Labour Economics 66 (2020) 101861

Contents lists available at ScienceDirect

Labour Economics

journal homepage: www.elsevier.com/locate/labeco

Intergenerational mobility across Australia and the stability of regional estimates

Nathan Deutscher[a,1,*], Bhashkar Mazumder[b]

[a] Australian National University
[b] Federal Reserve Bank of Chicago, University of Bergen and Australian National University

ARTICLE INFO

JEL classification:
J62
R0

Keywords:
Intergenerational mobility
Economic geography
Australia

ABSTRACT

We produce the first estimates of intergenerational mobility in Australia using administrative data, covering a million individuals born between 1978 and 1982. Australia emerges as one of the more mobile advanced economies, with an intergenerational elasticity of income of 0.185 and a rank-rank slope of 0.215. This picture of mobility remains under a range of exercises designed to test traditional methodological concerns. While mobility is rapid through most of Australia, there is meaningful dispersion: the mining boom in particular appears to have lifted incomes for those raised in affected regions over the period in question. More generally, regions with higher incomes and lower unemployment rates tend to have higher expected ranks for those raised there; while regions with less segregation and higher school attendance rates have weaker intergenerational persistence in income ranks. We end by extending a generalised error-in-variables model to provide a framework for thinking about the stability of these regional mobility measures over time. In line with this model, regional rank-rank slopes steadily increase over the period we observe, while the expected ranks of children in the national income distribution fluctuate in ways that partly mirror the changing economic fortunes of Australian regions.

1. Introduction

Equality of opportunity is central to many conceptions of a just society, and a perennial touchstone in public policy debates. Accordingly, there is a vast related literature in economics on intergenerational mobility — the extent to which economic outcomes persist from one generation to the next (Black and Devereux, 2011; Solon, 1999). Increasingly studies explore regional differences in intergenerational mobility within countries.[2] These differences are of interest in their own right, but may also shed light on the mechanisms that drive differences in outcomes (for example, see Chetty and Hendren (2018) and Deutscher (2020)).

We provide the first national and regional estimates of intergenerational income mobility using Australian administrative data. Australia emerges as among the more mobile advanced economies – and hence a potentially informative example of high mobility outside the Nordic countries. While mobility is rapid across most of Australia, there are meaningful differences within the country, and its states, territories and cities. These regional measures correlate with a variety of potential mechanisms for mobility, from good school attendance to strong local labour markets. In the final section of the paper we explore the stability over time, in theory and in practice, of this "geography of mobility" that is an increasingly common feature of the literature.

* Corresponding author at: Crawford School of Public Policy JG Crawford Building 132, Canberra 2601, Australia.
E-mail address: nathan.deutscher@anu.edu.au (N. Deutscher).

[1] Nathan Deutscher was supported by a Sir Roland Wilson scholarship, funded by the Australian Treasury and the Australian National University. We would like to thank Julia Neville, Thomas Abhayaratna, Bruce Bastian, Andrew Carter, Matt Power and Julia Rymasz for their invaluable expertise, advice and support accessing the intergenerational data. We thank Bob Breunig, Jonathan Davis, Peter Siminski, Satoshi Tanaka and seminar and conference participants at the University of Queensland and Asian and Australasian Society of Labour Economics 2019 Conference for their helpful comments and suggestions. This research uses data from the Australian Taxation Office. All findings, opinions and conclusions are those of the authors and do not necessarily represent the views of the Australian Government or any of its agencies. The views expressed here do not reflect those of the Federal Reserve Bank of Chicago or the Federal Reserve system or the Australian Treasury. The research plan was approved by the Australian National University Human Ethics Committee, protocol number 2017/832.

[2] For example, see regional estimates of intergenerational mobility within the United States (Chetty et al., 2014; Davis and Mazumder, 2018), Canada (Connolly et al., 2019; Corak, 2019), England and Wales (Bell et al., 2018), Italy (Acciari et al., 2019; Güell et al., 2018), Sweden (Heidrich, 2017) and Norway (Bütikofer et al., 2018).

https://doi.org/10.1016/j.labeco.2020.101861
Received 30 January 2020; Received in revised form 25 May 2020; Accepted 27 May 2020
Available online 17 June 2020

Data constraints have limited the quantity and quality of Australian evidence on intergenerational income mobility to date. Most existing studies lack data on income for long periods in both generations, often leading researchers to impute parental income using methods that have required strong assumptions.[3] There is also no evidence on geographic differences in intergenerational mobility within the country. National estimates for Australia are interesting in their own right and, when compared with other countries, may also shed light on the value of certain institutional features of the country and how these may enhance or impede mobility. Regional differences within Australia may also provide insights as to how differences in industry composition, economic growth or other unique aspects of different regions may matter for mobility. A key obstacle to producing definitive mobility estimates for Australia has been the lack of the kind of large, long and representative panel dataset that administrative data can provide. Large panels enable researchers to produce more precise national estimates as well as estimates for smaller geographic regions. Long panels hold promise for more effectively addressing the variety of measurement challenges that is endemic to the intergenerational mobility literature (e.g. Haider and Solon (2006); Mazumder (2005); Solon (1992)).

We address these challenges by producing the first estimates of intergenerational mobility based on Australian tax data. We find that intergenerational persistence in income is quite low and that intergenerational mobility is consequently quite high. We estimate the intergenerational elasticity of income (IGE) for Australia to be 0.185 and the Spearman correlation (or rank-rank slope) to be 0.215. By way of contrast, (Chetty et al., 2014) estimate an IGE and rank-rank slope of around 0.34 for the United States; while more recent estimates addressing limitations in their data put the IGE at 0.5 or higher and the rank-rank slope at 0.4 or higher (Mazumder, 2016). Our estimates for Australia are more comparable to recent Canadian estimates of an IGE of 0.20 and a rank-rank slope of 0.24 (Corak, 2019). The estimates are also close to Nordic countries such as Denmark and Norway where estimates of these parameters are typically around 0.2 or lower. While precise rankings of countries are fraught given differences in data sources, definitions and panel lengths, these findings nonetheless suggest Australia is among the more mobile countries in the world.[4]

We also estimate other rank-based measures of intergenerational mobility, such as conditional expected ranks. We find that the expected rank of individuals whose parents were at the 25th percentile is the 45th percentile. Those who started at the 75th percentile could expect to land at the 56th percentile.[5] This illustrates substantial upward and downward mobility. In the United States there are only 53 cities out of 381 that have higher rates of upward mobility than Australia taken as a whole.

Our qualitative results do not change when adjusting for potential sources of bias, including measurement of income over too short a period or too early or late in life, nonlinearities in intergenerational relationships, missing income observations, and missing or incorrect parent-child links. Consistent with recent research abroad and in Australia, rank-based measures appear less sensitive to many of these concerns (Mazumder, 2016; Murray et al., 2018; Nybom and Stuhler, 2017). Our most conservative estimates using our baseline sample lift our estimated IGE to 0.210 (from 0.185) and our rank-rank slope to 0.232 (from 0.215).[6]

We also present the first regional estimates of intergenerational mobility across Australia. While there is significantly less dispersion in mobility across Australian regions relative to the United States, meaningful differences nonetheless emerge — both within the country and within individual cities such as Sydney and Melbourne. Perhaps most notable is the influence of the mining boom, which appears to have lifted the expected ranks (in the national income distribution) of children raised in resource-rich regions. The boom appears to have ‘lifted all boats’, as although upward mobility was higher in the mining regions, the rank-rank slope was no lower than that experienced by other regions. We explore a range of correlates of intergenerational mobility, finding that regions with higher incomes and lower unemployment rates tend to have higher conditional expected ranks for those raised there; while regions with less segregation and higher school attendance rates have lower rank-rank slopes.

Given the recent proliferation of regional estimates of intergenerational mobility, we end with an exploration of the stability of regional estimates of intergenerational mobility. To provide a framework for this, we extend a generalised error-in-variables model to the setting of regional rank-based estimates of mobility. The model highlights the potential sensitivity of regional estimates to local income shocks. These shocks flow directly through to conditional expected ranks but only influence the rank-rank slope to the extent they are correlated with parental income. This suggests regional estimates of conditional expected ranks may be more volatile than regional estimates of the rank-rank slope – a possibility that is borne out in the Australian experience. To illustrate this, we produce regional estimates of intergenerational income mobility from 2000 onwards. National estimates are significantly attenuated (show more mobility) towards the start of this period, but stabilise towards the end of our panel in 2015. Regional rank-rank slopes follow a relatively similar pattern, and most of their variation over time is well explained by a common attenuation process. However, conditional expected ranks are much more volatile, and move in ways that can be at least partly explained by fluctuations in the local labour market.

Regional estimates of intergenerational mobility – at times referred to as the “geography of mobility” (Chetty et al., 2014; Corak, 2019) – are of significant public interest, and can help shed light on the mechanisms underlying intergenerational mobility (Chetty and Hendren, 2018; Deutscher, 2020). Regional estimates are also increasingly produced using rank-based measures, which have been shown to stabilise more quickly at a *national* level with the respect to the ages and lengths of time over which incomes are observed. We provide an important caveat to this growing literature in highlighting that stable national estimates may mask underlying volatility in regional estimates, particularly conditional expected ranks, potentially driven by local economic fluctuations. Future studies may benefit from keeping this in mind in interpreting and exploiting the geography of mobility.

2. Data

We use new intergenerational data drawn from Australian federal income tax returns from 1991 to 2015. The Australian Taxation Office (ATO) has produced the data as an extension of its existing research files, the ATO Longitudinal Income Files.[7] Family links primarily come from linking children to adults living at the same address when the child registers for a Tax File Number (TFN): a unique personal identifier issued by the federal government. The algorithm for linking is also informed by

[3] Studies to date include Leigh (2007), (Mendolia and Siminski, 2016) and Murray et al. (2018). The last of these provides the first estimates of intergenerational income mobility for Australia where both parent and child incomes are observed directly. The earlier studies imputed parental income based on parental occupation. A recent study using administrative data is Cobb-Clark et al. (2017), though the focus here is on intergenerational welfare dependency rather than income mobility.

[4] For example, administrative and survey data have differing strengths and weaknesses, and the Nordic countries are among the few with panels long enough to estimate lifetime incomes for both generations.

[5] If there were “perfect” mobility such that everyone’s rank was randomly distributed, then the expected rank would be at the 50th percentile.

[6] The IGE does rise further, to 0.24, when excluding the bottom and top deciles of the parent income distribution, or focusing only on a much smaller set of children for whom parent and child incomes can be observed closer to mid-life.

[7] For further information on these files, see https://alife-research.app/.

Table 1
Family characteristics in sample and population.

