Contents

Time Series Models for Small Area Estimation

There is a continuing demand for the Australian Bureau of Statistics (ABS) to provide statistics for small areas. The ABS has both good knowledge and practical experience with cross-sectional small area models. However, there are needs and opportunities that now motivate us to explore methods that can also borrow strength across time. There are current opportunities presented by existing ABS repeated surveys, in particular the Labour Force Survey with its rotating panel design. Future opportunities will arise for any household surveys that transition from periodic to continuous collection, so that sample is no longer concentrated in an occasional large survey but spread evenly over time. In response, we initiated a research project to investigate the application of time series methods for small area estimation.

In the literature two broad approaches are described; temporal multilevel models and structural time series models (Bollineni-Balabay et al, 2016). These methods were applied to the Australian Labour Force Survey to obtain and compare modelled estimates for Labour Force dissemination regions. For the structural time series approach multivariate models were fit, with borrowing strength over regions achieved by allowing potential non-zero correlations between the disturbances of the trend and seasonal components of the model. The Rao-Yu multilevel model (Rao and Yu, 1994) was implemented and estimated as a state space model with unemployment benefits data (provided by the Department of Social Services) used as a regression covariate. Estimation for the models was performed using the Dynamic Linear Model package in R.

The results may be considered preliminary in nature. However, strong improvements in the variances of the modelled estimates compared to the direct estimates were clear, under both the multilevel and structural time series approaches. The work was presented at the November Methodology Advisory Committee meeting. Further work is planned to refine the models, including the incorporation of a seasonal component into the Rao-Yu model, and fitting of Seemingly Unrelated Time Series (SUTSE) bivariate models that incorporate the unemployment benefits data.

References

Bollenini-Ballabay, O., van den Brakel, J., Palm, F. and Boonstra, H. (2016) 'Multilevel hierarchical Bayesian vs. state space approach in time series small area estimation: the Dutch travel survey'. Discussion paper, Statistics Netherlands.

Rao, J.N.K. and Yu, M. (1994) 'Small area estimation by combining time series and

cross-sectional data', *Canadian Journal of Statistics*, 22, pp. 511-528.

For more information:
Sean Buttsworth
Sean.Buttsworth@abs.gov.au

Estimation of Agglomeration Elasticities for Wider Economic Benefits for Transport Projects

The ABS and KPMG were recently engaged by the Department of Infrastructure and Regional Development (DIRD) to estimate agglomeration elasticities for measuring the agglomeration economies of transport projects in Australia. Agglomeration economies refer to the productive advantages that arise from the spatial concentration of economic activities (Mare and Graham, 2009). When firms locate in close proximity to each other, a number of productivity benefits can emerge, for example, increased opportunities for labour market pooling, knowledge and technology sharing, industry specialisation, and increased efficiency of input-output sharing.

A set of agglomeration elasticities by industry division in major Australian cities was calculated using econometric modelling of firm level information from ATO tax data for the years 2006/07 and 2011/12. The estimation was based on the production economic framework and utilised statistically robust methods. We estimated the effect of agglomeration on firm total factor productivity (TFP) by including the effective density[1] measure as a variable in a Cobb-Douglas production function (Graham, 2007 and Graham *et al*., 2009).

The detailed report of the econometric modelling work is provided in Appendix C of the KPMG discussion paper entitled, "Measuring WEBs in Australian cities" which was released for public consultation in June 2017 (KPMG 2017). The KPMG discussion paper also contains methodologies and parameters for estimating all three categories of WEBs expected to arise from large transport projects in urban areas, namely: Agglomeration impacts; Labour market impacts; and Output change in imperfectly competitive markets. These are important inputs to the updating and expansion of the *Australian Transport Assessment and Planning (ATAP) Guidelines,* which outline the best practice for transport planning and assessment in Australia. The update is being undertaken by DIRD in partnership with state and territory infrastructure and transport agencies.

[1]KPMG defines effective density as a quantitative measure of access to opportunities. An opportunity refers to any activity that a user wants to access. In the context of agglomeration, the relevant opportunity is typically jobs. Effective density is quantified using a measure of travel impedance, typically a function of distance, time or generalised cost of travel (KPMG 2017, p.40). KPMG estimated and provided area effective density measures to the ABS for use in the estimation of agglomeration elasticities.

References

Graham, D. J. (2007) 'Wider economic benefits of transport investment: Evaluating the application of UK agglomeration elasticity estimates in the New Zealand context', Report to Land Transport New Zealand, Wellington.

