The spatial patterns of ethnic diversity

Introduction

Migration has and continues to change the ethnic and cultural fabric of most developed nations. Established migrants and their descendants from earlier waves have been supplanted by new and large migrant groups from diverse origins, introducing greater complexity in the composition of cities, towns and local communities. Vertovec (2007) coined the concept of 'superdiversity' to capture how the influx of migrants from a growing array of origin countries has contributed to the diversification of host societies along multiple dimensions, including ethnic, linguistic, religious and socioeconomic. Capturing the spatial patterns of diversity though is made difficult by the sheer number of groups that have settled in host countries and the multiple ancestral, cultural and economic domains on which diversity exists. A great deal of methodological innovation has occurred in recent years to understand ethnic and migrant settlement patterns, including in typologising the ethnic make-up of local areas and measuring the geographic scale at which processes of segregation occur (Manley et al. 2019; Parisi et al. 2011; Poulsen et al. 2010). However, in most of this research, racial and ethnic groups are constructions that aggregate individuals and groups of very different origins. In US research, for instance, racial groups are typically based on people who identify as white, black, Hispanic, Asian or other. Particularly striking is the aggregation of all people of Asian background, a region comprised of many different ethnicities, languages and religions, as well as history of interethnic conflict and tension.

Aggregation is typically for practical and pragmatic reasons – if not, unavoidable. Many Census questionnaires compel respondents to self-select groups with tick box responses to racial, ethnic and ancestry-based questions, or have done so in the past. Researchers are then bound by those groups particularly if they want to construct a consistent time series. Questionnaires in other countries provide greater opportunities for free text responses. While advocated in public and scholarly debate, free text fields can produce a messy, highly variable, inconsistent and overlapping set of responses (Aspinall 2012). In the Australian Census, for example, respondents can tick a box corresponding to one of the eight largest ancestral groups in the country or write in their own self-described ancestry (ABS 2017). Large groups of respondents select or write in an ancestry corresponding to a nation state (e.g. English, Lebanese and Indonesian), a supranational ethnicity or pan-national identity (e.g. Arab, African, Asian and European) or a sub-national ethnicity (e.g. Javanese, Sinhalese and Punjabi). Notable non-state ethnic groups include Australian Aboriginal, Maori, Punjabi-Sikh,

Assyrian, Tamil, Sinhalese, Chaldean, Hazara, Fijian-Indian, Kurdish and Bengali. The variability and inconsistency of free-text responses creates issues in measuring ancestral diversity.

Data and methods

We use data from the 2006 and 2016 Australian Census to map the ethnic diversity of local communities across the country (ABS 2007, 2017). Ethnic diversity is measured from a mixture of multiple choice and free text responses to questions of ancestry, religion and language. Responses are organised into three separate three-level hierarchies. The hierarchies for ancestry and language are based on geography. The highest level refers to broad regions of the world, for example, North-West European, the Middle East and North Africa and South East Asia. The middle level refers to more refined regions of the world, for example, Britain, Arab and Maritime South-East Asia. The lowest level provides the highest level of detail, referring to specific national and sub-national ancestries and languages, including English, Lebanese/Arabic and Javanese/Indonesian. In the case of religion, the highest level of the hierarchy is based on the major religious traditions (e.g. Christianity, Islam, Buddhism, Hinduism, Judaism, other religions and no religion (including other spiritual beliefs)). Lower levels contain detailed Christian denominations (e.g. Catholic, Anglican), other religions (e.g. Sikh) and different forms of spiritual and non-religious belief (agnosticism and atheism). Importantly, information on religious sects and denominations are not given for religions other than Christianity.

