Theoretical focus

Our paper has two main theoretical focuses. First, we are attempting to understand how the geographic concentration of certain demographic groups in different areas helps explain spatial variation in economic indicators. It is a pertinent question whether migrant communities, Indigenous communities, or other marginalised groups, such as the disability community, live in geographically concentrated areas which may lead to different economic indicators for these areas than areas that are predominantly populated by Australian-born university graduates. Survey data may be limited in studying the economic outcomes of sub-populations if they live in geographically concentrated areas, if the sampling technique is meant to represent the entire country of Australia from a location perspective. Thus our research shows the potential of administrative data to supplement economic insights of survey data for certain demographic groups, which may result in location-based disadvantage.

Second, we are seeking to understand how trends in internal migration assist in explaining spatial variation in economic indicators. We are looking at ways administrative data can supplement the Census. At the moment the Australian Census is cross-sectional, except for 5% of the sample that is linked as part of the Australian Census Longitudinal Data (ACLD) set. The ACLD allows for limited geo-spatial analysis, due to its small sample size. We will present five-year change statistics. A confounding factor is that the area a person lived in in the year 2022 may not be the same area they lived in in 2017. We thus place individuals in their location in 2022 and calculate how their economic position has changed over the past five years, regardless of where a person lived in 2017. It seems intuitive that some of our five-year change indicators are also picking up trends in internal migration, such as young professionals moving to the inner parts of capital cities for work, and we are keen to seek feedback from conference attendees on how best to control for this confounding factor.

In order to study the abovementioned theories, our team at the Australian Bureau of Statistics has developed rigorous methods of scoping administrative data that includes Australia's ever-resident population to a point-in-time population estimate at 30th June 2022 and 30th Jun 2017. Furthermore, we have linked each person to a location at the SA2 level of geography. SA2s are geographic boundaries that contain between approximately 3,000 and 25,000 people. There are 2,473 SA2s covering Australia, without gaps or overlaps.

Data

Our data source is the Person Level Integrated Data Asset (PLIDA). PLIDA is a secure data asset combining information on health, education, government payments, income and taxation, employment, and population demographics (including the Census) over time, beginning in 2006. The primary data sources we utilise are government benefits (Centrelink) data, and Australian Taxation data (including income tax returns, payment summaries and Single Touch Payroll).

Research methods

There are four distinct steps we undertook to create a population snapshot of economic indicators:

1. **Creating an ever-resident administrative population:** The 'ever-resident' population of Australia is created through combining the populations from Medicare registrations, Services

Australia Administrative data (previously known as Centrelink data) and Personal Income Tax data. On the 30 June 2022 this included over 37 million people.

2. **Capturing the population at a point in time (population scoping):** Given we have an ever-resident population, we then 'scoped' the data to provide a population snapshot of Australia on 30 June 2022 and 30 June 2017. Firstly, we removed from the Spine those who were born after or died before this date. Secondly, we removed those for whom migration data revealed they were out of the country. Finally, we removed people who have no recent record of government activity (for 1-5 years, depending on their age). These people are assumed to have died or left the country but have not matched date of death or overseas migration record.
3. **Locating people at a point in time:** We placed individuals in a location for the 30 June 2022. People can report different addresses over time as they interact with government services. We prioritised residential addresses over postal addresses. Furthermore, we utilised the most recent address recorded if individuals had different addresses currently active with different administrative sources. We were able to place slightly over 99% of Australians in an SA2 on the 30 June 2022 using this method.
4. **Deriving information about people in the population snapshot, and then placing this information in interactive maps:** Finally, we derived economic indicators to explore our theoretical questions of interest. These include the proportion of individuals in an area whose income has increased faster than inflation between 2017-2022 (i.e. experienced real wage growth); if an area has high-government/low-employment or low-government/high-employment income; the Gini coefficient of an area; and the proportion of individuals in an area who have been continuously on the welfare payments of JobSeeker, Disability Support Payment or both between the period 2017-2022. Each interactive map has three main features that convey different pieces of information: the background colour of the SA2, the colour of the circle indicating the SA2, and the size of this circle. Putting out data in maps enabled us to identify clear geo-spatial population trends across Australia.

Preliminary findings

Across all our maps, there seemed to be a clear distinction between capital cities and the remainder of the country. Furthermore, across all economic indicators, the Australian Capital Territory appeared relatively more well-off, whilst Tasmania and the Northern Territory were the poorest states. We highlight below some selected analytical findings.

