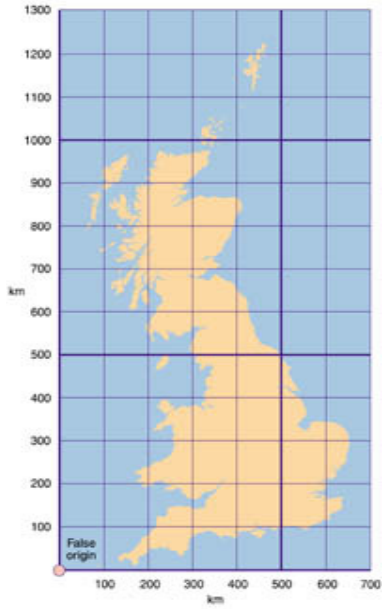
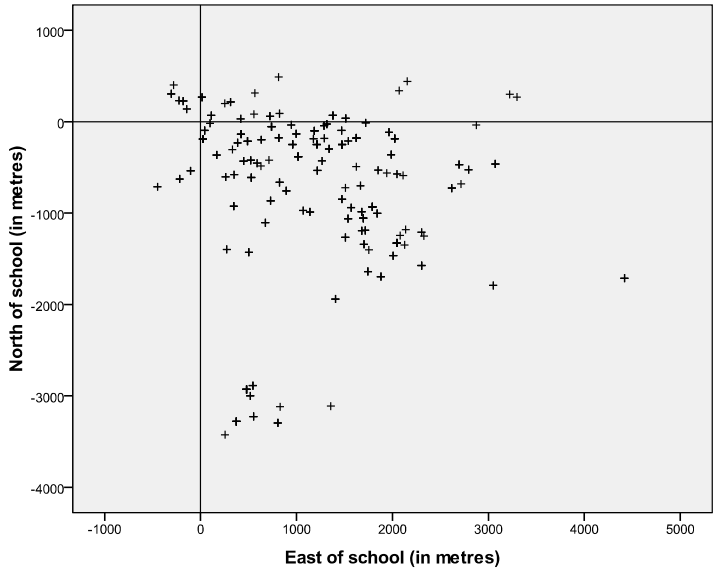
Documentation of postcode coding in the Glasgow data

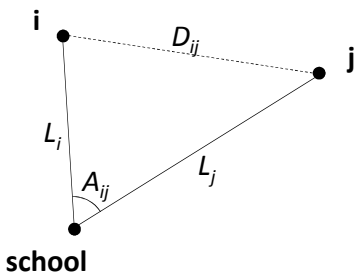


At https://www.ordnancesurvey.co.uk/ it is possible to download detailed geographical information about postcode areas. For all postcodes in the (almost complete) list, so-called *northings* and *eastings* are provided in metres, which give the coordinates in the Cartesian system shown on the right (scaled in kilometres).

Two geographical outliers were recoded as missing to avoid bias in analyses. These outliers seem to be students who either moved into the school district during the first year of the study, or moved out of the district during the last year of the study. The scatterplot below shows the centre of the distribution of geographical locations in the data, with the school location as origin.