Birth cohort	Full sample 1978-1982	Population Various
Family structure (%)		
Couple	84	81
Lone mother	11	15
Lone father	5	4
Median parental age at birth (years)		
Mother	27	26
Father	29	29
Family size (%)		
1	13	8
2	38	38
3	30	34
4	13	15
5	4	4
6	2	1
7 or more	1	1
Mean family size	2.7	2.8
Number of children	1,136,900	1,100,000
Number of children linked to parents	1,025,800	NA
Number of families	792,900	835,800

Notes: Population estimates are based on: Family Characteristics Survey 1997 (Australian Bureau of Statistics, 2010), (family structure, 1973-82 birth cohorts); Births, (Australian Bureau of Statistics, 2017b) (median parental age at birth, 1978-82 birth cohorts); and the 1991 Census, (Australian Bureau of Statistics, 1991) (family size, 1978-82 birth cohorts).

a subset of families in which children are directly claimed as dependents on tax returns.[8]

The universe for our baseline sample includes 1.1 million individuals born in Australia between 1 July 1978 and 30 June 1982 who registered for a TFN and remained resident in Australia through 2015. Of these 90% are linked to parents. This is comparable to the matching rate attained by Chetty et al. (2014) for their core sample of children born in the United States from 1980 to 1982. Our baseline sample closely mirrors population benchmarks for family structure, median parental age at birth and family size (Table 1). Compared to the population, our sample contains a slightly higher share of two parent families and a slightly lower share of families headed by single mothers. Our sample is also slightly skewed towards smaller families.

Our primary measure of income is individual total pretax income. This is the most commonly used income measure across Australian Bureau of Statistics household surveys and is commonly used in the literature (e.g. (Chetty et al., 2014)). In years where an individual has filed a tax return, this is their reported total income or loss. In years where an individual has not filed a tax return, it is the sum of individual salary and wages reported by employers, and taxable allowances, benefits and pensions reported by government welfare agencies, where available. This 3rd party information is only available in the latter half of our panel, and hence not included in our measures of parental income. Those who have no return or 3rd party information are recorded as having zero total pretax income. This income measure includes labour and capital income, and taxable government payments such as unemployment and study benefits. It is prior to any tax deductions or offsets.

Income variables are measured in 2015 dollars, adjusted for inflation using the headline consumer price index published by the Australian Bureau of Statistics (Australian Bureau of Statistics (2017c)).

In our baseline analysis, we calculate parental household income as the average of the combined annual income of the parent(s) over eleven years from 1991 to 2001. This choice of years balances the benefits of comparability with international studies of intergenerational mobility within a country (such as Chetty et al. (2014)), where shorter time periods are used, with the importance of averaging over many years to generate a better proxy for lifetime income (Mazumder, 2005; 2016). This choice also ensures parental income is in the vast majority of cases observed in the middle of the working life (in the 30s, 40s or 50s), given the distribution of parental age at birth, which closely mimics that seen in population data (Appendix Table B.1). Child household income is defined similarly as the average of the combined annual income of the child and their most recently reported spouse (as at 2015) over the five years from 2011 to 2015. We examine the sensitivity of our national estimates to both of these choices.

Average incomes in our sample tend to be slightly lower than those observed in population survey data, though typically only a few thousand Australian dollars less (Appendix Table B.2). This is consistent with the tax data being less comprehensive, even where 3rd party reported information is available. Whereas a negligible proportion of survey respondents has zero income in a given year, this ranges from 5 to 10% in our data. In contrast, the proportion of children with a spouse is consistently around 70% in both datasets. The largest discrepancy is observed is in the final year of the panel, when child household income averaged $94,600 in the tax data and $112,500 in the survey data. Our suspicion is that this reflects a small proportion of late tax filers being missed at the point in time when the data was extracted. We obtain similar estimates of mobility with a variety of different approaches to treating years with zero income, and when dropping the last year of the panel.

Finally, for the purposes of defining geography for our regional estimates of mobility, children are assigned to the first geographic location associated with their primary parent. These locations either arise from a geocoded address or a residential postcode for the parent. In both cases we assign children to the associated Statistical Area 4 (SA4), as defined in Australian Bureau of Statistics (2011). These SA4s delineate broad labour markets, and are the closest Australian analogue to the commuting zones of Chetty et al. (2014).

3. Methodology

In this section we describe the various measures of intergenerational mobility that we use for our national and regional estimates.

3.1. Intergenerational elasticity

The intergenerational elasticity (IGE) has been the most commonly used measure of intergenerational mobility in economics. The IGE characterizes the rate of intergenerational persistence in a particular outcome (measured in logs such as log income) and one minus the IGE can be viewed as a gauge of intergenerational mobility. The IGE is the estimate of β obtained from the following regression:

$$y_{1i} = \alpha + \beta y_{0i} + \varepsilon_i \tag{1}$$

where y_{1i} is the log of income in the child's generation and y_{0i} is the log of income in the parents' generation.[9] The estimate of β provides a measure of intergenerational persistence in log income and $1 - \beta$ can be used as a measure of mobility. A value of 0.2, for example, suggests that if the difference in income between two families is 10 per cent, in

[8] More details on the linking procedure and features of the data can be found in Deutscher (2020). Note there are strong institutional incentives to acquire a TFN well before adulthood – without one, individuals face higher income tax withholding rates and are unable to receive welfare payments or concessional loans for higher education. Most individuals in our birth cohorts get a TFN while young: of those born in Australia in the 1980 financial year and with a TFN by the time they were 30, over 90% had registered by age 17. As a result co-residency biases — which can arise when children are more likely to be observed and/or linked to parents if they live with their parents as adults — are unlikely to be significant in this setting.

[9] Often the regression will include age controls but few other covariates since β is not given a causal interpretation but rather reflects all factors correlated with parent income. We include financial year of birth dummies to control flexibly for the age of the child when income is measured.

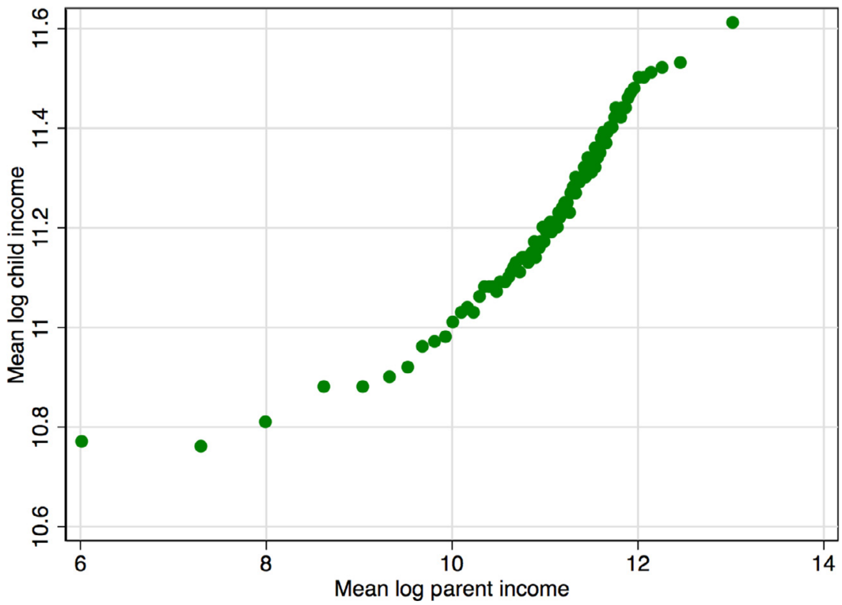

Fig. 1. Intergenerational income mobility - log of income
Notes: Chart plots the mean log child and parent total household income for each percentile of the parent income distribution.

the parent generation, then on average, approximately 2 per cent of this gap would be expected in the income of the children's generation.[10] An IGE of 0.2 would be indicative of a low degree of persistence and a fairly high degree of mobility compared to an IGE of 0.6.

3.2. Intergenerational correlation

In contrast to the IGE, the intergenerational correlation (IGC), or Pearson correlation, is calculated by first standardizing log incomes in both generations to have equal variance. This abstracts from any changes in the variance in log income across generations. Formally, the IGE is equal to the IGC times the ratio of the standard deviation of log income in the child's generation to the standard deviation of log income in the parents' generation:

$$IGE = IGC\frac{\sigma_{y_1}}{\sigma_{y_0}} \tag{2}$$

From this it is readily seen that if the standard deviation of log incomes in the child generation exceeds that in the parent generation ($\sigma_{y_1} > \sigma_{y_0}$) then it follows that $IGE > IGC$. This could result from a secular increase in inequality, but may also occur due to other differences between the parents and their children (for example, driven by when incomes are observed).

3.3. Rank-based measures

A closely related measure to the IGC is the rank-rank slope or Spearman correlation, which is obtained from the following regression:

$$r_{1i} = a + br_{0i} + \varepsilon_i \tag{3}$$

where r_{1i} and r_{0i} now represent the percentile rank in income in each respective generation. In this case, b provides an estimate of persistence in rank position and $1 - b$ provides a measure of positional mobility. In addition to estimates of rank persistence, following Chetty et al. (2014) we also use the rank-rank regression framework to calculate expected ranks at the 25th and 75th percentiles.[11] These statistics are useful for thinking about 'directional' mobility. For example, if the expected rank of individuals whose parents were at the 25th percentile is the 40th percentile then this would suggest average upward mobility of about 15 percentiles.

A key advantage of using the rank-based measures is that when using ranks based on the national income distribution, they can be used to make "apples to apples" comparisons of various subgroups of the population. Most notably for our purposes, we can compare regions within Australia to one another and be confident that our intergenerational rank mobility estimates mean the same thing in all places.

Finally, we also produce a matrix of transition probabilities across quintiles of the income distribution. This approach has also commonly been used to summarize intergenerational income mobility. Transition probabilities also provide measures of directional mobility as well as showing how mobility may differ at different points of the income distribution.

4. National estimates

We begin by showing some descriptive figures. Fig. 1 plots, for each percentile in the parent income distribution, the mean log of total household income of adult children against the mean log of total household income of their parents. What is immediately evident is that the relationship is nonlinear with a much flatter slope at lower and higher levels of parent income and a steeper slope in the middle of the distribution. This pattern is similar to that seen in Canada (Corak and Heisz, 1999) and the United States (Chetty et al., 2014), where the IGE is also highest in the interior of the income distribution.[12] To some extent this may be driven by a failure to capture full incomes in the parent generation – for example, for those not filing tax returns – as it seems unlikely that the bottom percentiles of parents have annual incomes of only a few hundred or thousand dollars (as implied by Fig. 1).

The patterns in the data shown in Fig. 1 suggest that in addition to estimating the IGE, which characterizes the entire distribution in one summary statistic, it is also useful to consider how mobility might differ at different points in the distribution, a point we return to when discussing nonlinearities and estimating transition probabilities. It also highlights the value in careful consideration of how estimates may vary

[10] Since the data is measured in logs, the difference in log income approximates the percentage difference in income.

[11] We present these two percentiles purely for descriptive purposes. Given the rank-rank slope and the intercept (or expected rank at any given percentile) it is of course possible to calculate the expected rank at any percentiles of interest.

[12] A similar pattern is observed through in Sweden, outside the top percentile, where the IGE again increases (Björklund et al., 2012). A similar, though less pronounced kick up in the top percentile is also apparent in Fig. 1.

