Mobility metrics for Glasgow City Region

UBDC 14/12/22

Aim

To produce small-area aggregate measures of mobility over time

Data

Mobile phone application data from **Huq[1] and Tamoco**[2].

Geography

Glasgow city-region with results broken down by **Intermediate Zone** (417 in the city-region).

Temporal detail

1st July 2019 – 31st December 2021 (2.5 years) with results aggregated **quarterly** (10 periods).

Measure

Mobility is measured following **Cuebiq Mobility Index**[3] approach:

- Mobility analysis is limited to mobile phone users determined to live within the Glasgow city-region (see Appendix 1 for details).
- For each mobile phone user with a home location, on each day:
    1. Subset available data points for the user if they are active. Mobility is only recorded when users are active in the dataset. If a user is not active on a given day, they are not assumed to have 0 mobility but given a NULL value for that day.
    2. Add a point for their home location so that is always present [effectively assumes they always travel from home each day, and don't stay over somewhere else].
    3. Draw a bounding box around the remaining observations [min x, min y; and max x, max y].
    4. Measure the diagonal distance (D) of the bounding box [using Euclidean distance – in metres].
    5. **Device Mobility (DM)** = $\log_{10}(D + 1)$ [addition of 1 to avoid log(0) probs].
- For each Intermediate Zone and quarter, the **Mobility Index (MI)** = weighted median of Device Mobility (DM).
    - The weight applied to each user is based on the Scottish Index of Multiple Deprivation quintile of their home Datazone and the Council in which they live (see Appendix 2 for details).
- Other measures recorded for each Intermediate Zone and quarter:
    - **Number of users (N)** – used in the calculation of median for zone/period [to allow if to include a minimum threshold for reporting]
    - **Number of Active days** – how many active days were there in the quarter, where each day can have multiple active users.
    - **MI_{mean}** = mean (DM)
    - **MI_{25}** = 25th percentile value (DM)
    - **MI_{75}** = 75th percentile value (DM)

[1] http://www.ubdc.ac.uk/data-services/data-catalogue/transport-and-mobility-data/huq-data/
[2] https://ubdc.ac.uk/data-services/data-catalogue/transport-and-mobility-data/tamoco-data/
[3] https://help.cuebiq.com/hc/en-us/articles/360041285051-Mobility-Insights-Mobility-Index-CMI-

Notes

All results are reported. Where user numbers or active days are low in reporting mobility for an intermediate zone in a given quarter, users of the results should consider employing a threshold cutoff depending on the analysis being undertaken to ensure the uncertainty of low user numbers and/or low active days do not skew results.

Unique user IDs in the Tamoco dataset are rehashed monthly but the results are aggregated quarterly. The number of distinct users in a quarter may therefore be slightly overestimate given the same users may be active over multiple months but with different user IDs.

Variables

Variable name	**Definition and/or commentary**
Name	Intermediate zone name
Code	Intermediate zone code
Quarter	Year and quarter of the year, E.g. 2019.Q1 begins in January 2019.
Users	The number of mobile phone users used in the calculation of MI [can be used to set a minimum threshold for analysis using the data]
Active_days	The number of active days used in calculating MI, where each day can have multiple active mobile phone users.
Median_weighted	Weighted median of mobility from mobile phone users active in an intermediate zone and a given quarter. See data description for details on calculation.
Lower25	25th percentile value of weighted mobility
Upper75	75th percentile value of weighted mobility
Mean_weighted	Weighted mean of mobility from mobile phone users active in an intermediate zone and a given quarter. See data description for details on calculation.

Appendix 1

Calculating home location for mobile phone users by incorporating land use data – v1

UBDC 1/12/22

Aim

To estimate the home location of each unique mobile phone device and thus locate the user geographically and assign area socio-demographic data for further analysis.

Data

Mobile phone application data from **Huq[4] and Tamoco**[5]. Each mobile phone device has the personal identifiers replaced with non-reversible hashed identifiers in the supplied datasets so that points from the same user can be linked. With Huq, the points from individual users can be linked over time through a hashed (anonymous) identifier so estimates of home location can be made across one or multiple years of data. Tamoco resets its hashed identifiers every month, so home locations need to be estimated monthly.

Datazone boundaries are used in home location detection as they are a key geography for the small area or neighbourhood statistics in Scotland, and are available under Open Government License. They are composed of groups of Census Output Areas. Datazones nest within council areas and are designed to have a population of 500-1000. In the study area there are 2336 Datazones.

