

Understanding Victorian small area models of legal problems: A practical guide —

Research Brief

Victorian Legal Services
BOARD + COMMISSIONER

Dr Georgina Rychner and Prof Nigel J Balmer

Legal Services
Research Centre

What are small area models?

Small area models are sophisticated statistical tools that help us understand where everyday legal problems are likely to occur across Victoria. These models work by combining legal needs survey data about people's experiences of legal problems from the Public Understanding of Law Survey (PULS),¹ with detailed Census information² to predict the relative likelihood of problems in small geographic areas.³

Rather than simply mapping where civil legal problems have been reported (sometimes referred to in our context as 'expressed need'), small area models use statistical relationships to estimate where problems are likely to exist, including where they may not be acted upon or even recognised as being legal in nature.

The result is an interactive map which shows where problems are more or less likely, at various geographic levels. You can see the [map and access the underlying data here](#).

The aim is to create a picture of everyday legal problem prevalence across Victoria that can help to make service planning easier and more effective.

Who are these models for?

The models serve anyone interested in understanding everyday legal problem prevalence across Victoria. Key user groups include legal service providers (legal aid, community legal centres, and private practitioners), government agencies, community organisations, researchers, and policy makers.

They may also fill a range of purposes, including strategic planning, resource allocation, legal needs assessments, testing assumptions about where legal problems occur, and supporting funding applications.

How do I use the interactive maps?

The map allows users to view estimates of everyday legal problem prevalence at seven geographical levels that are set out below in Table 1. There are two sets of models. The first includes all everyday legal problems. The second excludes consumer problems, given their volume.

Once you have selected an area on the menu, double click on any area on the map to retrieve the small area estimates of legal problem prevalence. Double click on multiple areas to compare estimates between areas.

[lsbc.vic.gov.au/
research-centre](https://lsbc.vic.gov.au/research-centre)

Table 1. Small area levels⁴

Small area level	Use and reach
Statistical area 1 (SA1)	The smallest area level, containing between 200-800 people
Statistical area 2 (SA2)	Medium-sized general-purpose areas built from SA1s, containing between 3,000-25,000 people
Statistical area 3 (SA3)	Larger general-purpose areas built from SA2s, containing between 30,000-130,000 people
Postcode	Mail code that scales with population, containing between 25-90,000 people
Local Government Area (LGA)	Approximation of gazetted local government boundaries as defined by the state, containing between 850-300,000 people
State electorate division	Electoral divisions at the state level, containing between 50,000-110,000 people
Commonwealth electorate division	Electoral divisions at the federal level, containing between 115,000-200,000 people

How do I use the accompanying datasets?

Underlying data used to create the interactive map is available in seven .csv files, one for each small area level. All boundary data is from 2021, with the exception of the Local Government Area (LGA) and State Electoral Division (SED22P) data, which used 2022 boundaries.

Each file has the variables shown in Table 2, with some exceptions for SA1 and SA2 data (which are denoted in Table 2 with an asterisk).

Some rows in the SA1, SA2 and postcode datasets were omitted completely due to ABS privacy requirements, for example, where an area had fewer than ten persons or households. On crosschecking the SA1 data, we removed three additional SA1s where the results had been skewed by a combination of few households paired with large industrial sites.⁵

Each estimate variable in a dataset appears twice – the first variable includes all problem types, the second variable (labelled with NC, e.g. LGA_MeanProblemsNC) omits consumer problems.

Table 2. Variables in the Small Area Modelling Datasets

Variable	Variable information
[Area], e.g. SA1	ABS area code
[Area]name, e.g. SA3name	Area name (if available)
[Area]_MeanProblems e.g. SA1_MeanProblems	Mean estimated problems per person for each area
[Area]_MedianProblems	Median estimated problems per person for each area
[Area]_MaxProblems	Maximum problems per area
[Area]_MinProblems	Minimum problems per area
[Area]_Spread	Spread of problems per area
[Area]_StdDev	Standard deviation of problems per area
[Area]_IQR	Interquartile range of problems per area
[Area]_SumProblems*	Sum of problems per area
[Area]_Persons*	Population of area

How were these models created?