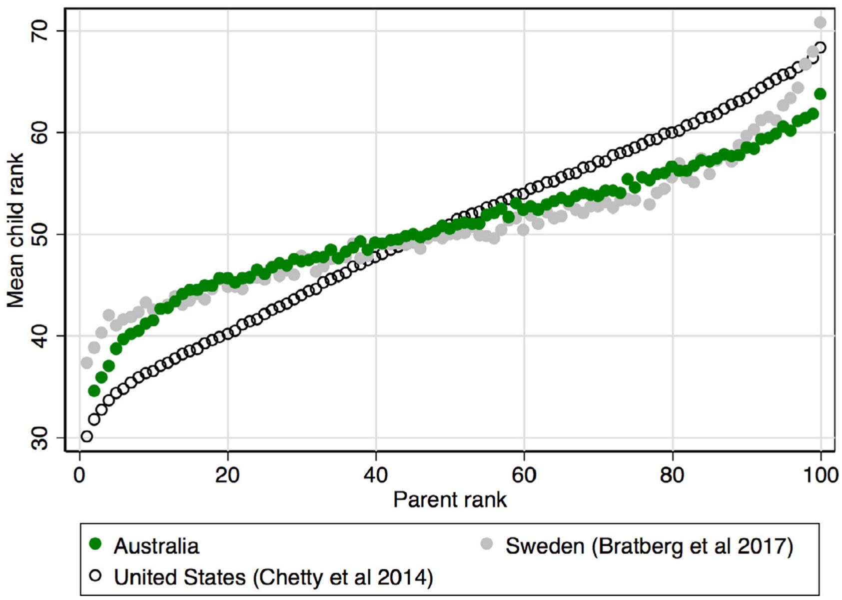

Fig. 2. Intergenerational income mobility – income ranks Notes: Chart plots the mean child and parent total household income rank for each percentile bin of the parent income distribution.

with the treatment of missing values, a point we return to later. An advantage of working with rank-based measures is the added robustness that comes with the resulting bottom-coding of low incomes (a point made in Dahl and DeLeire (2008) and Chetty et al. (2014)).

Fig. 2 plots intergenerational income mobility in terms of income ranks. Working with income ranks allows us to characterize regional upward mobility and downward mobility in a more readily interpretable way. In addition, the slope of this line, the rank-rank slope, provides an additional measure of intergenerational persistence. Fig. 2 also compares the rank-rank relationship in Australia to those in United States and Sweden — two advanced economies with vastly different experiences of mobility (results are from Chetty et al. (2014) and Bratberg et al. (2017)). This picture highlights different experiences of purely *positional* mobility – an important dimension of public discourse on equality of opportunity – as captured in rank-rank relationships.

Consistent with the existing literature, the intergenerational relationship is more linear in ranks than it is in log incomes (Chetty et al., 2014), making it easier to summarize in a single statistic. That said, there are interesting nonlinearities. In Australia, the rank-rank slope is flatter between the 15th and 95th percentile than it is below the 15th percentile and above the 95th percentile. A greater slope and hence greater persistence in outcomes towards the bottom of the parent income distribution is in fact a feature across all three countries. This may point to different mechanisms underlying entrenched disadvantage. Similarly, Sweden has a pronounced increase in the slope at the top of the distribution, which may again point to different mechanisms, such as intergenerational wealth transfers (Björklund et al., 2012). A comparison of the three countries suggest that expected ranks are highest for the bottom 10 per cent and top 1 per cent in Sweden, for the lower-middle income earners in Australia and for the upper-middle income earners in the United States. While the mapping of ranks to actual living standards and social standing will differ across all three countries, this nonetheless highlights some nuances that can be lost in comparing rank-rank slopes alone.[13]

In Table 2 we present our baseline national estimates of intergenerational income mobility. In Panel A we consider four different measures of income: wages and salary, all private income (including investment income), total income (including government transfers) and disposable income (total income income minus taxes). These different measures of income capture different concerns. Wage and salary income may be viewed as most indicative of earnings ability as it presumably reflects productivity and excludes passive sources of income. Private or market income includes passive sources of income as an additional mechanism by which advantage may be passed from one generation to the next. Finally total and disposable income will result in measures accounting for redistribution first through transfers (total income), and then through the tax system (disposable income). Disposable income might also be most reflective of consumption possibilities and therefore of welfare.

In column (1) we find that the estimates of the IGE range from 0.107 for wages and salary to as high as 0.192 for private income. Our estimate for total income is 0.185. The estimates for the IGC (Pearson correlation) shown in column (2) are consistently lower, ranging from 0.114 to 0.159. Given the mechanical relationship between the IGE and IGC, this implies that the variance in the income measures is higher in the child generation than in the parent generation. The IGC abstracts from this increase in inequality, while the IGE captures it and is higher as a consequence.[14] Finally, the estimates of the rank-rank slope (Spearman correlation) presented in column (3) range from 0.186 to 0.222. Our baseline estimate for the rank-rank slope in total income is 0.215. Across both the IGE and rank-rank slope, persistence is lowest when looking purely at wages and salaries, highest when looking at all private income and then progressively less for total and then disposable income.[15] This aligns with the typical effects of capital income, transfers and taxes on static

[13] In Appendix Chart A.1 we present an alternative mobility measure for the United States and Australia – the probability that child income exceeds parent income by at least 50%, after incomes are normalised for their generational mean. This is a measure of upward mobility that abstracts from economic growth. While a different construction of mobility, a broadly similar story emerges, with a higher probability of such upward transitions for lower income earners in Australia (roughly the bottom three deciles) and for middle and high income earners in the United States (roughly the top seven deciles). Interestingly, a similar pattern is seen in Italy (Acciari et al., 2019), where this comparison is also used.

[14] This increase in inequality could reflect, but need not imply, a secular increase in inequality. For example, the set of those who are parents in one generation may be more homogenous than the full set of offspring in the next. Child incomes are also measured over a different and shorter period of the working life, and may have higher variance for this reason.

[15] The elasticity estimates for wages and salaries may appear strikingly low, particularly in comparison to the international literature on earnings mobility. They are likely underestimates of true persistence in earnings since labour income, for the self-employed in particular, may appear under a number of different tax return labels. As a rough correction, if we restrict attention to those children and parents for whom wages and salaries constituted at least 80 per cent

Table 2
National measures of intergenerational income mobility.

	IGE	Pearson correlation	Rank-based		
			Rank-rank slope	$E[r_{1i} \mid r_{0i} = 25]$	$E[r_{1i} \mid r_{0i} = 75]$
Panel A: Income definition					
Wages	0.107	0.114	0.186	45.8	55.1
	(0.001)	(0.001)	(0.001)	(0.0)	(0.0)
Private	0.192	0.150	0.222	44.8	55.9
	(0.002)	(0.001)	(0.001)	(0.0)	(0.0)
Total	0.185	0.159	0.215	45.0	55.8
	(0.001)	(0.001)	(0.001)	(0.0)	(0.0)
Disposable	0.175	0.148	0.211	45.1	55.7
	(0.001)	(0.001)	(0.001)	(0.0)	(0.0)
Panel B: Household, own or spousal income					
Women – household	0.181	0.156	0.211	46.5	57.0
	(0.002)	(0.002)	(0.001)	(0.1)	(0.1)
Women – own	0.166	0.149	0.174	37.4	46.1
	(0.002)	(0.002)	(0.001)	(0.0)	(0.1)
Women – spouse	0.117	0.136	0.126	55.7	62.0
	(0.002)	(0.002)	(0.002)	(0.1)	(0.1)
Men – household	0.188	0.161	0.217	43.7	54.5
	(0.002)	(0.002)	(0.001)	(0.1)	(0.1)
Men – own	0.181	0.159	0.209	53.4	63.8
	(0.002)	(0.002)	(0.001)	(0.1)	(0.1)
Men – spouse	0.117	0.138	0.100	40.0	45.0
	(0.003)	(0.002)	(0.001)	(0.0)	(0.0)

Notes: Presents estimates of five different measures of intergenerational persistence for different income definitions and units of observation. The default is to estimate using total income at a household level for the full sample. In Panel A we vary income only: wages income is the self- or third-party reported individual salary and wages; private income is total income minus self- or third-party reported government payments; total income is as defined in the text; and disposable income is taxable income minus gross tax. In Panel B we split the sample into women and men (based on child gender) and vary whether the child's adult household income, own income or spouse income is the outcome of interest.

measures of inequality. These various measures however, all tell a fairly consistent story that intergenerational persistence in Australia is quite low with estimates typically around 0.2 or lower. By way of contrast, estimates of intergenerational persistence in total income in the United States are typically around 0.5 or higher (Mazumder, 2016).

In columns (4) and (5) we present estimates of the expected rank at the 25th and 75th percentiles to illustrate the movements in income ranks implied by the high level of intergenerational mobility. We find that upward mobility is quite high as children whose parents were at the 25th percentile can expect to rise nearly to the median at the 45th percentile. Similarly those born into the 75th percentile can expect to fall to the 56th percentile. Naturally, behind these average experiences there is a diversity of outcomes, which we will later explore through the probabilities of transitioning between given quintiles of the income distribution.

In Panel B we consider how the estimates differ by the gender of the child, and whether it is the child themselves, their spouse or their combined outcomes that are considered.[16] This allows us to consider the role of household formation in driving the results. For both women and men, measures of persistence based on individual income are modestly lower when looking at their individual income, rather than their household income. This gap is slightly larger for women, pointing to a slightly greater role for household formation in driving their observed household outcomes. Nonetheless, the differences between the genders are mostly modest. While women have lower expected ranks in the income distribution, the relationships between their income levels and position in the income distribution and those of their parents are similar to those of men. There is only a slightly stronger connection between parent and child outcomes for men versus women, and the connection between parent and child spouse outcomes are fairly similar. This finding contrasts with the United States literature, which has found more substantial differences, with lower persistence in outcomes for women, and driven more through household formation, though these differences tend to be smaller in studies examining more recent birth cohorts (Chadwick and Solon, 2002; Chetty et al., 2014; Mitnik et al., 2015).[17] In Appendix Table B.3 we provide estimates based on individual parent-child pairings – father-son; father-daughter; mother-son; and mother-daughter – persistence is generally greatest between same-sex pairings.

Finally, for robustness, we also estimated the IGE and the other parameters in total income using several other approaches shown in Appendix Table B.4. First, given the nonlinearities at the tails of the income distribution, we produced estimates just using the middle 80 percentiles (10th to 90th percentile). As expected based on Figs. 1 and 2, this boosts our estimate of the IGE slightly higher to 0.241 and lowers our estimate of the rank-rank slope to 0.181. We also produce estimates where we use inverse probability weights to account for differences in the probability that a child is successfully linked to their parents.[18] This has a relatively modest effect on our results producing an IGE of 0.191 (up

of their total income, the intergenerational elasticity rises from 0.107 to 0.131. A more comprehensive measure of earnings mobility is outside the scope of this paper, but could potentially draw on the net income from working measure that underpinned the (now abolished) Mature Age Workers Tax Offset, a targeted Australian earned income tax credit discussed in some detail by Carter and Breunig, 2019.

