City Research Online

City, University of London Institutional Repository

Citation: Lahtinen, K., Slingsby, A., Dykes, J., Butt, S. & Fitzgerald, R. (2015). Informing Non-Response Bias Model Creation in Social Surveys with Visualisation. Paper presented at the VIS 2015, 25-10-2015 - 30-10-2015, Chicago, USA.

This is the accepted version of the paper.

This version of the publication may differ from the final published version.

Permanent repository link: https://openaccess.city.ac.uk/id/eprint/12333/

Link to published version:

Copyright: City Research Online aims to make research outputs of City, University of London available to a wider audience. Copyright and Moral Rights remain with the author(s) and/or copyright holders. URLs from City Research Online may be freely distributed and linked to.

Reuse: Copies of full items can be used for personal research or study, educational, or not-for-profit purposes without prior permission or charge. Provided that the authors, title and full bibliographic details are credited, a hyperlink and/or URL is given for the original metadata page and the content is not changed in any way.

City Research Online: http://openaccess.city.ac.uk/ publications@city.ac.uk

Informing Non-Response Bias Model Creation in Social Surveys with Visualization

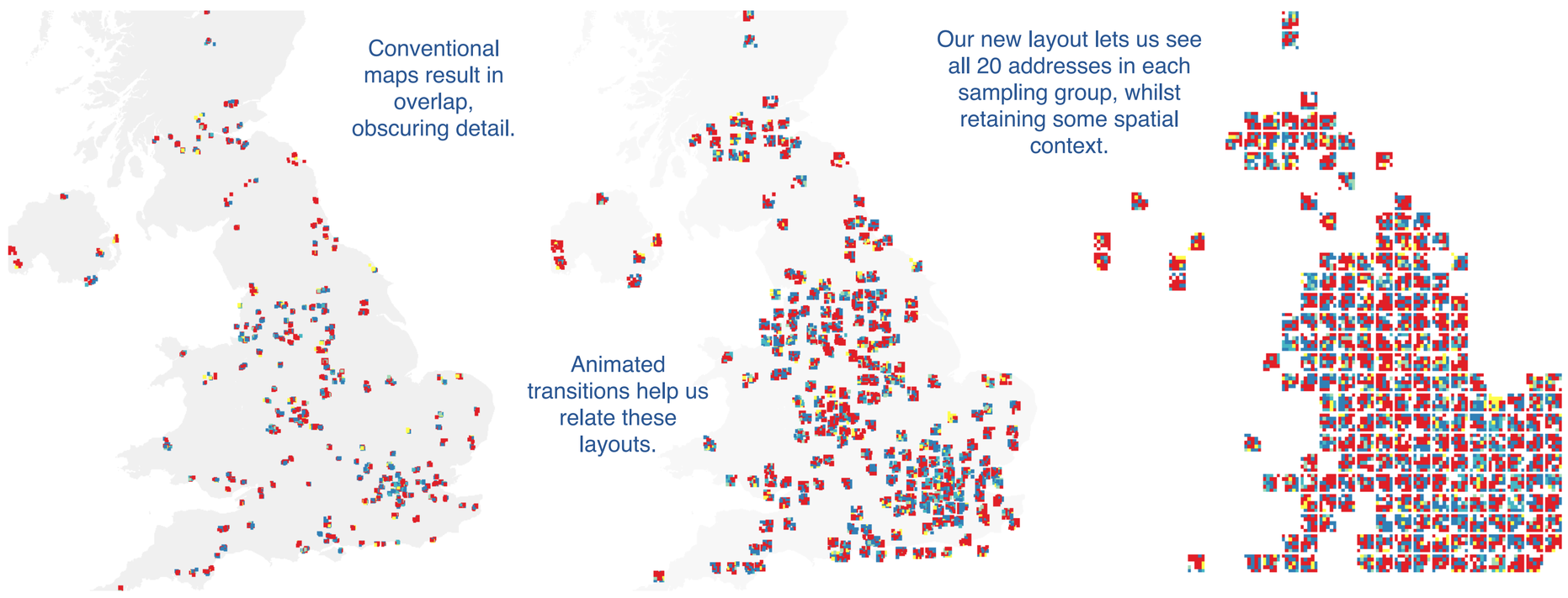
http://openaccess.city.ac.uk/12333

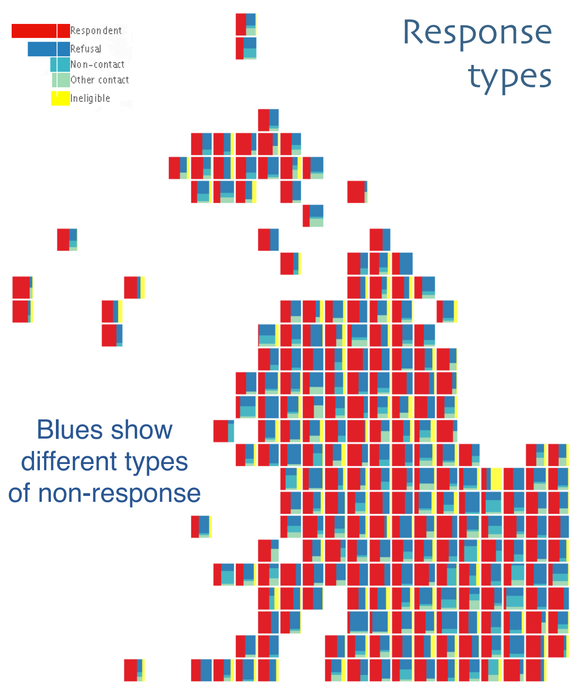
Kaisa Lahtinen, Sarah Butt, Rory Fitzgerald - CCSS, City University London
Jason Dykes, Aidan Slingsby - giCentre, City University London