Creating the small area models involved five key phases:

PHASE 1 - Model Development: Having matched variables between the PULS and the 2021 Census, statistical models were fitted using PULS data to estimate legal problem prevalence. These models sought to optimise the variables to be included in the model to ensure everything played a role, avoiding unnecessary complexity.⁶

PHASE 2 - Multiple imputation: Missing values in Census records for our matched variables were estimated and added to a new Census dataset.⁷ This ensured that we could calculate estimated problem prevalence for all Census respondents rather than only those with complete data on matched PULS/Census variables.

PHASE 3 - Microdata Application: Final PULS model coefficients were applied to individual-level Census records, generating predicted number of legal problems for each adult Victorian based on their demographic and socioeconomic characteristics.⁸

PHASE 4 - Geographic Aggregation: Individual predictions were aggregated to the various geographies in Table 1.

PHASE 5 - Visualisation and Planning Tools: Having applied ABS treatment requirements to ensure individuals could not be identified, the aggregated estimates were extracted from DataLab, transformed into our interactive map and made available as datasets.

Important caveats and limitations

Small area modelling represents an iterative process that continuously evolves as new data becomes available, and methods improve. These models reflect our current best understanding, but will be refined using our ongoing Legal Understanding and Lawyer Use (LULU) Survey data. They are most effective when supplemented with local knowledge and complementary information sources. They can also fail to capture local factors (e.g. significant job loss, demographic shift, population growth/decline) which may affect legal problem experience.

What's next?

The Legal Services Research Centre will continue to develop and refine the models using data from our ongoing LULU Survey. We will also shortly publish an accompanying report setting out the development of the models, detailed statistical output, and their research and policy context. We are happy to discuss them further with any prospective users.

¹ Nigel J. Balmer, Pascoe Pleasence, Hugh M. McDonald and Rebecca L. Sandefur, The Public Understanding of Law Survey (PULS) Volume 1: Everyday Problems and Legal Need (Victoria Law Foundation, 2023); Nigel J. Balmer, Pascoe Pleasence, Hugh M. McDonald and Rebecca L. Sandefur, The Public Understanding of Law Survey (PULS) Volume 2: Understanding and Capability (Victoria Law Foundation, 2023); Nigel J. Balmer, Pascoe Pleasence, Hugh M. McDonald and Rebecca L. Sandefur, The Public Understanding of Law Survey (PULS) Volume 3: A new perspective on legal need and legal capability (Victoria Law Foundation, 2024). For more on legal needs surveys, see OECD/Open Society Foundations, Legal Need Surveys and Access to Justice (OECD Publishing, 2019).

² The 2021 Census was administered by the Australian Bureau of Statistics and unit-level data can be accessed by application through the ABS Datalab. See Australian Bureau of Statistics, “Datalab”, released 3 November 2021: <https://www.abs.gov.au/statistics/microdata-tablebuilder/datalab>.

³ The final models incorporate variables including age, sex, Aboriginal or Torres Strait Islander status, main language spoken, family status, work status, highest education, long-term illness or disability, mental distress, housing status, income, whether born in Australia and business ownership. Full technical details including complete variable specifications, our approach, rationale for steps taken and contextualisation will be provided in Georgina Rychner and Nigel J. Balmer, “Small Area Modelling of Everyday Legal Problems” (forthcoming).

⁴ For more information on these area levels, see Australian Bureau of Statistics, “Geography Standard (ASGS) Edition 3”, released 20 July 2021: <https://www.abs.gov.au/statistics/standards/australian-statistical-geography-standard-asgs-edition-3/latest-release>.

⁵ Located in Moorabbin, Clayton South and Avalon Airport.

⁶ In practice, this meant Least Absolute Shrinkage and Selection Operator (LASSO) for feature selection and regularization, and Zero-Inflated Poisson (ZIP) regression for handling count data with excess zeros. Full details in Rychner and Balmer, “Small Area Modelling of Everyday Legal Problems” (forthcoming). Variables included in the models were: sex, age group, Aboriginal or Torres Strait Islander, place of birth, language other than English spoken at home, highest education, work status, business owner, income, housing status, family status, carer, long-term illness or disability, and mental distress.

⁷ Using a process called multiple imputation.

⁸ This step would not be possible in most jurisdictions and is a testament to the way in which ABS maximises the utility of the public data while maintaining safety and security.