[16] Those without a spouse are coded as having zero spouse income, which means they are dropped in the IGE and Pearson correlations, but included in the rank specifications.

[17] The most recent and precise estimates of mobility to date in Australia (Murray et al., 2018) do not estimate mobility by gender due to sample size constraints.

[18] To do this we first calculate the percentile income ranks for all children, including those not linked to parents. We then calculate for each percentile bin the proportion of children who are linked to parents. The inverse of this provides the weight for the subsequent regressions. More complex approaches are possible, accounting for a wider set of potential covariates (for example, sex and location). However, this risks a false sense of precision given the inability

Table 3
Intergenerational transition matrix.

		Parent quintile				
		1	2	3	4	5
Child quintile	5	12.3	15.9	18.6	22.5	**30.7**
	4	15.5	19.0	21.1	**22.6**	21.9
	3	18.5	21.1	**21.5**	21.0	18.0
	2	22.7	**22.3**	20.9	18.7	15.4
	1	**31.0**	21.8	18.0	15.3	14.0

Notes: Shows the per cent frequency with which a child with parents in a given income quintile (column) ends up in given income quintile (row) themselves. The main diagonal is shaded grey, with figures in bold.

from 0.185) and a rank-rank slope of 0.217 (up from 0.215). Finally, we produce a set of estimates using only the links we are most confident about, and weighting on the probability that a child has one of these high quality parent links.[19] Incorrect parent-child links may be expected to bias down our mobility estimates, but again the changes are relatively modest and in varying directions, with the IGE rising to 0.195 but the rank-rank slope falling to 0.208. In Appendix Appendix C we explore in more detail the treatment of missing values for either parent or child incomes – whether they are bottom coded to some minimum plausible annual income, dropped whenever they occur, or dropped if they occur across all the years of measurement. As expected, the rank-rank slope changes little with these choices, while the IGE is much more sensitive. Nonetheless, when consistent treatment is applied to parents and children the IGE reaches at most 0.24, which still stands as a relatively high degree of mobility.

As noted earlier, one disadvantage of summary statistics of intergenerational income mobility is the loss of finer detail about underlying movements. A world in which all children born into the 25th percentile end up at the 45th percentile is very different from one where this is simply the average across outcomes that span the full income distribution. A common way to capture this nuance is to examine transition probabilities.[20] Table 3 presents the probability a child born into a given quintile of the parent income distribution transitions to each quintile in the child adult income distribution. While the most common outcome is that a child stays in the same quintile they were born into, there are large proportions moving both up and down. The transition probability of 12.3 per cent from the bottom quintile to the top quintile again marks Australia out as among the more mobile of the advanced economies. An Australian child born into the bottom quintile is over 60 per cent more likely to reach the top quintile than a child born in the United States (where the transition probability is 7.5 per cent). Further, while Denmark has more mobility than Australia as measured by a rank-rank slope of 0.180 relative to 0.215, they have slightly less upward mobility on this measure with a probability of transition from the bottom to the top quintile of 11.7 per cent (Boserup et al., 2013). Even finer detail on transitions is captured in a ventile transition matrix in Appendix Fig. A.3. Again, all transitions are realised – with a transition probability of 1.9 per cent from the bottom ventile to the top ventile. However, the stickiness of the bottom and top of the income distribution is now more apparent, with 14.6 per cent of those born into the top ventile remaining there.

to weight on the (unobserved) joint distribution of child and parent incomes, and given this limitation we have not conducted further weighting exercises.

[19] The ATO data includes a variable that captures the quality of the parent-child link on the interval [0,1]. We include only those links for which this is 0.95 or greater, which drops around 10 per cent of the sample.

[20] Another approach to illustrating the distribution of child outcomes across the parent income distribution is to replicate the plot of mean child rank by parent rank in Fig. 2, but present percentiles of the child rank outcomes instead. Appendix Fig. A.2 provides such an illustration.

4.1. Lifecycle and attenuation biases

As the intergenerational mobility literature has matured, increasing attention has been paid to the importance of the length of time and ages over which incomes are observed. Intergenerational mobility can be greatly overstated if measured over too short a period, or too early or late in the lifecycle (Haider and Solon, 2006; Mazumder, 2005; Solon, 1992). These are frequently referred to as attenuation and lifecycle biases respectively.

In Figs. 3a and 3b we show the influence of the age and window over which parent or child incomes are measured for the IGE and the rank-rank slope. While the age and window of observation is varied for one generation it is held fixed for the other. Each series is centered on a given year of observation, with the length of the window of observation increasing along the x-axis. By centering on a given year of observation, we fix the average age at which child or parent income is measured, allowing us to consider the influence of age at and length of observation separately. A number of points can be made.

First, persistence is higher when income is measured in mid-to-late working life. For children, measuring income in a window centered later in life (in their 30s in this case) yields notably higher measures of persistence. For parents, measuring income in a window centered on their late 40s to early 50s produces the highest measures of persistence. This latter result contrasts slightly with international findings that have suggested measuring earnings at slightly earlier ages. For example, Mazumder (2005) notes that from the late 40s incomes tend to become more volatile, introducing a potential downward bias to measures of persistence. However, such findings are typically for much earlier birth cohorts, and the results here may be consistent with the extension of working lives. In addition, unlike most detailed examinations of attenuation and life-cycle biases, we present results based on both parents' incomes — not just fathers. The incomes of women later in life may well be a better reflection of the endowments passed on to their children than those earlier in their working lives when child care may limit their labor force participation.

Second, persistence is higher when income is observed over a longer period of time. While this is particularly true for parents, it also holds for children, pointing to the existence of non-classical measurement error (Nybom and Stuhler, 2017).[21] It remains common for mobility studies to concern themselves primarily with the length of time over which parent incomes are observed, and the age at which child incomes are observed. The results in Figs. 3a and 3b suggest the influence of length and age should be examined thoroughly for both generations.

Finally, comparing the two measures reveals some interesting differences. As has been noted in past work, the rank-rank slope is less sensitive to the age and window over which incomes are measured (Mazumder, 2016; Nybom and Stuhler, 2017). However, when faced with the trade-off between age of measurement and window width, the rank-rank slope tends to rise more with window width — the rank-rank slope is greatest when parent income is measured over the full 25 years, even though the average parental age will be 63 at the end of this window. Similarly, the rank-rank slope is greatest when child income is measured over 9 years, even though this includes incomes observed in their 20s. On the other hand, the intergenerational elasticity estimates are greatest for somewhat shorter windows that are centered closer to the middle of parent and child working lives.

A remaining concern regarding our national estimates may be that while the individual potential biases may be small, their cumulative effect may be more significant. In Appendix Table B.5, we present a set of 'conservative' estimates, where we measure parent income over the full 25 years (centered on an average age of 51) and child income over 9 years (centered on an average age of 31) as well as weighting for

[21] As classical measurement error in the left hand side variable would not bias the estimated coefficients.

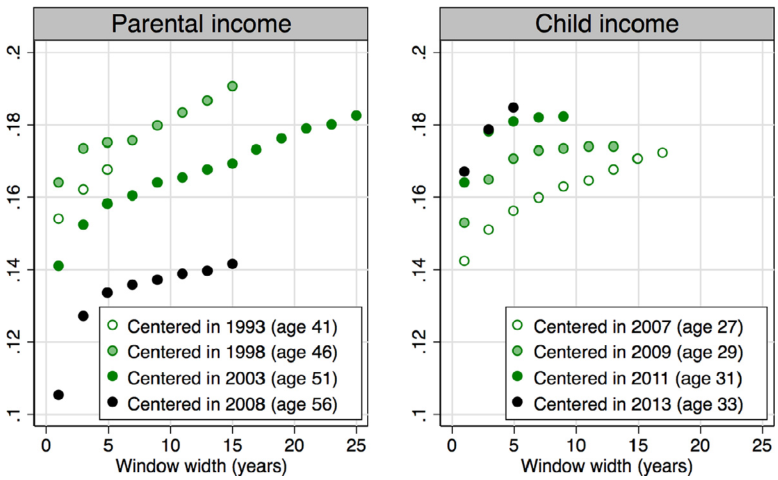

Fig. 3a. Lifecycle and attenuation biases in the IGE.

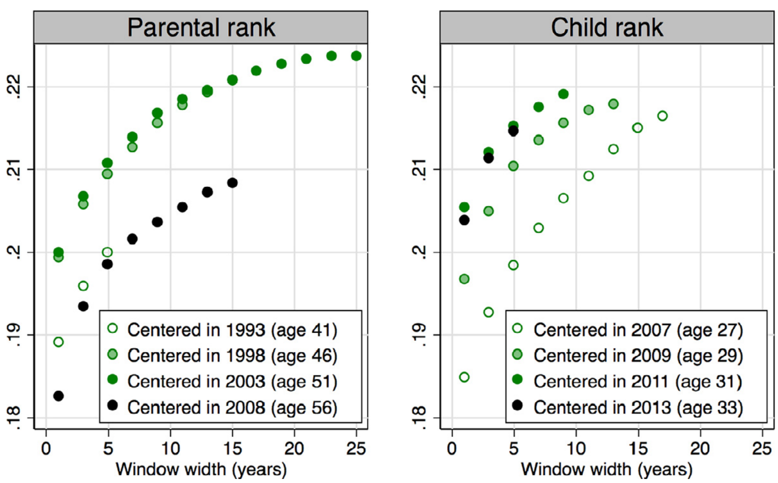

Fig. 3b. Lifecycle and attenuation biases in the rank-rank slope.
Notes: Presents estimates of the IGE and rank-rank slope, varying the center year and width (in years) over which parent or child log incomes or income ranks are observed. The corresponding average ages of the parents and children are shown in brackets.

the probability of inclusion in the sample and, in the second row, using only the highest quality links. We also repeat this exercise with the slightly smaller windows, and ages closer to the mid-life, that appear better suited to the IGE. The IGE estimates range from 0.188 to 0.210 (compared to 0.185 in Table 2), while the rank-rank slope ranges from 0.220 to 0.232 (compared to 0.215 in Table 2). Once again, these estimates place Australia among the more mobile advanced economies.[22]

It is worth noting that our estimated levels of persistence are somewhat less than the current benchmark estimates for Australia. Murray et al. (2018) estimate an intergenerational elasticity of 0.28 (s.e. 0.05) and a rank-rank slope of 0.27 (s.e. 0.05) using the Household Income and Labour Dynamics in Australia (HILDA) survey. They note their elasticity estimate in particular is likely biased down as parent income is only observed over five years. These are the first direct estimates of intergenerational mobility in Australia and improve upon earlier estimates for which parent income had to be imputed (Leigh, 2007; Mendolia and Siminski, 2016). Nonetheless, the estimates remain imprecise and subject to potentially complex attrition biases. For example, it is plausible that more mobile children — be it upwardly or downwardly mobile — are more likely to be lost in surveys, which would bias upwards the estimated persistence. For these reasons, at a minimum, we favor using our estimates when comparing Australia to other countries' mobility estimates based on administrative data.[23] It would be useful for future research to better explore the differences between estimates produced using survey and administrative data in Australia.