The variability and inconsistency of free-text responses creates issues in measuring ancestral diversity. In many cases, it may be appropriate to treat responses as they are provided, thereby respecting and placing value in how individuals identify themselves – or are identified by their households. In other cases though, treating each self-identified set of responses as a separate group overstates the number of groups and apparent diversity. In the 2016 Census (ABS 2017), for example, the ancestry of 42,100 people was coded as ‘Arab not further defined’, while, 367,900 people cited their ancestry as belonging to a specific Arab nation state, most commonly Lebanon (230,900), Egypt (50,500) and Iraq (42,900). If ‘Arab not further defined’ is treated as a separate group, artificial distance is created between people in this group and Arab populations who nominate a nation state as their ancestry, thus exaggerating the level of diversity in areas where they co-exist.

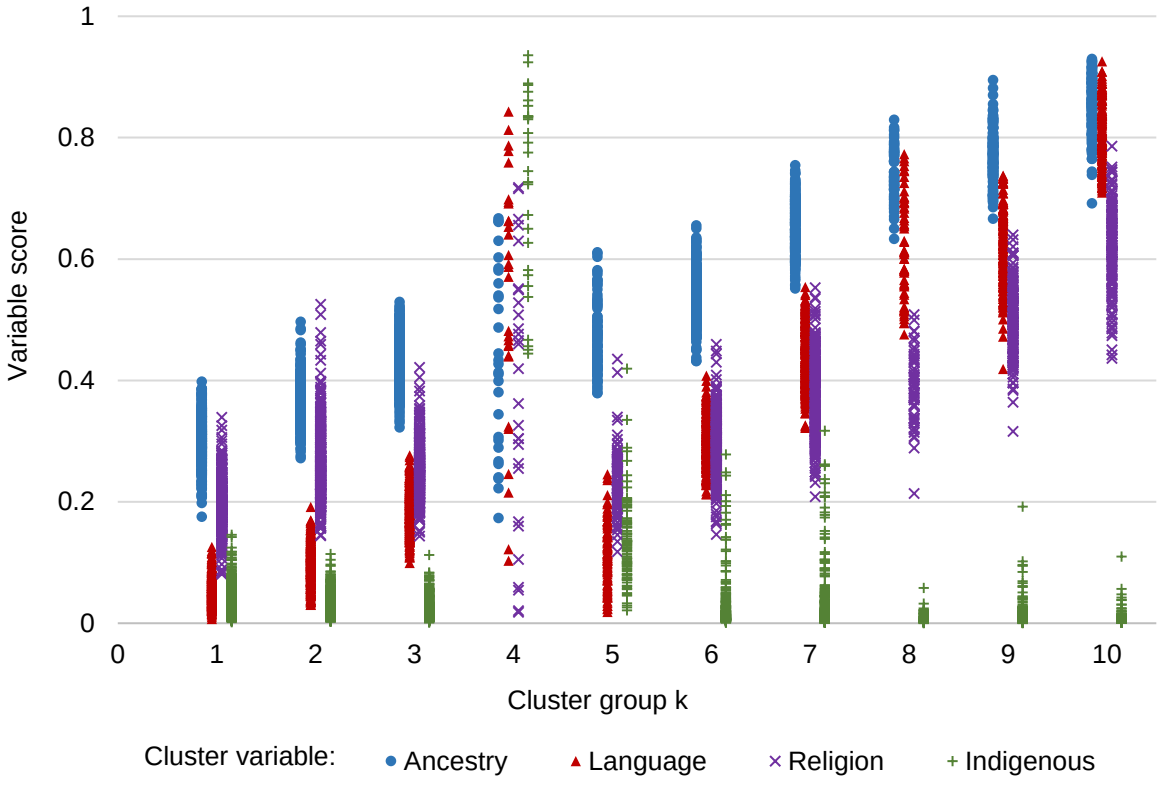
We take advantage of the hierarchical ancestry, religion, language and geography structures to estimate revised population counts. A series of procedures based on Iterative Proportional Fitting (IPF) are used to revise and fit counts at progressively more detailed levels

of geography and ethnic domain. In this way, counts at the most detailed levels are revised based on the strength and greater reliability of counts at increasingly less detailed levels. This serves several purposes: a) it removes or reduces the effect of perturbation on small numbers; b) it proportionately reassigns people who identify with relatively broad pan-national ancestries (for example, African and Arab) to more specific ethnic or national groups; and c) it imputes ethnic domains for missing Census records and those where the domain is not stated. The latter is particularly pertinent for responses to the religion question, as this is the only non-compulsory question on the Census and usually has a relatively high rate of non-response.