First, Figure 1 illustrates the proportion of individuals whose income has increased faster than inflation between 2017-2022. The colour brown indicates few individuals experienced real wage growth, whilst the colour teal indicates the majority of individuals experienced real wage growth. Figure 1 (a) and Figure 1 (c) reveal that, respectively, for regional New South Wales and regional Victoria real wage growth has slowed. However, in the inner-city Sydney and Melbourne (i.e. Figures 1 (b) and 1 (d), respectively), there was high real wage growth. This effect could be due to retirees moving away from inner-city areas to regional and coastal areas, or university graduates and young professional moving to city centres for job opportunities.

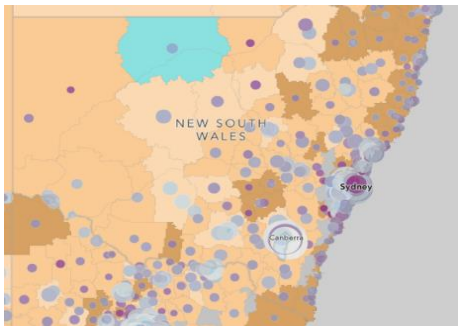

Figure1 (a): Income growth (NSW)

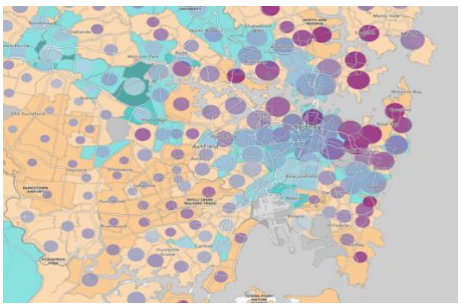

Figure 1 (b): Income growth (Sydney)

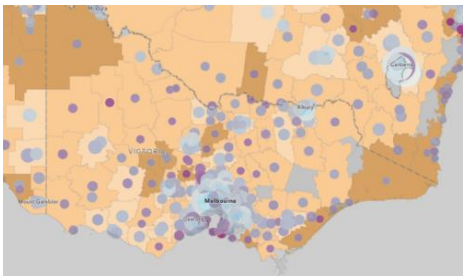

Figure 1 (c): Income growth (Victoria)

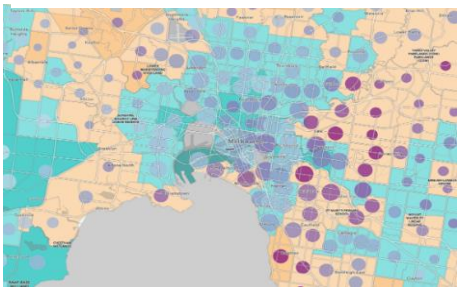

Figure 1 (d): Income growth (Melbourne)

Second, we would like to draw your attention to a map illustrating the Gini coefficient and main source of income. These economic indicators can be seen in Figures 2, (a)-(d). A low Gini, or relatively homogenous income within an SA2, is given by a yellow/green colour whilst a high Gini, or high inequality, is given by a red colour. Areas outside of capital cities are consistently relatively homogenous in income. However, there is great inequality in the inner-city areas. Our initial analysis reveals this is upper-tail inequality, that is, there are few people whose primary source of income is government income in inner-city areas. The pink dots in Figure 2 reveal high-government/low-employment income, whilst the blue dots on Figure 2 reveal low-government/high-employment income. As can be observed, regional areas are high in government income, and capital cities are high in employment income. Figure 2 (c) reveals that in the inner south-west Sydney, predominantly populated by CALD Australians, there is high government income.

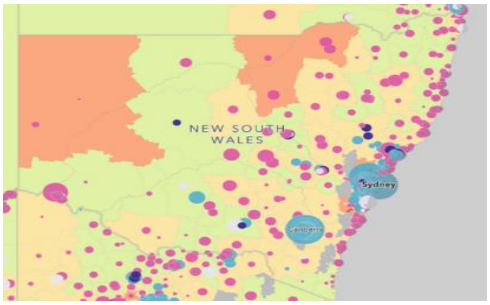

Figure 2 (a): Gini coefficient and main income source (NSW and Victoria)

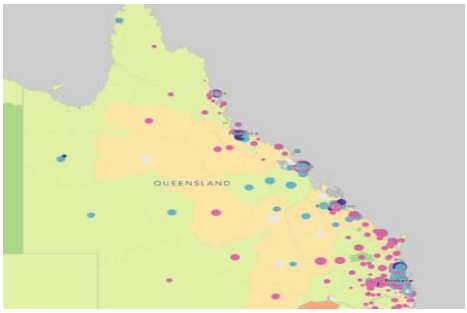

Figure 2 (b): Gini coefficient and main income source (QLD)