5. Regional estimates

We now present the first regional, rather than purely national, estimates of intergenerational mobility for Australia. As noted earlier, variations in intergenerational mobility within nations are increasingly being explored in the literature, with an eye to shining a light on potential transmission mechanisms — for example, in the work of Chetty et al. (2014) and Davis and Mazumder (2018) for the United States; Corak (2019) and Connolly et al. (2019) for Canada; Bell et al. (2018) for England and Wales; Güell et al. (2018) and Acciari et al. (2019) for Italy; Heidrich (2017) for Sweden; and Bütikofer et al. (2018) for Norway.

In this section we focus on two measures of intergenerational mobility — the rank-rank slope and the expected rank at the 25th percentile. As is standard in the literature, ranks are measured with respect to the

[22] An alternative exercise is to restrict the sample to those children and parents where income is measured closest to mid-life. To this end Panel C in Appendix Table B.5 restricts the sample to those born in 1978 and whose parents were aged 40–55 when incomes are observed (1991–2001) and had at least five non-missing income observations. Further, child incomes are only measured over the last three years of the panel (2013–2015). The resulting IGE is higher (as high as 0.24) but the rank-rank slopes are lower (as low as 0.19): a mixed picture that may be as much driven by the dramatic change in the sample (now a tenth of the full sample).

[23] We are not aware of any systematic comparison of mobility estimates derived from survey data with estimates derived from administrative data.

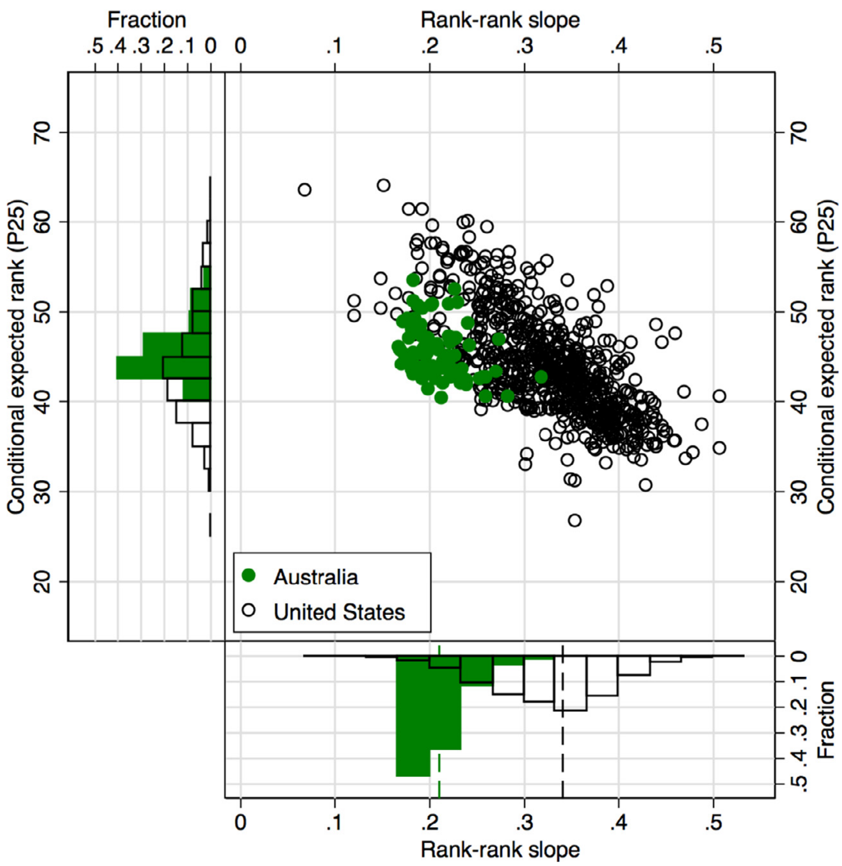

Fig. 4. Distribution of intergenerational mobility measures across regions in Australia and the United States Notes: Presents estimates of the expected rank, conditional on being born into the 25th percentile of the national parent income distribution, against the rank-rank slope for similarly-sized regions in Australia (87 regions) and the United States (741 regions). A scatter plot of the joint distribution and histograms of the individual distributions of the two intergenerational mobility measures are shown.

national income distribution. Appendix Table B.8 presents these estimates for all 87 Australian regions, alongside the sample size, and estimates of the intergenerational elasticity and the transition probability from the bottom to the top quintile. In future work, we will consider a still broader set of mobility measures – the focus in this paper is on those measures most commonly reported in the literature.

Fig. 4 presents three charts characterising intergenerational mobility for regions in Australia (shown in green) and the United States (shown in black, and based on Chetty et al. (2014)). The central chart shows a scatter plot of the expected national rank of a child born to parents at the 25th percentile against the rank-rank slope in the region. These two measures are negatively correlated in both countries (but positively correlated in the sense that more mobility on one measure is correlated with more mobility on the other). In the histograms we show the dispersion in the individual measures. While there is a notable dispersion in the Australian estimates, it is much less than that seen in the United States. For example, a child born to parents at the 25th percentile in a mobile Australian region (at the 90th percentile of regions ranked by mobility) can expect to end up only 8 percentile rank points higher than if they were born in an immobile Australian region (at the 10th percentile of regions). For the United States, the gap in expected outcomes for a poor child between high and low mobility regions is nearly double this, at 15 percentile rank points.

The fact that there is less variation in mobility within Australia than within the United States is perhaps unsurprising for two reasons. First, Australia is a more centralised federation than the United States, with less geographic variation in policies that might influence mobility. Second, Australia's high level of mobility may mean more regions are near upper bounds on mobility determined by factors such as levels of assortative mating and the heritability of income-earning potential. This latter possibility would be consistent with the histogram of the rank-rank slopes accompanying Fig. 4, where there appears to be a missing left tail of regions where the rank-rank slope is in the low 0.10s.[24]

Fig. 5 maps intergenerational mobility for Australia and its two largest cities — Sydney and Melbourne. Once again we can see variation in the estimates, within the nation and the cities, though there appears to be less dispersion in mobility within Melbourne. An interesting set of contrasts between the two maps are the high expected national ranks for children from poor families in some regional areas of Queensland and Western Australia that have fairly unexceptional rank-rank slope mobility. (These states occupy the north-eastern corner and western third of Australia respectively.) This likely reflects the influence of the mining boom driving strong local labour markets in these regions over the period of observation. Norwegian research has shown that resource shocks can improve outcomes for children for poor families (Bütikofer et al., 2018). A similar mechanism may well be at play here. Finally it is worth noting that this high-level visual 'ranking' of regions implied by these maps is robust to accounting for the uncertainty in the underlying metrics (though many fine distinctions are not). In Appendix Fig. A.4 we use the approach introduced in Mogstad et al. (2020) to identify those regions that are, with 95% confidence, in the bottom or top

[24] The narrower dispersion of Australian regional estimates of mobility seems highly unlikely to be driven by differences in the regional disaggregations. For example, the regions have similar population sizes, on average, as can be seen by the fact that the United States, roughly an order of magnitude more populous than Australia, has roughly an order of magnitude more regions. A more detailed comparison of United States commuting zones and Australian SA4 can be found in Deutscher (2020). Perhaps the most notable difference is that the cities are only ever represented by a single commuting zone in the United States (e.g. New York is a single commuting zone) whereas they may contain several SA4 in Australia.

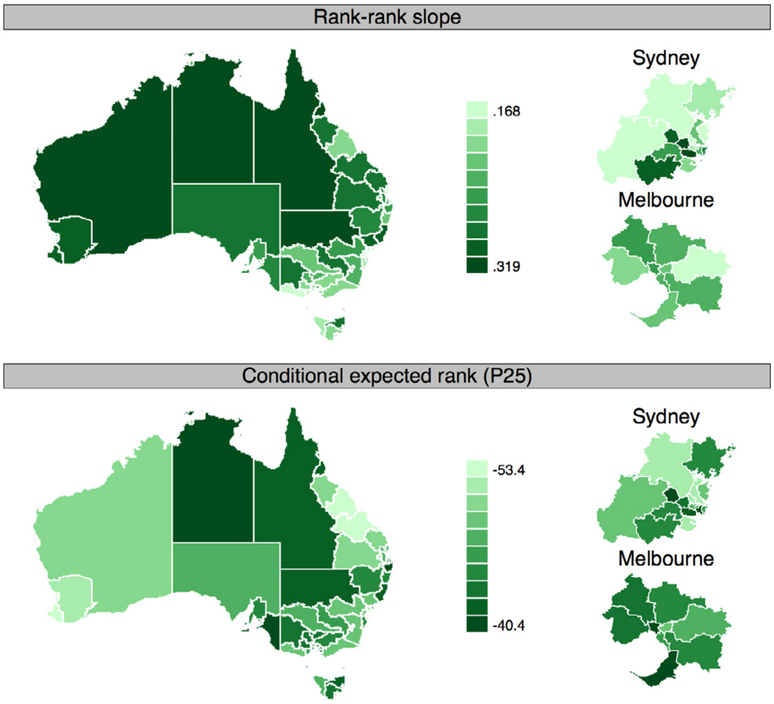

Fig. 5. Intergenerational mobility within Australia and its two largest cities Notes: Regional (SA4) estimates of the expected rank, conditional on being born into the 25th percentile of the national parent income distribution, and the rank-rank slope. Maps for Australia and its two largest cities, Sydney and Melbourne, are shown.

Category	Correlate	IGE	Rank-rank slope	$E[r_{1i}\|r_{0i}=25]$	$E[r_{1i}\|r_{0i}=75]$	$P[r_{1i}>80\|r_{0i}\leq 20]$
Diversity and distance	Frac. Indigenous	**0.64**	**0.60**	-0.20	0.10	-0.20
	Ethnic diversity	-0.04	-0.06	-0.02	-0.05	**0.29**
	Ethnic segregation	**0.41**	**0.37**	**-0.25**	-0.07	-0.06
	Income segregation	**0.49**	**0.36**	-0.08	0.10	0.06
	Mean commute	-0.06	0.05	0.08	0.11	-0.13
Labour market	Mean income	0.05	-0.17	**0.49**	**0.41**	**0.76**
	Unemp. Rate	0.05	**0.35**	**-0.42**	**-0.26**	**-0.50**
	NILF rate	0.16	**0.44**	**-0.27**	-0.06	**-0.47**
Education	School attendance	**-0.54**	**-0.60**	0.05	**-0.25**	0.05
	Tertiary attainment	-0.21	**-0.34**	0.00	-0.17	**0.30**
Inequality	Gini	-0.12	-0.21	**0.22**	0.12	**0.40**
	Top 1% share	-0.15	**-0.31**	0.10	-0.06	**0.41**
Social	Frac. Volunteers	-0.02	-0.11	0.07	0.02	-0.10
	Frac. Religious	0.10	**0.22**	-0.03	0.08	-0.07
	Frac. Married	**-0.28**	**-0.31**	**0.40**	**0.25**	0.16

Fig. 6. Regional correlates of intergenerational mobility metrics Notes: Shows the bivariate correlation between the given correlate and intergenerational mobility metric. Darker shading indicates larger absolute values. Correlations significant at 5% level are bolded. Details on construction of the correlates are in Appendix D.

half of Australian regions when ranked on these measures of mobility. Once again, several regions in Queensland and Western Australia are notable for their high expected ranks.