Land use data from Geomni's UKBuildings layer is used as part of the home location detection (Digital Map Data © The GeoInformation Group Limited (2022), created by Geomni, a Verisk company). Data from 2020 was used. This multi-polygon spatial dataset represents the use, characteristics, and extent of commercial, public and residential buildings across the whole of the UK.

Geography

Data contained with the Glasgow city-region.

Temporal detail

2019 – 2021.

Method

For each user, the home location for a given year (Huq) or month (Tamoco) is determined as the Datazone which maximises the number of active evenings withing residential and mixed residential space based on Geomni data, where an evening is determined as the time between 20.00 and 06.00.

The process is as follows. For each user and each year (Huq) or each month (Tamoco):

1. Subset data points to only those which are <=100m in accuracy.
2. Subset data points to only those falling within the footprint of a residential or mixed residential building (from Geomni data).
3. Count active evenings (between 20.00 and 06.00) in each Datazone.
4. Assign the home location as the Datazone which maximises active evenings[6].
5. Subset evening residential data points to those located within the home Datazone.

---

[4] http://www.ubdc.ac.uk/data-services/data-catalogue/transport-and-mobility-data/huq-data/

[5] https://ubdc.ac.uk/data-services/data-catalogue/transport-and-mobility-data/tamoco-data/

[6] If two Datazones have the same number of active evenings or if the user only have 1 active evening in the Datazone determined to be their home, do no assign a home location.

6. Calculate the location which minimises the distance between these points.
7. Assign a Postcode based on 6.

Appendix 2

Calculating mobility weights for mobile phone data – v1

UBDC 1/12/22

Aim

To produce a set of weights which can be applied to each unique mobile phone device based on its estimated home area. These weights attempt to correct for uneven coverage by socio-demographic status and geography over time. For example, If we find less/more mobile phone users from certain socio-demographic groups and/or certain geographic areas than expected (based on a comparison to adult population), weights help to increase/decrease the impact in any analysis using the data.

Data

Mobile phone application data from **Huq[7] and Tamoco**[8]. Each unique mobile phone has a de-identified device ID. For each de-identified device ID, a home area has been calculated for each year (see the technical document 'Calculating home location for mobile phone data').

Socio-demographic data and population data used in the weighting calculation come from Scottish Index of Multiple Deprivation (SIMD) 2020. For socio-demographic weighting, we use the SIMD quintile grouping which is aggregated based on data for all of Scotland. For geographic weighting we use Councils and Intermediate zones.

Geography

Weights are calculated for data within the extent of the Glasgow city-region. Weights are calculated at two levels: Council (8) and Intermediate Zone (417).

Temporal detail

2019 – 2021, with weights calculated per calendar year for each data set.

Method

Council level weights and **intermediate zone level weights** are calculated in three stages:

1. A **socio-demographic weight** is calculated based on the composition of the mobile phone population **within** SIMD quintiles of the council or intermediate zone.
2. A **geographic weight** is calculated based on the composition of the mobile phone population **between** councils or intermediate zones.
3. The total weight is the combinations of **socio-demographic weight** and **geographic weight** and is calculated for each council or intermediate zone.

Socio-demographic weight

For each **council or intermediate zone** for each year:

A **socio-demographic weight** is calculated for each SIMD quintile as the ratio between the percentage of the mobile phone population in that SIMD quintile and the percentage of the adult population that SIMD quintile (using 2020 data).

[7] http://www.ubdc.ac.uk/data-services/data-catalogue/transport-and-mobility-data/huq-data/
[8] https://ubdc.ac.uk/data-services/data-catalogue/transport-and-mobility-data/tamoco-data/

Geographic weight

For each **council or intermediate zone** for each year:

> A **geographic weight** is calculated as the ratio between the percentage of the mobile phone population in a council or intermediate zone and the percentage of the adult population in that council or intermediate zone (using 2020 data).

Total weight

For each **council or intermediate zone** for each year:

> The **total weight** is calculate by multiplying the socio-demographic weight and the geographic weight.

Weight code

A **weight code** is assigned based on the combination of four pieces of information:

1. Mobile phone data source
2. Year
3. Name of council or intermediate zone code
4. Scottish Index of Multiple Deprivation (SIMD) quintile

E.g.	**Weight code**	**Weight**	**Impact of weight**
	huq_2019_South Lanarkshire_SIMD1	1.234	23.4% increase
	tamoco_2021_S02001506_SIMD5	0.934	16.6% decrease

Each mobile phone user is assigned the **weight code** based on the datazone where they are estimated to live in a given year. This datazone provides an SIMD quintile and can be located within an intermediate zone and council.