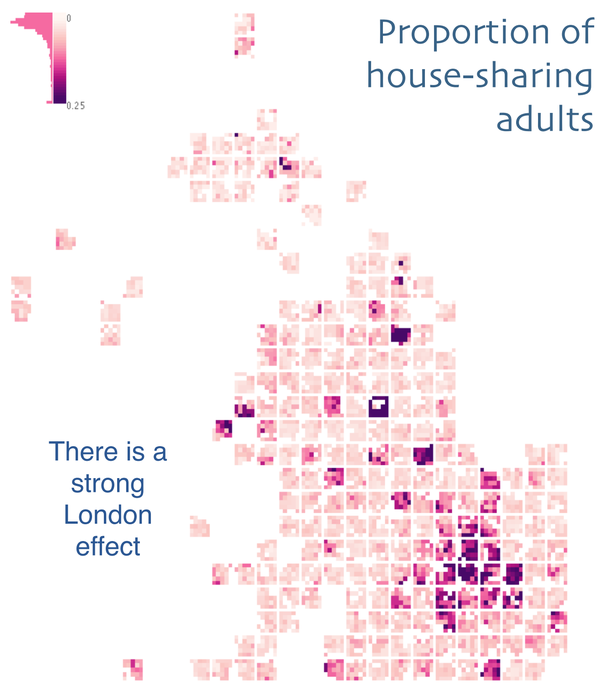
Non-response bias in social surveys is one of the most challenging issues facing social science research. We are building a model to try and predict non-response bias for the European Social Survey. Selecting variables for use in the model is difficult as the pool of variables is large, geographic a unfamiliar. A collaborative design process involving social scientists and visualization is developing solutions.

CREATING A SUITABLE LAYOUT

The 226 sets of 20 tightly clustered household locations are challenging to display in a way that both retains geographical relationship and avoids occlusion. Our carefully-designed layout preserves this two-level hierarchy, in a non-occluding manner.





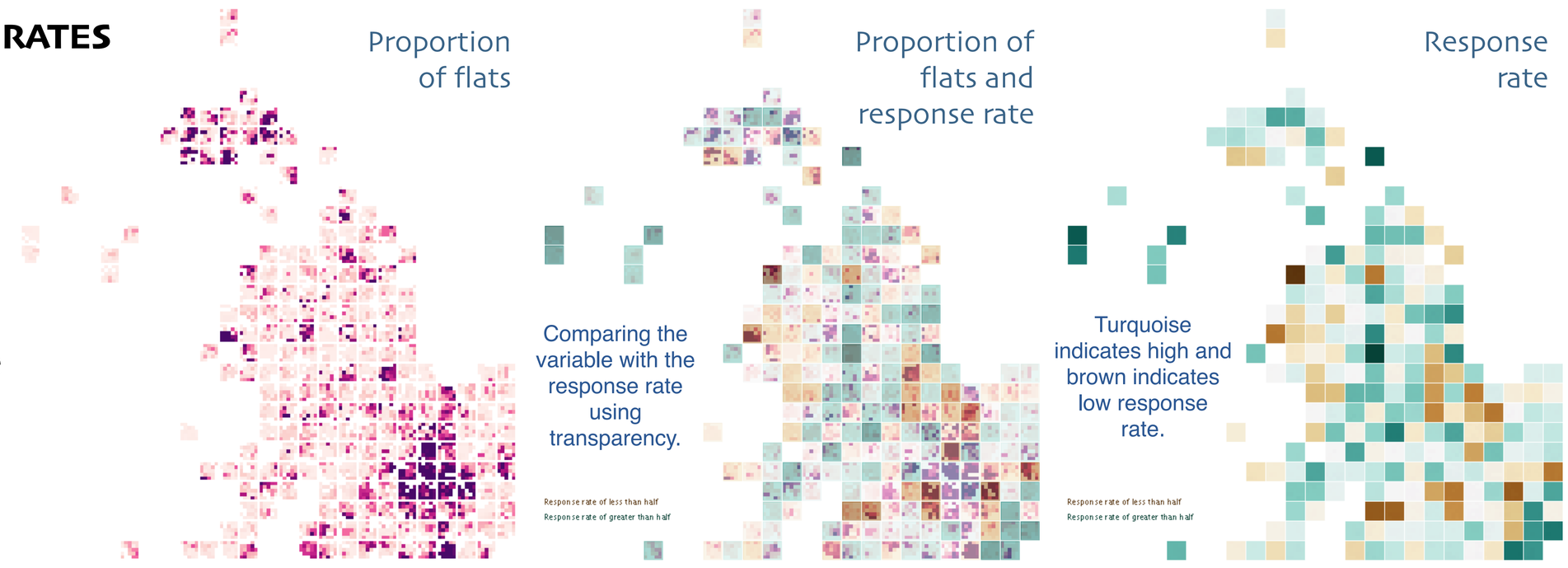


UNDERSTANDING DATA DISTRIBUTIONS

This layout enables us to quickly flick through hundreds of variables, helping us choose those variables with suitable distributions. For example, London-centric density of house-sharing adults makes this a potentially problematic variable for inclusion in a national non-response model.

COMPARING TO RESPONSE RATES

Since we are modelling non-response, we want to know how our candidate variables relate to non-response. Our diverging colour scale emphasises both low (brown) and high (turquoise) response rates. Overlaying it on our variable maps with varying degrees of transparency helps us relate variable to response as we consider local models



Collaborative co-design is helping us sift through hundreds of variables that might help us predict non-response, leading to a better understanding of factors affecting non-response and more informed models.

CITY UNIVERSITY LONDON
EST 1894

European Social Survey

giCentre