5.1. Regional correlates of intergenerational mobility

A natural next step is to look at correlations between regional mobility and commonly posited explanations for intergenerational persistence and its reverse. In Fig. 6 we display the bivariate correlations across the 87 regions and such a set of correlates (described in detail in Appendix D).

The strongest correlation for the two slope coefficients – the IGE and rank-rank slope – are with the fraction Indigenous and school attendance. Regions with a higher fraction of Indigenous Australians and lower school attendance have greater intergenerational persistence.[25] In contrast, the strongest correlations for the metrics relating to expected ranks or transition probabilities are with mean income in the region. In part this likely reflects the fact that many of those raised in a region will end up living there and hence benefiting from any associated earnings premium – less than a third of those in our sample are filing from a different region in the final year of our sample, and less than a tenth are filing from a different state (Appendix Table B.6). Interestingly, the correlations between the expected ranks and the factors highlighted in

[25] This need not imply greater intergenerational persistence among Indigenous Australians – whom we are unable to identify in our data – for example, (Davis and Mazumder, 2018) find that low mobility in the south east of the United States is driven by low mobility whites.

Chetty et al. (2014) – segregation, income inequality, schools, social capital and family stability – are much weaker or nonexistent. Some of this likely reflects further genuine differences in intergenerational mobility between the two countries, though measurement issues and omitted variables complicate such comparisons.[26]

Finally, we test an alternate explanation for differences in regional estimates of intergenerational mobility – differences in local price levels. The concern here is that if higher incomes in the child generation in a region are also accompanied by higher prices, then mobility in real incomes may be overstated. To examine whether this influences the regional patterns observed above we re-estimated our regional estimates of mobility having first adjusted parent and child incomes for local price levels based on Australian Bureau of Statistics (2017c).[27] This adjustment has very little impact on our estimates. The correlation between our baseline and local-price-adjusted measures is over 0.99 for both the rank-rank slope and expected rank at the 25th percentile. This provides some comfort that, as in other comparable studies (e.g. Chetty et al. (2014)), local price variations do not drive the patterns observed here.[28]

What then will we be able to learn about the *causal* mechanisms underlying intergenerational mobility from variations across Australian regions? A threshold question is the extent to which these variations arise from the influences of the regions themselves versus simple differences in the types of families that live there. To this end, (Deutscher, 2020) replicates and extends the work of Chetty and Hendren (2018) in the Australian setting, finding that most of the differences in mobility across Australian regions is indeed causal. A child moving at birth between two Australian regions can expect to pick up around 70 per cent of the gap between the outcomes of permanent residents of those regions who have the same birth cohort and parent income rank. Further analysis of the regional estimates presented here may thus shed helpful light on the mechanisms underlying intergenerational mobility.

5.2. The stability of regional estimates of intergenerational mobility

Even where differences in regional estimates of intergenerational mobility seem likely to reflect a causal effect of place, the precise mechanisms and persistence of these place effects remains of significant interest. When can regional estimates of mobility be viewed as reliable estimates of lifetime mobility outcomes?

As noted earlier, a significant body of work has considered how proxying lifetime income by income observed too early or late in life, or over too short a window, can bias *national* measures of intergenerational mobility. Given the increasing array of subnational estimates of intergenerational mobility, it seems prudent to consider how sensitive they may be to such measurement error. To date, explorations of measurement error in the intergenerational mobility literature have implicitly considered national estimates (Böhlmark and Lindquist, 2006; Haider and Solon, 2006; Mazumder, 2005; Nybom and Stuhler, 2017; Solon, 1992). In the following section we extend a generalised error-in-variables framework to the regional estimation of (national) rank-based measures of intergenerational mobility.

5.2.1. An error-in-variables framework

We begin with the assumption that rather than measuring the lifetime income rank of a child i born in location j (r_{1ij}), we instead capture an income rank (r^*_{1ijt}) over some shorter time period t. Consider the linear projection of the observed income rank onto the lifetime income rank:

$$r^*_{1ijt} = \lambda_t r_{1ij} + u_{1ijt} \tag{4}$$

This is simply the generalised error-in-variables model in Haider and Solon (2006), with income ranks rather than log incomes as the variable of interest. This model's implications for national measures of mobility are explored in Nybom and Stuhler (2017). As noted by Nybom and Stuhler (2017), the linear functional form of this model closely matches the observed deviations from lifetime income ranks observed in their Swedish data: they also observe that, due to the use of income ranks, we must have $\lambda_t \leq 1$. With percentile income ranks we can make use of the fact that $\bar{r}^*_{1ijt} = \bar{r}_{1ij} = 50.5$ (and hence $\bar{u}_{1ijt} = (1-\lambda_t)50.5$) to rewrite this as:

$$r^*_{1ijt} = \lambda_t r_{1ij} + (1-\lambda_t)50.5 + v_{1ijt} \tag{5}$$

where v_{1ijt} is mean zero, at a *national* level. We assume parent lifetime income rank (r_{0ij}) is observed – both to simplify the analysis and as data limitations are often felt most keenly for the child generation.

The question now is how rank-based measures of intergenerational mobility with child outcomes measured over the shorter period t relate to those based on lifetime income ranks. We denote the estimated rank-rank slopes based on observed and lifetime incomes as $\hat{b}_{jt}$ and $\hat{b}_j$ respectively, and the analagous conditional expected ranks for a child born to parents at percentile p as $\widehat{CER^p}_{jt}$ and $\widehat{CER^p}_j$. First, for the rank-rank slope we have:

$$\begin{aligned}\hat{b}_{jt} &= \frac{\mathrm{Cov}\left(r_{0ij}, r^*_{1ijt}\right)}{\mathrm{Var}\left(r_{0ij}\right)} \\ &= \lambda_t \hat{b}_j + \frac{\mathrm{Corr}\left(r_{0ij}, v_{1ijt}\right)\sigma_{v_{1ijt}}}{\sigma_{r_{0ij}}}\end{aligned} \tag{6}$$

While for the conditional expected rank for a child with parents at rank p we have:

$$\begin{aligned}\widehat{CER^p}_{jt} &= \bar{r}^*_{1ijt} + \hat{b}_{jt}\left(p - \bar{r}_{0ij}\right) \\ &= \left(\lambda_t \bar{r}_{1ij} + (1-\lambda_t)50.5 + \bar{v}_{ijt}\right) + \left(\lambda_t \hat{b}_j + \frac{\mathrm{Corr}(r_{0ij}, v_{1ijt})\sigma_{v_{1ijt}}}{\sigma_{r_{0ij}}}\right)\left(p - \bar{r}_{0ij}\right) \\ &= \lambda_t\left(\bar{r}_{1ij} + \hat{b}_j(p - \bar{r}_{0ij})\right) + (1-\lambda_t)50.5 + \bar{v}_{ijt} + \frac{\mathrm{Corr}(r_{0ij}, v_{1ijt})\sigma_{v_{1ijt}}}{\sigma_{r_{0ij}}}(p - \bar{r}_{0ij}) \\ &= \lambda_t \widehat{CER^p}_j + (1-\lambda_t)50.5 + \bar{v}_{ijt} + \frac{\mathrm{Corr}(r_{0ij}, v_{1ijt})\sigma_{v_{1ijt}}}{\sigma_{r_{0ij}}}(p - \bar{r}_{0ij})\end{aligned} \tag{7}$$

To develop an intuition for these relationships it is instructive to consider the case where the measurement error in child income rank is uncorrelated with parent income rank, that is, $\mathrm{Corr}(r_{0ij}, v_{1ijt}) = 0$. In this case equations (6) and (7) simplify to:

$$\hat{b}_{jt} = \lambda_t \hat{b}_j \tag{8}$$

$$\widehat{CER^p}_{jt} = \lambda_t \widehat{CER^p}_j + (1-\lambda_t)50.5 + \bar{v}_{ijt} \tag{9}$$

Equations (8) and (9) have a simple and intuitive interpretation. For the rank-rank slope, measuring child income over period t rather than over the lifetime simply attenuates regional estimates by a common factor λ_t. For the conditional expected rank, the regional results are both attenuated, reflected in the weighted sum of the lifetime estimate and

[26] For example, we do not have access to regional measures of school quality. In Appendix Table D.2 we examine the multivariate correlations between regional measures of intergenerational mobility and explanatory factors. The associations between the fraction Indigenous and intergenerational persistence, and higher incomes and higher conditional expected ranks, remain. However, due to the small number of regions and large number of correlates we do not put much weight on this analysis.

[27] In particular, consumer price indices are available for all eight capital cities, and we apply these to all regions in the relevant state or territory. Parent incomes are converted to national 2014-15 dollars based on the average index applying over the period parent incomes are observed, in the primary parent's first state or territory of residence. Child incomes are converted similarly based on their state or territory of residence in each year of observation.

[28] Chetty et al. (2014) conduct a similar exercise (see Online Appendix Table VII, row 16). They find a correlation between their baseline estimates and local-price-adjusted measures of 0.98 for the conditional expected rank of a child born to parents at the 25th percentile and 0.99 for the rank-rank slope.

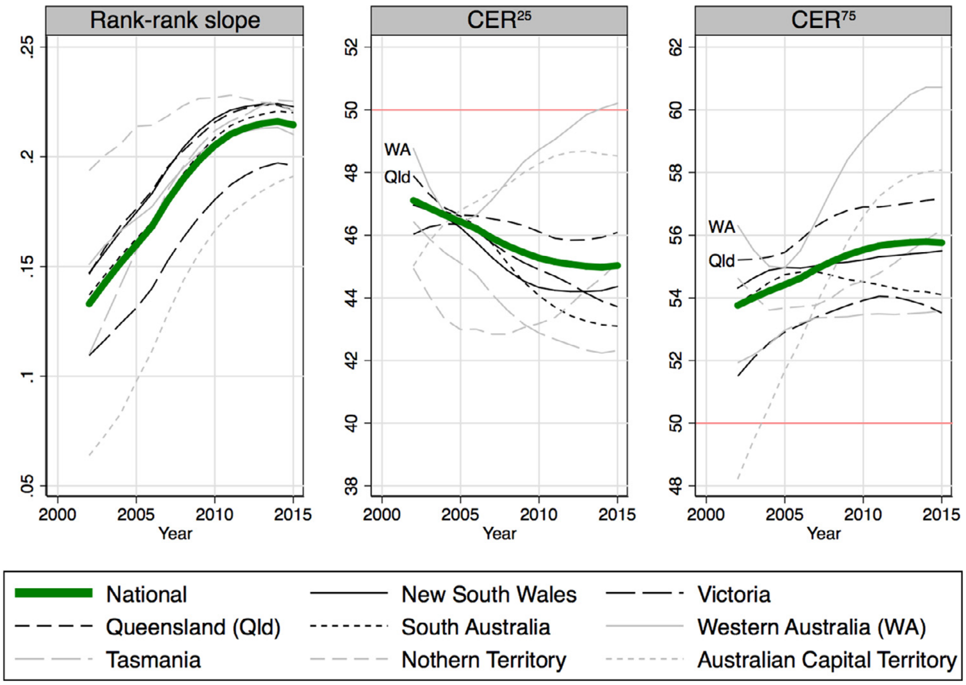

Fig. 7. National and state estimates of the rank-rank slope and conditional expected ranks by final year of observation Notes: Presents the evolution of national and state/territory estimates of the rank-rank slope and conditional expected ranks of a child born to parents at the 25th or 75th percentiles of the national income distribution. Estimates are based on child income observed over a five-year window ending in the given year. The final year corresponds to the baseline estimates presented in earlier sections.

the mean percentile rank (50.5), and shifted, by $\bar{v}_{ijt}$. While $\bar{v}_{ijt} = 0$ at the national level, this equality need not hold within regions. Regional estimates of the conditional expected rank will also capture any average tendency towards higher or lower income ranks (relative to lifetime ranks) in the region over the period t. Local economic shocks — booms or busts — are a possible driver of such tendencies. Importantly, even when $\lambda_t = 1$ and hence national estimates of intergenerational mobility are measured without error, regional estimates of the conditional expected rank will still be shifted by $\bar{v}_{ijt}$ over the relevant period. Stability in national estimates of intergenerational mobility thus need not imply stability in regional estimates.