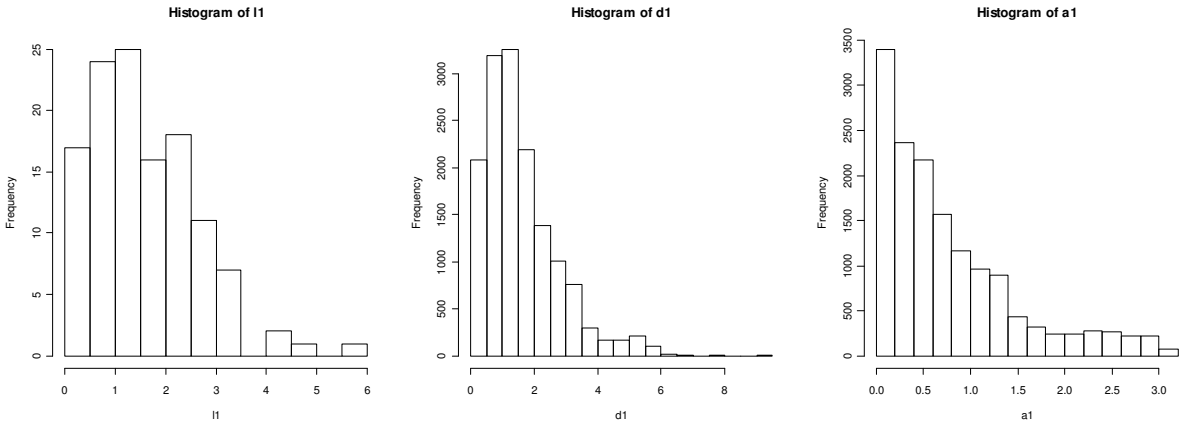


The northings and eastings were transformed into one individual level variable L_i *"length of way to school"* (in kilometres) and two dyad level variables D_{ij} *"distance between student i's home and student j's home"* (also in kilometres) and A_{ij} *"angle between student i's way to school and student j's way to school"* (in radian); the calculations are based on beeline connections, not the actual street grid. One could think of calculating travel distances in minutes with software such as http://www.routenet.nl/, but the added value seemed questionable compared to the effort.

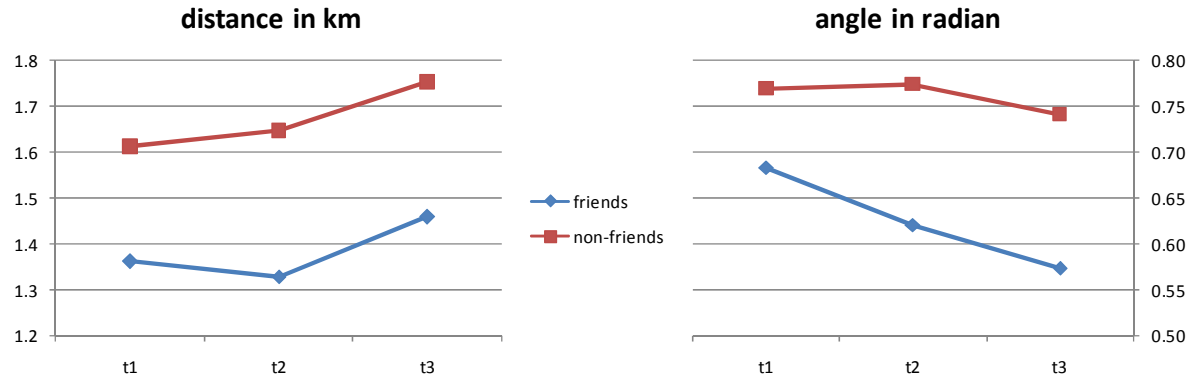


Over time ($t_1 \rightarrow t_2 \rightarrow t_3$) next to the outliers, another 11 students changed their postcode. Three of them moved at t_2, five at t_3, and another three temporarily moved at t_2 and returned to their old t_1 postcode at t_3. Based on the quantitative data, it is impossible to tell the reasons for these change patterns. Because the huge majority (147 of 160) doesn't show a change, one might think of working with at most one geographical location per respondent; in order to avoid causality blunders, it would be advisable to take the first valid response rather than the mode. But as the longitudinal data are available, the better choice may be to just keep them. Overall, there are 122 valid postcode responses for the 150 respondents at t_1, 116 (for 146) at t_2, and 113 (for 137) at t_3.

Below is an impression of the distributions of the variables *L*, *D*, and *A* as calculated for the t_1 observation:



The dyadic variables *angle* and *distance* show the expected patterns of association with friendship. Friendship coincides with shorter distances between students' homes, reflecting ease to meet outside school. On average, non-friends live about 300 metres further apart than friends do. Also, friends' homes lie in more similar geographical direction when seen from school than non-friends' homes do, reflecting opportunity to interact on the way to school or back home.



Christian Steglich, 3 August 2011.