Ethnic diversity is computed by separately calculating the Hirschman-Herfindahl Index for each local area and dimension of ancestry, religion and language. The Index approximates the probability of two randomly selected (with replacement) individuals belonging to a different group. Cluster analysis is used to combine index scores into a single measure of ethnic diversity. We also include the proportion of the population in the area who are Indigenous Australian. The cluster analysis offers at least two key benefits. Firstly, it allows us to use very detailed data on ancestral groups while adjusting for the relative salience of ethnic boundaries. For example, we can treat two seemingly similar ancestral groups as distinct, relying on their religious and lingual similarities to reduce the measured diversity of areas in which they co-reside. Secondly, the cluster allows for the examination of the ways in which ancestral, religious and language diversity correlate and for the identification of areas that may be diverse across one or two but not all ethnic domains. In this way, we may detect areas of high ancestral diversity that are bound by a common religion and/or language.

Results

The cluster analysis is performed with ten groups. The distribution of diversity scores is shown in Figure 1. Group 10 represent the most diverse areas. These are highly ancestrally, linguistically and religiously diverse. They are predominantly located in Sydney and Melbourne (Figures 2 and 3) and consist of long established minority groups, as well as substantial shares of new migration. Interestingly, group 9 is also very ancestrally diverse but have substantially less language diversity. These are primarily areas that have high diversity but where minorities are more likely to speak only English at home. Group 8 is comprised of areas that are substantially less religiously diverse, largely because ancestrally diverse, they are relatively more likely to be made of Christian and non-religious populations. Group 4 are Indigenous communities, which have wide distributions in terms of their ancestral, language and religious diversity.



Source: authors' calculations from ABS (2017)