In the more general case, when we allow for $\mathrm{Corr}(r_{0ij}, v_{1ijt}) \neq 0$, then regional rank-rank slopes and conditional expected ranks experience a further shift, and one not necessarily uniform across regions, as described in equations (6) and (7). There can be good reasons to believe that this term may be nonzero — for example, as noted by Nybom and Stuhler (2016) heterogenous income profiles can lead measurement error when child incomes are measured too early to be negatively correlated with parent incomes. However, in the following section we see a simplified model along the lines of equations (8) and (9) describes the evolution of regional estimates of intergenerational mobility in Australia relatively well.

5.2.2. An application – the (ongoing) evolution of regional estimates in Australia

The potential instability of regional conditional expected ranks is reflected in the Australian data. Fig. 7 shows the evolution of regional estimates of the rank-rank slope and expected ranks of a child born to parents at the 25th or 75th percentiles of the national income distribution. We measure child income ranks over a five-year window ending in progressively later years – and hence at later ages and different points in local economic cycles. For clarity we present estimates at the state level, though similar patterns emerge when using SA4 estimates. National estimates are also shown.

Regional estimates of the rank-rank slope rise steadily over time, as apparent in the left panel of Fig. 7. While the dispersion in the estimates falls slightly, the ranking of the states with respect to this measure appears relatively stable. In contrast, regional estimates of the conditional expected ranks move in different directions both between states and within them over time, as apparent in the middle and right panels of Fig. 7. Western Australia has the highest conditional expected ranks by the end of the period, coinciding with the height of the mining boom, but had a much more middling performance at the start of the period. It is also notable that while national conditional expected ranks settle down in the last few years of the period (only increasing slightly in their implied level of persistence over time), this stability still masks significant movements at the regional level.

The visual impression that regional estimates of the rank-rank slope are better approximated as suffering from a common attenuation factor than conditional expected ranks can be confirmed by regression analysis. For this exercise we again produce regional estimates of intergenerational mobility, this time at an SA4 level, measuring child incomes over five-year windows ending in any one of the years from 2002 to 2015 (the latter corresponding to our baseline estimates). Estimates based on earlier windows of observation are expected to be attenuated due to the younger age of the children at measurement, but may also be influenced by cyclical differences across the local labour markets. We run nonlinear least squares regressions of the regional estimates of the rank-rank slope and conditional expected rank on expressions of the form given by equations (8) and (9) – namely a nonlinear combination of year and regional fixed effects – but do not initially model $\bar{v}_{ijt}$. The year fixed effects capture the common attenuation factor, while the regional fixed effects capture lifetime mobility (though these are not uniquely identified from the expressions).

In Table 4 we present model fit statistics for these simple attenuation models. As expected, the evolution of the rank-rank slopes is better explained by these models than the conditional expected ranks — with a correlation between observed and fitted values of 0.915 as opposed to 0.758 and 0.743 (column (1)). Interestingly, in both cases the fit is only marginally worse if we instead measure child incomes over a one-year window (column (4)). Simply measuring over modestly larger windows does not appreciably change the differing evolution of these regional measures of intergenerational mobility.

As speculated above, one potential source for the unexplained variation in regional conditional expected ranks is the influence of local economic shocks, which will be reflected in $\bar{v}_{ijt}$ and feed directly into these metrics. A natural test of this is to control for the local labour market cycle. To do this we include the deviation over the five- or one-year

Table 4
Explaining the evolution of regional estimates of intergenerational mobility.

	Window over which child income is measured					
	Five years			One year		
	(1)	(2)	(3)	(4)	(5)	(6)
Panel A: Rank-rank slope						
$[\mathrm{Corr}(y,\hat{y})]^2$	0.915	0.915	0.955	0.896	0.896	0.920
Coefficient (or average) on local		-0.16	0.44		0.03	0.19
labour market cycle proxy		(0.06)	(0.08)		(0.03)	(0.04)
Panel B: Conditional expected rank (P25)						
$[\mathrm{Corr}(y,\hat{y})]^2$	0.758	0.758	0.865	0.738	0.742	0.808
Coefficient (or average) on local		-4.2	-26.0		-13.8	-26.6
labour market cycle proxy		(5.9)	(8.5)		(3.2)	(4.2)
Panel C: Conditional expected rank (P75)						
$[\mathrm{Corr}(y,\hat{y})]^2$	0.743	0.744	0.858	0.713	0.716	0.785
Coefficient (or average) on local		-14.1	-10.6		-12.9	-19.6
labour market cycle proxy		(6.3)	(9.1)		(3.5)	(4.6)
N	1218	1218	1218	1392	1392	1392
No. of regressors	100	101	187	102	103	189

Notes: Presents model fit statistics and the coefficient or average of the coefficients on a proxy for the local labour market cycle for three models for the evolution of regional estimates of intergenerational mobility. Estimates are at an SA4 level (87 Australian regions) and centred around the years from 2000 through to 2013 (or 2015 when using a one year window). The first model (columns (1) and (4)) regresses the regional estimates on a nonlinear combination of regional and time fixed effects given by equations (8) and (9). The subsequent models add an additional explanatory variable — deviation over the five- or one-year period in the regional fraction unemployed from the national fraction unemployed (minus the mean over the full sample) — as a proxy for the local labour cycle. This is either included by itself (columns (2) and (5)) or interacted with regional fixed effects (columns (3) and (6)).

period in the regional fraction unemployed from the national fraction unemployed (minus the mean over the full sample) as an additional explanatory variable.[29] Movements in this variable reflect deviations from national labour conditions that are beyond the average experience of the given region: they may thus give rise to higher or lower national income ranks for workers in these regions, beyond what would be observed with lifetime incomes. In columns (2) and (5) we include this as a single additional covariate with a common slope coefficient. In columns (3) and (6) we allow region-specific slope coefficients. Region-specific slopes allow this proxy to differ in how it translates to outcomes for those raised in different regions, which could arise from regional differences in out-migration rates and labour market structures, for example.

The inclusion of a proxy for local labour market conditions improves our ability to explain the evolution of regional estimates of intergenerational mobility. The improvements are particularly notable for conditional expected ranks (though also present for the rank-rank slope). As expected, the coefficient on local labour market conditions is negative — a weaker local labour market relative to national conditions, beyond what is typical for the region, leads to lower conditional expected ranks for those raised there. The coefficients are also meaningful in size and generally highly statistically significant — an idiosyncratic one percentage point rise in the local unemployment rate would translate to a loss of 0.1-0.3 percentile rank points in the conditional expected ranks based on columns (3) and (6).

Finally, while we have ruled out local price differences as explaining the static pattern of mobility across Australia it is worth considering whether they could nonetheless be part of the explanation for the *evolution* of regional estimates. Certainly, the mining boom was accompanied by local price inflation in official statistics. In Appendix Table B.7 we replicate Table 4 after first adjusting all regional mobility estimates for local price levels as before. Local price changes appear to play only a modest role in the instability of conditional expected ranks, only slightly narrowing the gap between how well the simple attenuation model fits the evolution of rank-rank slopes and conditional expected ranks. This gap once again narrows, and more substantially, with the inclusion of the local labour market proxy.

The potential for local economic shocks to influence regional estimates of intergenerational mobility suggests extra caution in their interpretation and use. In the Australian setting, the high conditional expected ranks in some regional areas of Queensland and Western Australia may not persist – particularly given unemployment has risen sharply in these regions in more recent years as the Australian mining boom has entered a less labour-intensive phase. More generally, regional estimates of intergenerational mobility can be thought of as reflecting both permanent regional characteristics – such as school quality or structural labour market advantages – and transient ones – such as cyclical effects. Even with a panel long enough to render standard lifecycle and attenuation bias relatively modest (as is the case here), regional estimates may still contain a sizeable transient component. In some cases this may be of genuine interest, but in others it may be something researchers wish to abstract from. Explicitly accounting for local economic shocks, by drawing on frameworks such as the one presented here, may provide a way forward, much as error-in-variables models have been used to remove bias from national estimates of mobility elsewhere.[30]

6. Conclusions

Intergenerational mobility is of key interest to policy-makers in Australia and beyond — it motivates many public policies, and is often thrown into focus in the face of concerns around trends in inequality, globalisation and technological change. We present a new Australian evidence base that aids domestic policy-makers but also adds to the international literature by presenting a detailed picture of intergenerational mobility in a country with different demographics, institutions and economic circumstances.

This paper provides the most precise and comprehensive set of estimates of intergenerational income mobility for Australia to date. Australia emerges as one of the more mobile of the advanced economies.

[29] Regional labour force statistics are routinely published by the Australian Bureau of Statistics (Australian Bureau of Statistics, 2017d).

[30] See Nybom and Stuhler (2016) for a recent proposed approach.

Having a parent 10 percentile rank points higher in the income distribution increases a child's expected rank by a little over 2 percentile rank points. This relationship is a little more pronounced at the bottom 15 per cent of the parent income distribution, possibly reflecting different transmission mechanisms behind entrenched disadvantage.

We also examine differences in mobility across Australian regions. While Australia has much less dispersion in mobility relative to the United States, differences still emerge, both within the country and individual cities. Both macroeconomic and more finely-grained factors appear to be in play. For example, the mining boom appears to have lifted the expected ranks in the national income distribution of children in resource-rich states. But even within individual cities, such as Sydney, regions with mobility measures at either end of the Australian experience sit alongside one another.

Finally, we highlight the potential instability of regional estimates of mobility both in theory and in practice – an additional contribution to a rapidly growing literature and associated set of estimates. In line with an extension to a generalised error-in-variables model, regional rank-rank slopes steadily increase over the period we observe, while the expected national income ranks of children fluctuate in ways that partly mirror the changing economic fortunes of Australian regions.