Graham, D. J., Gibbons S. and Martin, R. (2009) 'Transport investment and the distance decay of agglomeration benefits', Report prepared for the UK Department of Transport, Imperial College, London.

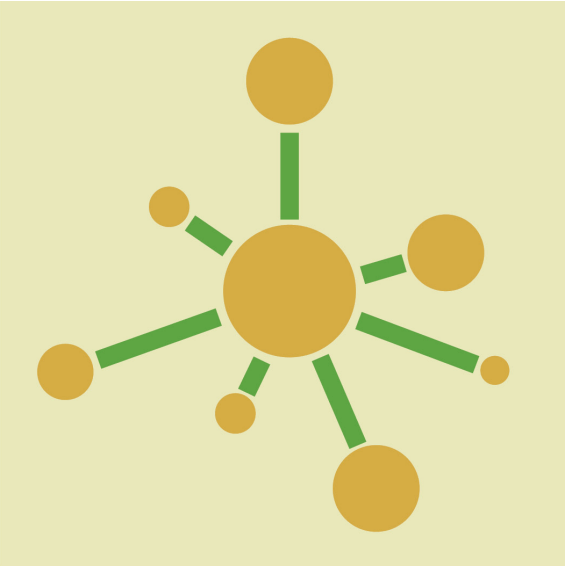
KPMG (2017) 'Measuring WEBs in Australia cities', KPMG Discussion Paper, 01 June 2017, Melbourne.

Maré, D. C. and Graham, D. J. (2009) 'Agglomeration elasticities in New Zealand', NZ Transport Agency Research Report 376, Wellington.

For more information:
Franklin Soriano
Franklin.Soriano@abs.gov.au

Graph-Based Framework for Data Linking Projects

The ABS has developed a prototype graph-based approach for linking heterogeneous multisource data in statistical production. This work is motivated by the need for a more agile analytical practice to support evidence-based decision making and to utilise diverse new sources of human and machine-generated data associated with digital presence, connectivity and interaction in the global network. Such sources include commercial transactions, remote imagery, sensor measurements, geospatial positioning, web content, and online user activity. They are often collectively referred to under the rubric of big data.



Graph-based data linking involves a paradigm shift in the conceptualisation of complex data from tables, records and fields to graphs (network structures) of entities and relationships. Entity-relationship graphs are similar to the sociograms used in social network analysis to visually depict interpersonal relations. However, entity-

relationship graphs can depict any entity or relationship type of analytical utility, and the entities can be multiply connected by different relationships. For this reason, entity-relationship graphs are intrinsically dynamic as their topological properties change over time in response to evolving patterns of interaction in complex systems. The ABS prototype approach makes use of the Semantic Web (Web 3.0) model defined in a set of published standards by the World Wide Web Consortium (W3C) as the operational framework for graph-based data linking.

A feature of the graph-based data linking approach developed by the ABS is the introduction of temporal segments demarcated by events that change the measurable state of individuals and groups in systems of statistical interest. Data entities that correspond to observations of these individuals and groups over time are then explicitly resolved to canonical (or base) entities, for which the temporal segments and events form distinct trajectories through a series of life states. Similar data entities are associated by a set of logically precise equivalence relations that extend the implicit notion of identity in traditional data linking approaches.

The ABS graph-based approach also embeds strong computable semantics in the description of entity and relationship types. This creates a substrate for the execution of new reasoning methods that utilise the inherent connectedness of the entity-relationship graph to make collective linking decisions. These either draw on the logical properties of entity and relationship types through a process of deductive reasoning based on First-Order Logic (FOL), or they generalise the patterns of association for existing entities through an inductive method such as statistical learning on the graph structure. Deduction and induction are synergistic, and may be interleaved in an iterative process. The ABS is currently evaluating automated FOL reasoning based on description logics and horn-clause logic, as well as a range of graph-based inductive reasoning schemes involving relational machine learning and kernel functions.

For more information:
Ric Clarke
Ric.Clarke@abs.gov.au

How to contact us and Email Subscriber List

Methodological News features articles and developments in relation to methodology work done within the ABS Methodology Division By its nature, the work of the Division brings it into contact with virtually every other area of the ABS. Because of this, the newsletter is a way of letting all areas of the ABS know of some of the issues we are working on and help information flow.

We hope the Methodological Newsletter is useful and we welcome comments.

If you would like to be added to or removed from our electronic mailing list, please contact:

Nick Husek
Methodology Division
Australian Bureau of Statistics
Locked Bag No. 10
BELCONNEN ACT 2617

Email: methodology@abs.gov.au

The ABS Privacy Policy outlines how the ABS will handle any personal information that you provide to us.