Figure 1 Distribution of ethnic diversity scores by cluster, SA2 Australia, 2016

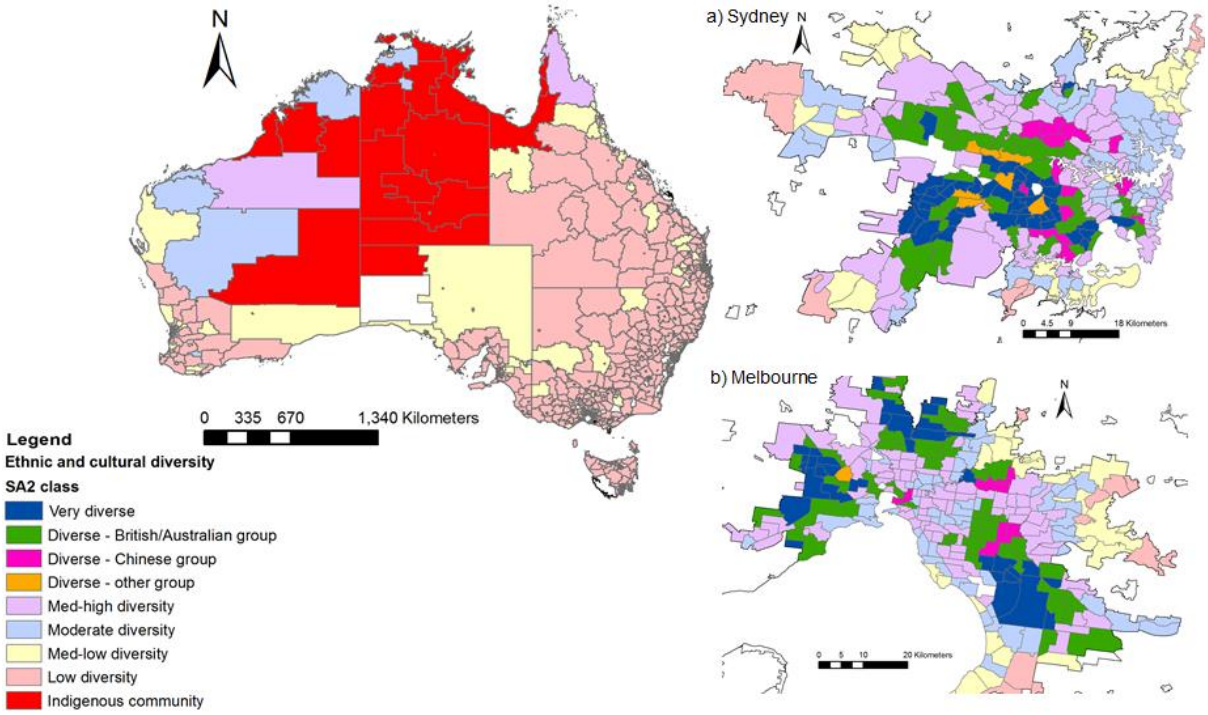
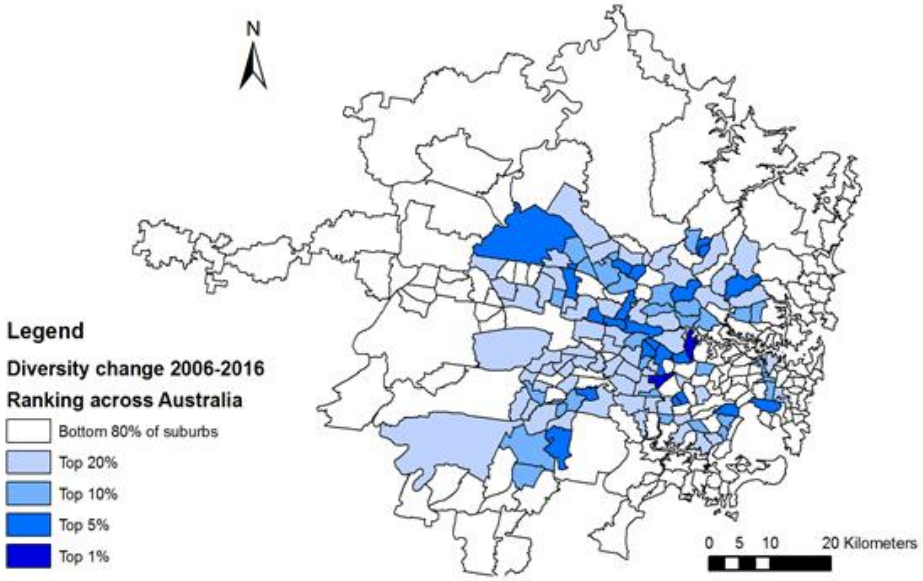


Figure 2 Ethnic cluster across Australia, 2016

The most diverse areas (group 10) have large populations of established and new migrants. Many of these were already highly diverse areas before the arrival of new migrant waves. They are largely highly disadvantaged areas with high rates of low income households, unemployment and low rent housing and low rates of education. New migrants moving into these areas have substantially shifted the ethnic composition of the population (measured by the Index of Dissimilarity), resulting in the fact that many of the most ethnically diverse parts of the major cities are also those that are changing the most. This is particularly the case in Sydney (Figure 3a), while in Melbourne, new migrants are more likely to settle in the outer fringe where the city is expanding into.

