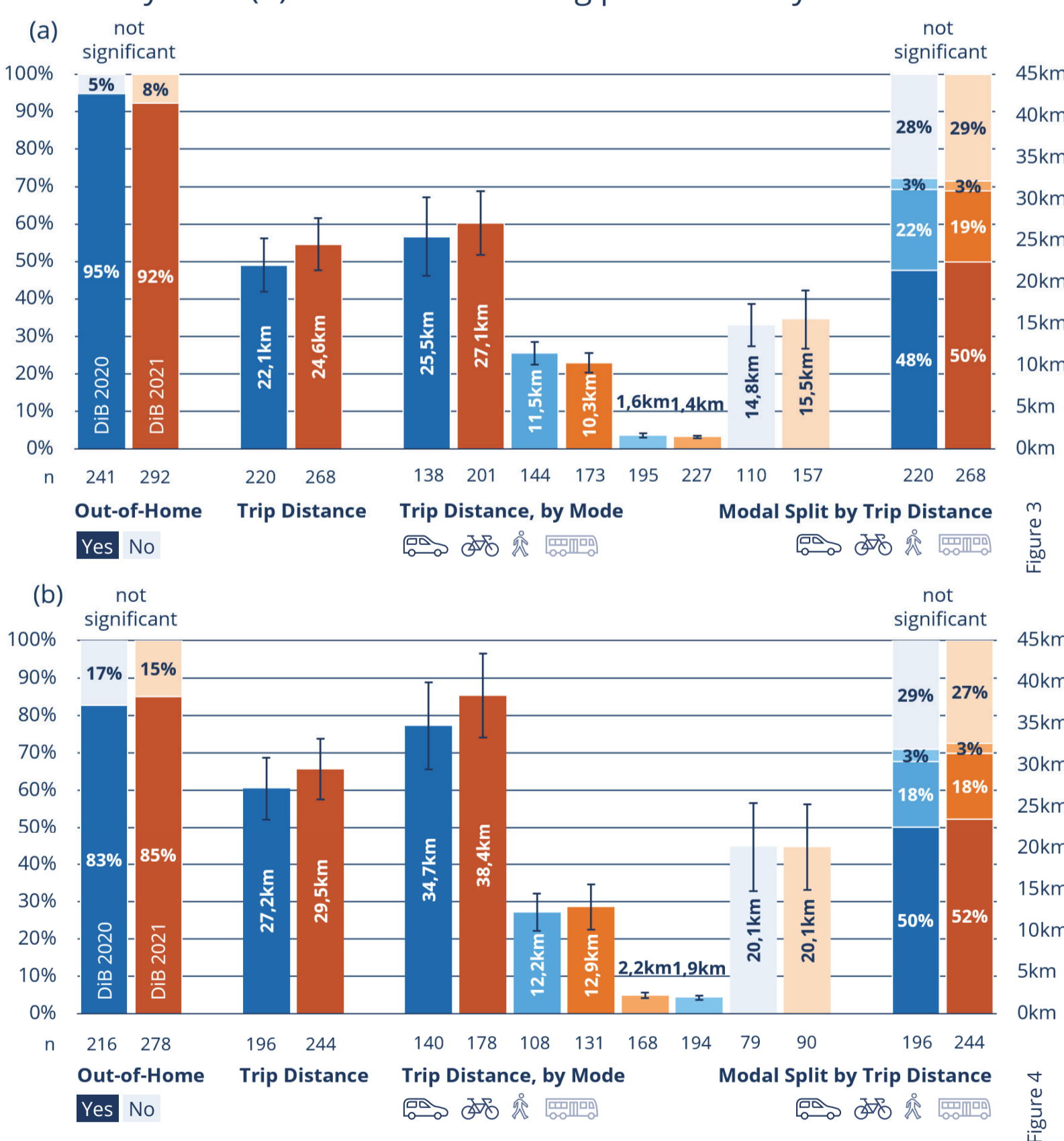
Worth Taking a Closer Look — Similar Sample Structures



