Active commuting has been widely identified as beneficial, both through direct association of physical activity and health and through a wider linkage of commute mode to satisfaction and well-being (Smith, 2017). However, many commutes are multi-modal, and to ignore walking or cycling trips as the first or last stage of a commute is to ignore a significant amount of activity.

Commuting data is collected in the 'long form' variant of the census of Canada in a similar manner to that used by many other countries – which is to say, it asks for one mode of transport, and does not distinguish between outward and return journeys. The 2016 Census added some additional questions about time taken and time of departure, as well as vehicle occupancy. The data produced are thus typical of census journey to work data – on the one hand censuses can furnish larger volumes of data than surveys, but on other hand they are much less detailed than dedicated travel surveys. This paper is an interesting attempt to look at how we can bridge these two sources, in order to estimate walking as an element of other modes.

An underlying characteristic of journey to work data, when presented as residence-workplace matrices, is that they are very sparse, leading to concerns about disclosure and the need to strike a balance between ease of access to data and utility (Duke-Williams, 2019). The need to limit the number of questions in a census means that there is no scope to look at trip-chaining for multiple purposes (as discussed by Criado Perez (2019), a feature of travel that differs for men and for women). Due to the sparsity of journey to work data, disclosure thresholds are much more problematic than they are in typical aggregate data, regardless of whether taken from censuses or surveys. The authors of this paper use census data with a disclosure threshold of ten. Analysis of 2011 Census data from England and Wales (Special Workplace Statistics Table WU03BEW) using a relatively fine spatial resolution (MSOA Level), shows that around 17% of all journeys would be excluded with a publication threshold of ten. More worryingly, the exclusion level would vary by transport mode, from 6% for 'on foot' to 32% for 'train'. This effect is mostly related to relatively long distance journeys – hence the low level for pedestrians. The analysis in this paper is limited to a single (though large) urban area and is unlikely to be affected to the same degree, but it may be relevant to note that the transport mode that is least well modelled (Table 2) is train. The authors do not dwell on the possible differential effects by mode of transport of the confidentiality threshold used.

The processing of the census and travel survey data to estimate walking distance from home to public transport and from public transport to work is interesting, and highlights the types of analysis that are now possible, especially given increasing access to open data catalogues of bus stop locations etc. and of actual transport performance. It should be borne in mind of course, that each individual will choose their own route, which might not necessarily be the most obvious. Analysis using WiFi connection logs by Transport for London of individuals' routes through their network (TfL, 2017) showed that for one sample pair of stations, at least 18 different routes were used between them. We might similarly expect that people will not necessarily take the most obvious walking route, and estimates of walking distance based on rationality will be the lower limit; actual walking distance may be longer, as might trips by any other mode (Dalton, Jones, Panter, & Ogilvie, 2015). Recognising that commutes might involve walking at both ends of the journey might also be a useful qualifier to the modelling of walkability (Dannenberg, Kraft, & Alvanides, 2017) – it may be that we need to simultaneously consider walkability near workplaces (Adams, Bull, & Foster, 2016) in conjunction with walkability close to the home in conjunction with overall flow pattern analysis better understand individual choices.

Relatively easy adjustments to census questionnaires could be and are made that would deliver richer data (at the expense of greater noise, perhaps) that would facilitate future work of the sort

described here, whether in a Canadian context or in other settings. Rather than asking about just one mode of transport in censuses, we could ask respondents to identify multiple modes (as done for example in Australia and in Switzerland, the latter also asking for the time spent on each mode) and then attempt to model the likely split between modes. Greater use could be made of intra-household information on travel. In Scotland and Northern Ireland (but not the rest of the UK), the 'place of work' and 'method of travel' census questions are modified to also refer to place of study for school-children and students. Whilst this introduces ambiguity for the (non-trivial number of) students who both study and work (as only one location can be identified), it creates a wealth of data, and allows potential analysis of transport at a household level of cases wherein children are car-passengers to school.

In many countries, there is interest in full or partial replacement of censuses with administrative data sources. This may be problematic for journey to work (and to place of study) data currently generated by these censuses, which would restrict the future ability to continue to carry out the sort of analysis described by the authors. Whilst administrative sources will have a good record of each individual's home address (although this might not be the point of departure for the journey to work, something recognised by a number of censuses that ask for the address from which people leave for work, if it is different to their usual residence), the workplace may or may not be accurately know: it may be the case that admin sources only record the stated address of the employer. More pressingly they are unlikely to know much about the method of travel to work, which can potentially vary from day to day, especially in urban environments where there might be a number of alternative routes and types of transport. There is therefore a note of caution to be had in looking at how studies like this might be taken forward; we should not assume that current census data sets will always be produced in the same way as now. This is in contrast to the optimism we might have through greater publication of open data on use of public transport.

The consideration of how far people will walk to public transport is interesting. Clearly, it depends on infrastructure density and population distribution, but also on personal choice. Studies of children walking to school (Chillón, Panter, Corder, Jones, & Van Sluijs, 2015; Rodríguez-López et al., 2017) suggest higher thresholds exist for children and young people than suggested in this paper; thus adults might indeed be encouraged to walk further than we currently suspect (Durand et al., 2016).

Bibliography

Adams, E. J., Bull, F. C., & Foster, C. E. (2016). Are perceptions of the environment in the workplace 'neighbourhood' associated with commuter walking? *Journal of Transport & Health*, *3*(4), 479–484. https://doi.org/10.1016/j.jth.2016.01.001

Chillón, P., Panter, J., Corder, K., Jones, A. P., & Van Sluijs, E. M. F. (2015). A longitudinal study of the distance that young people walk to school. *Health & Place*, *31*, 133–137. https://doi.org/10.1016/j.healthplace.2014.10.013

Criado-Perez, C. (2019). *Invisible women: Exposing data bias in a world designed for men*.

Dalton, A. M., Jones, A. P., Panter, J., & Ogilvie, D. (2015). Are GIS-modelled routes a useful proxy for the actual routes followed by commuters? *Journal of Transport & Health*, *2*(2), 219–229. https://doi.org/10.1016/j.jth.2014.10.001

Dannenberg, A. L., Kraft, K., & Alvanides, S. (2017). Special Issue Editorial: Tools and practices for understanding and promoting walking and walkability. *Journal of Transport & Health*, *5*, 1–4. https://doi.org/10.1016/j.jth.2017.05.365

Duke-Williams, O. (2019). Location Tracing and Potential Risks in Interaction Data Sets. *Applied Spatial Analysis and Policy*, *12*(2), 377–394. https://doi.org/10.1007/s12061-017-9247-1

Durand, C. P., Tang, X., Gabriel, K. P., Sener, I. N., Oluyomi, A. O., Knell, G., … Kohl, H. W. (2016). The association of trip distance with walking to reach public transit: Data from the California Household Travel Survey. *Journal of Transport & Health*, *3*(2), 154–160. https://doi.org/10.1016/j.jth.2015.08.007

Rodríguez-López, C., Salas-Fariña, Z. M., Villa-González, E., Borges-Cosic, M., Herrador-Colmenero, M., Medina-Casaubón, J., … Chillón, P. (2017). The Threshold Distance Associated With Walking From Home to School. *Health Education & Behavior*, *44*(6), 857–866. https://doi.org/10.1177/1090198116688429

Smith, O. (2017). Commute well-being differences by mode: Evidence from Portland, Oregon, USA. *Journal of Transport & Health*, *4*, 246–254. https://doi.org/10.1016/j.jth.2016.08.005