a) Sydney



b) Melbourne

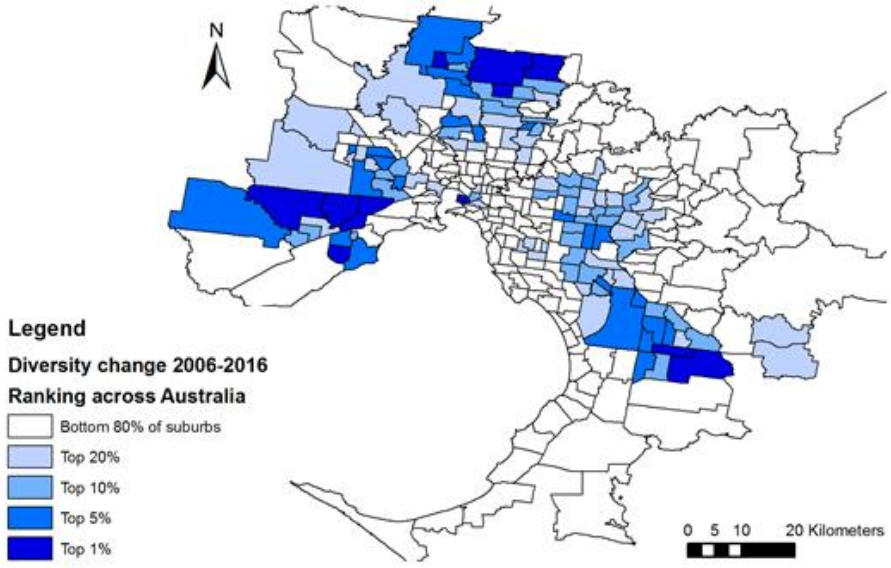


Figure 3 Ethnic composition change in Sydney and Melbourne, 2006-2016

Conclusion

In this paper, we set out to contribute to existing knowledge in three key areas. Firstly, we offer an approach for estimating and harmonising open text data on ethnicity so as to better capture the effects of superdiversity on residential settlement patterns. Secondly, we combine information on ancestral, lingual and religious diversity to characterise the composition of local areas and explore their correlations and differences. Thirdly, we associate the ethnic diversity of areas with their recent migrant intakes and socioeconomic profiles to understand where new migrants are settling and how they are changing the profile of their host communities. The results have valuable insights for migration and assimilation theory, particularly in adding an understanding of the heterogeneity of migration populations and the spatial interaction between new migrants and established minority groups and the mediating role of socioeconomic disadvantage. In view of the literature on ethnic diversity and social capital and cohesion (e.g. Putnam 2007), there are also important though largely unexplored policy implications for understanding how new migrant groups interact and integrate with diverse host communities.

References

ABS (2007) *Census of Population and Housing 2006*. Canberra, Australian Capital Territory: Australian Bureau of Statistics

ABS (2017) *Census of Population and Housing 2016*. Canberra, Australian Capital Territory: Australian Bureau of Statistics

Aspinall, P.J. (2012) "Answer Formats in British Census and Survey Ethnicity Questions: Does Open Response Better Capture 'Superdiversity'?" *Sociology*, 46(2), pp. 354-364

Manley, D., Jones, D. and Johnston, R. (2019) "Multiscale segregation: Multilevel Modeling of Dissimilarity - Challenging the Stylized Fact That Segregation Is Greater the Finer the Spatial Scale" *The Professional Geographer*, 71(3), pp. 566-578

Parisi, D., Lichter, D.T. and Taquino, M.C. (2011) "Multi-Scale Residential Segregation: Black Exceptionalism and America's Changing Color Line" *Social Forces*, 89(3), pp. 829-852

Poulsen, M., Johnston, R. and Forrest, J. (2010) "The intensity of ethnic residential clustering: exploring scale effects using local indicators of spatial association" *Environment and Planning A*, 42, pp. 874-894

Putnam, R.D. (2007) "E Pluribus Unum: Diversity and Community in the Twenty First Century The 2006 Johan Skytte Prize Lecture" *Scandinavian Political Studies*, 30(2), pp. 137-174

Vertovec, S. (2007) “Super-diversity and its implications” *Ethnic and Racial Studies*, 30(6), pp.1,024-1,054