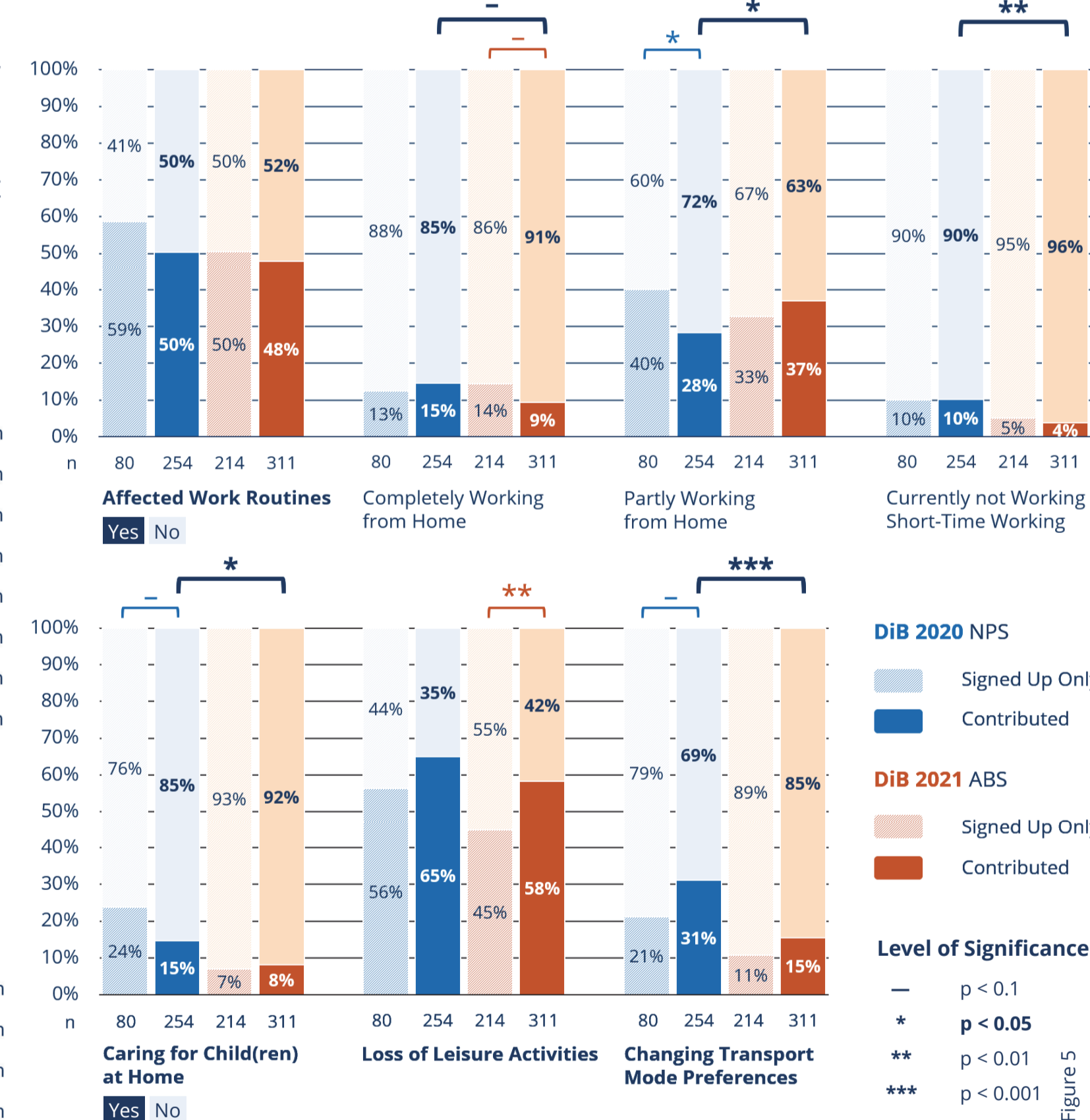
Focusing on Travel Estimates of the Employed (18–59 years)

Emerging from the net samples above, main differences exist in primary occupation and age distribution. To exclude these major influence factors and to check if NPS is feasible for studying specific subgroups of the population, net participants are filtered by full-time employment and an age range (18–59). In this subset, no significant differences concerning sample structures were found any more.

The question rises if the sampling methods lead to differences in collected travel behavior, which has been analyzed for (a) average workdays and (b) weekends including public holidays:



Contrasting with Stated Affectedness by SARS-CoV-2



Both, the travel estimates on (a) average workdays and (b) weekend days incl. public holidays do not show any significant differences as well. Hence, the data reflect travel behavior without considerable deviants and regardless of the kind of days evaluated (Figure 3 & 4). To better understand if indeed no significant behavioral differences in the selected subgroups can be distinguished or if contrary effects might be involved, the individuals’ stated affectedness by the SARS-CoV-2 crisis was analyzed — showing specific differences (Figure 5).