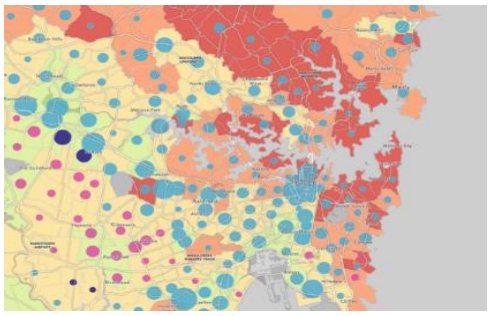

Figure 2 (c): Gini coefficient and main income source (Sydney)

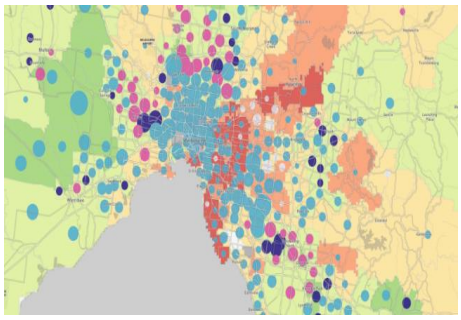

Figure 2 (d): Gini coefficient and main income source (Melbourne)

Finally, we would like to share our analysis on persistent welfare receipt. A darker colour in Figures 4, (a)-(d) reveal a higher proportion of individuals have remained on either JobSeeker, the Disability Support payment or both for a 5-year period. The darkest colours are known to be areas with a high Indigenous population, or a high CALD-population, such as inner south-west Sydney in Figure (c). It is notable to compare the dark coloured purple in some SA2s in Figure 4 (c), which indicates more than 45% of adults have persistently remained on either of these two welfare payments, and the light colour of SA2s near Sydney Harbour in Figure 4 (d), which indicates less than 5% of individuals in these SA2s have remained on these payments for four years.

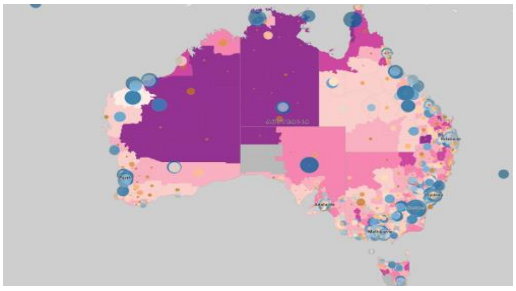

Figure 4 (a): Persistent Welfare (Australia)

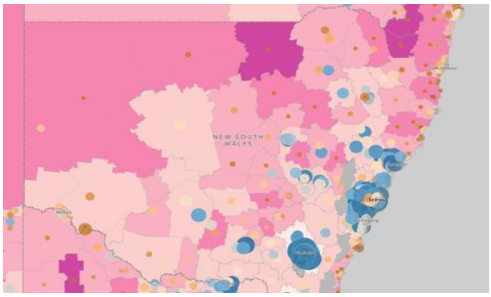

Figure 4(b): Persistent Welfare (NSW)

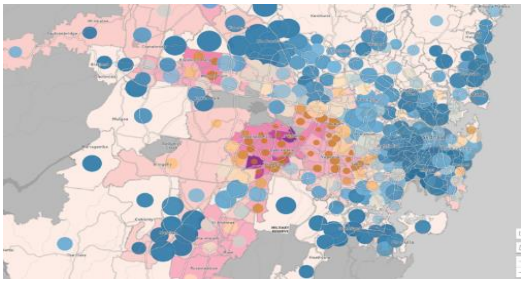

Figure 1 (c): Persistent welfare (Sydney)

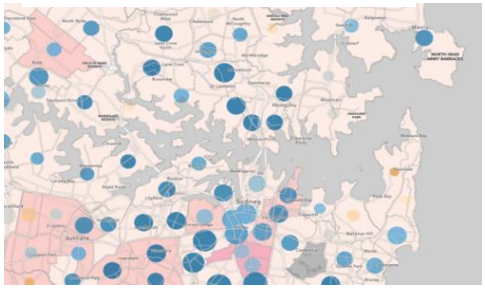

Figure 4 (d): Persistent Welfare (Inner Sydney)

Remarks

We would greatly appreciate conference participants feedback on our methodology, the type of experimental demographic statistics we have decided to map, and whether there is appetite for the Australian Bureau of Statistics to make these interactive spatial maps publicly available.