Conclusion

While NPS benefited from targeted ad-hoc recruitment yielding higher contribution rates and willingness to stay in the survey, ABS obtained randomly drawn residents from register sticking to representativeness in terms of probability theory.

Remarkably similar sample structures could be revealed according to background variables. This deviates from research findings in literature. Some differences could be identified with respect to survey break-off between participants that signed up only and those contributed with travel data. These did not result in serious deviations due to sample losses.

However, significant differences across the surveys appeared according to age distribution and primary occupation. Focusing only on full-time employed persons (18–59 yrs.) as the largest subgroup possible to select, no significant differences for any of the considered variables could be identified, both in the subsample structures and in the travel estimates. This is the key finding of this study.

Concerning stated affectedness by the pandemic in everyday life, differences occur in more detail. This is not surprising as the respective pandemic situation and its subjective perception of effects may have varied between 2020 and 2021, despite similar relief phases of pandemic restrictions. Based on the key travel estimates used above, the GPS-based measurement of travel behavior was not affected.

To conclude, applying NPS recruitment among full-time employed target groups is a viable way, particularly in the case of disruptive events, e.g. a pandemic context. Benefits of NPS may outweigh challenges of ABS in declining response rates and increasing costs and efforts in the future. Thus, for studying specific population subgroups only (e.g., young adults, middle-aged employees), NPS may be reasonable for specific research tasks. They might still not be a substitute for general population surveys, but could be complementary, particularly for hard-to-reach groups. However, further empirical research is needed regarding the effects of NPS recruitment strategies on an entire survey population — and how these samples can be processed to meet the requirements of core quality criterion of representativeness and reliable survey data.

Figures:

- Figure 1: Characteristics of the two app-based travel surveys in Dresden, Germany (individuals between 18 and 80 years)
- Figure 2: Survey sample characteristics, by participation level (individuals between 18 and 80 years)
- Figure 3: Key travel estimates (95% CI) on average workdays (Tue, Wed and Thu, no public holidays), full-time employed, 18–59 years
- Figure 4: Key travel estimates (95% CI) on weekends and public holidays, full-time employed, 18–59 years
- Figure 5: SARS-CoV-2 affectedness in everyday life, by participation level (full-time employed, 18–59 years)

Footnotes:

- Dresden register of residents 2020, 2021
- “Mobility in Cities – SrV” household travel survey 2018 in Dresden, Germany — weighted by gender, age, household size and city districts
- Dresden Housing Market Report 2018, data refers to entire housing stock
- EQF — European Qualifications Framework
- Individuals and respective city districts’ motorization, related to mean city value of 354 vehicles/1,000 people
- Individuals and respective neighborhoods’ equalized disposable income, related to mean city value of 1,800 €

References:

- ERIKSSON, J.; LINDBORG, E.; ADELL, E.; HOLMSTRÖM, A.; SILVANO, A. P.; NILSSON, A.; HENRIKSSON, P.; WIKLUND, M.; DAHLBERG, L. (2018): New ways of collecting individual travel information: Evaluation of data collection and recruitment methods.
- FAGHIH IMANI, A.; HADDING, C.; SRIKUNTHIRAN, S.; MILLER, E. J.; NURUL HABIB, K. (2020): Lessons from a Large-Scale Experiment on the Use of Smartphone Apps to Collect Travel Diary Data: The “City Logger” for the Greater Golden Horseshoe Area. *Transportation Research Record: Journal of the Transportation Research Board*.
- KUHNNIMHOF, T.; BRADLEY, M.; ANDERSON, R. S. (2018): Workshop Synthesis: Making the transition to new methods for travel survey sampling and data retrieval. *Transportation Research Procedia*.
- WEBER, J.; HUBRICH, S.; WITTEW, R.; GERIKE, R. (2020): Non-probability recruitment strategies for innovative smartphone-based travel surveys. *Internationales Verkehrswesen*.
- ZEGRAS, P. C.; LI, M.; KILIC, T.; LOZANO-GRACIA, N.; GHORPADE, A.; TIBERTI, M.; AGUILERA, A. I.; ZHAO, F. (2018): Assessing the representativeness of a smartphone-based household travel survey in Dar es Salaam, Tanzania. *Transportation*.