Appendix A. Additional charts

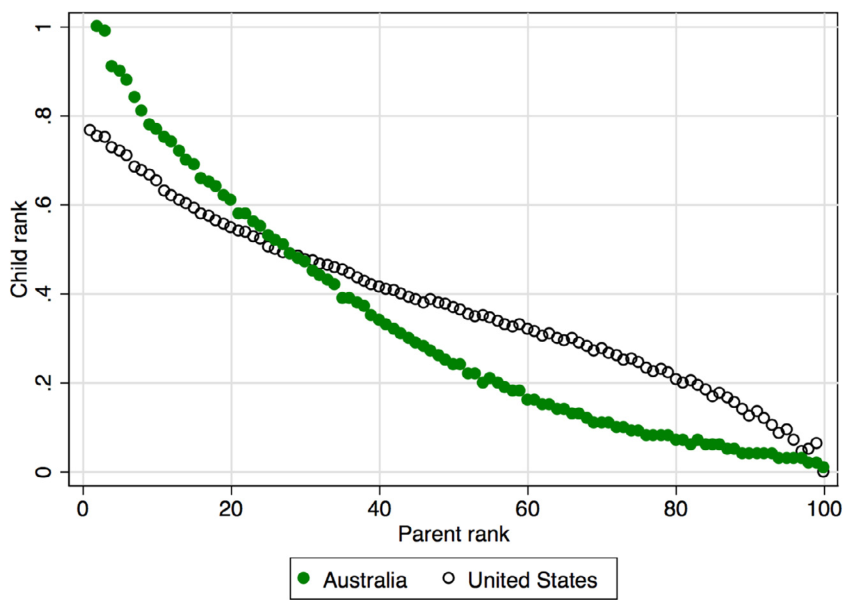

Fig. A.1. Probability income (relative to generational mean) is 50% higher than parents Notes: Chart plots the probability that child total household income is 50% higher than parent total household income, after both have been divided by the respective generational means. Plotted for each percentile bin of the parent income distribution. This is a measure of upward mobility that abstracts from economic growth. United States data is derived from Chetty et al. (2014).

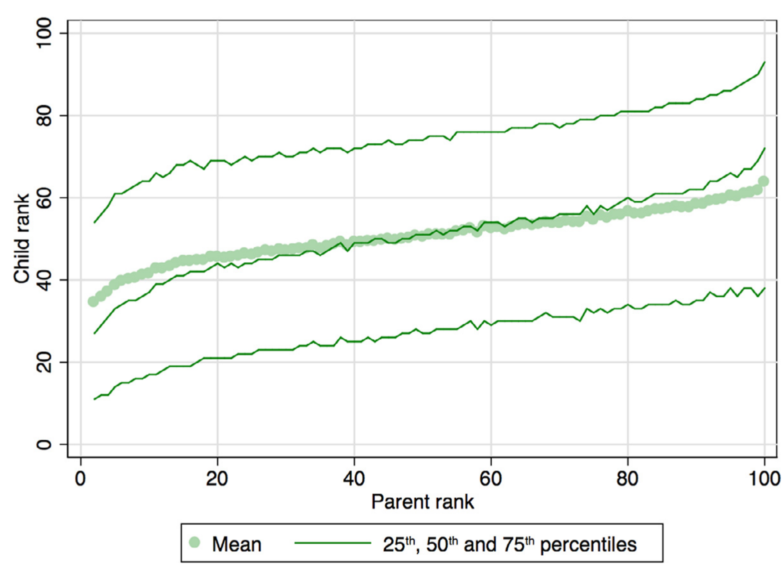

Fig. A.2. Mean and percentiles of child income rank by parent income rank Notes: Chart plots the mean, and 25th, 50th (median), and 75th percentiles of child total household income rank, for each percentile bin of the parent income distribution.

Child ventile \ Parent ventile	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20
20	1.9	2.4	2.9	3.2	3.2	3.4	3.7	3.7	3.9	3.7	4.1	4.5	4.6	5.0	5.4	6.1	6.7	7.7	9.3	14.6
19	2.2	2.8	3.3	3.7	3.6	3.8	4.0	4.2	4.5	4.5	4.7	4.9	5.4	5.3	5.8	6.2	6.8	7.3	8.0	9.0
18	2.5	3.2	3.6	3.8	3.9	4.1	4.1	4.4	4.5	4.9	5.0	5.2	5.4	5.5	5.9	6.3	6.4	6.7	7.4	7.1
17	2.6	3.4	3.8	3.9	4.0	4.2	4.4	4.7	4.7	4.9	5.1	5.4	5.2	5.7	6.0	6.1	6.3	6.6	6.6	6.2
16	2.8	3.4	3.9	4.1	4.3	4.6	4.6	4.6	4.7	5.2	5.3	5.4	5.6	5.7	5.6	5.9	6.3	6.2	6.3	5.5
15	3.1	3.7	4.1	4.3	4.4	4.5	4.8	4.8	5.2	5.3	5.3	5.5	5.5	5.5	5.7	6.0	5.8	5.8	5.8	5.0
14	3.2	3.9	4.4	4.5	4.7	4.8	4.7	4.9	5.1	5.2	5.3	5.5	5.6	5.8	5.7	5.6	5.5	5.4	5.2	4.7
13	3.6	3.9	4.3	4.8	4.8	5.0	4.9	5.5	5.1	5.2	5.4	5.4	5.6	5.7	5.4	5.5	5.5	5.2	4.9	4.5
12	3.6	4.4	4.5	4.8	4.9	5.1	5.3	5.1	5.5	5.5	5.4	5.5	5.5	5.6	5.4	5.2	5.1	4.9	4.8	4.0
11	4.0	4.4	4.8	5.0	5.0	5.3	5.3	5.4	5.4	5.3	5.4	5.4	5.5	5.3	5.2	5.1	4.9	4.9	4.6	4.0
10	4.2	4.6	4.9	5.1	5.2	5.2	5.3	5.5	5.3	5.4	5.4	5.4	5.2	5.3	5.1	5.1	4.7	4.6	4.4	3.9
9	4.4	4.8	5.0	5.3	5.4	5.3	5.5	5.5	5.3	5.4	5.2	5.2	5.3	5.3	5.0	4.8	4.7	4.4	4.4	3.9
8	4.8	5.3	5.2	5.2	5.4	5.5	5.5	5.4	5.5	5.3	5.5	5.2	5.0	5.0	5.1	4.7	4.4	4.4	4.2	3.5
7	5.2	5.6	5.6	5.5	5.7	5.7	5.5	5.4	5.2	5.4	5.3	5.1	4.9	4.9	4.5	4.6	4.5	4.2	3.8	3.5
6	5.8	5.9	5.8	5.9	5.9	5.6	5.6	5.4	5.2	5.3	5.1	5.1	4.8	4.7	4.7	4.2	4.2	4.0	3.6	3.2
5	6.6	6.5	6.2	6.0	5.9	5.8	5.6	5.5	5.4	5.3	4.9	4.7	4.5	4.4	4.4	4.1	3.8	3.7	3.4	3.2
4	7.8	7.1	6.6	6.2	6.0	5.7	5.5	5.3	5.2	5.0	4.7	4.5	4.4	4.1	4.0	3.9	3.9	3.5	3.4	3.3
3	9.5	8.2	7.0	6.5	6.1	5.7	5.6	5.2	5.0	4.6	4.5	4.2	4.1	3.8	3.7	3.4	3.4	3.3	3.1	3.1
2	11.8	8.8	7.7	6.4	5.9	5.5	5.3	4.9	4.8	4.5	4.0	3.8	3.8	3.8	3.5	3.3	3.3	3.1	2.9	3.1
1	10.4	7.5	6.4	6.0	5.6	5.1	5.0	4.7	4.5	4.3	4.3	4.0	4.0	3.7	3.9	3.7	3.9	4.0	4.1	4.8

Fig. A.3. Intergenerational transition matrix Notes: Shows the per cent frequency with which a child with parents in a given income ventile (column) ends up in given income ventile (row) themselves.

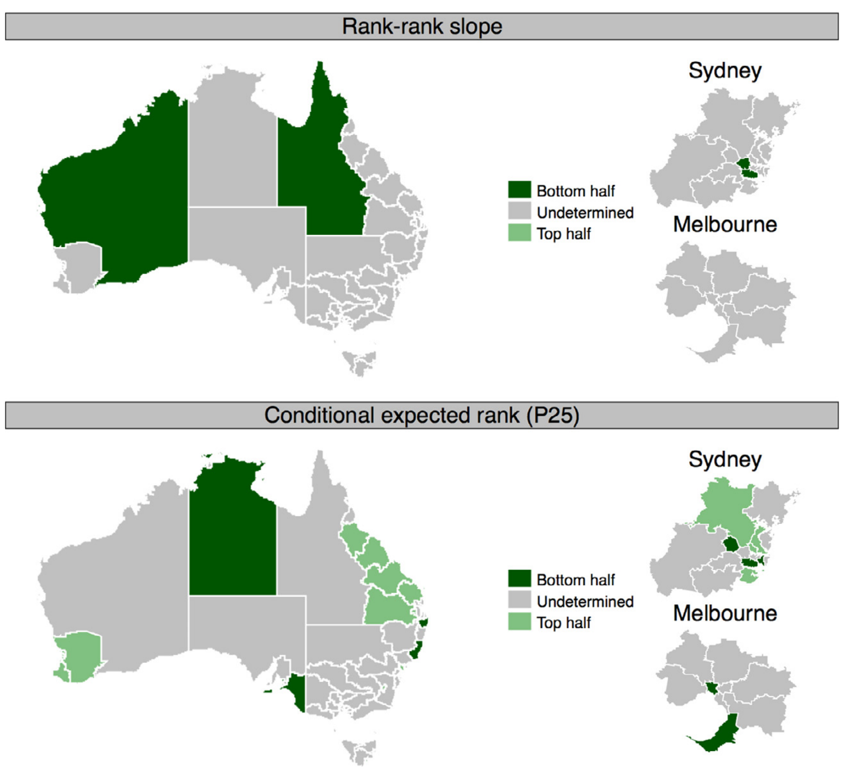

Fig. A.4. Intergenerational mobility within Australia and its two largest cities – regions in the bottom and top half when ranked by mobility (with 95% confidence) Notes: Shows whether regional (SA4) estimates of mobility fall, with 95% confidence, into the bottom or top half of all regions when ranked on that metric. The metrics are the expected rank, conditional on being born into the 25th percentile of the national parent income distribution, and the rank-rank slope. The confidence sets are calculated using the algorithm introduced by Mogstad et al. (2020). We thank the authors for providing their code (available on Github at: https://github.com/danielwilhelm/R-CS-ranks). Maps for Australia and its two largest cities, Sydney and Melbourne, are shown.

Appendix B. Additional tables

Table B.1
Distribution of parental ages at birth and during income observation.

	Full sample		Population	
	At birth	At observation	At birth	At observation
Panel A: Distribution of maternal ages				
p10	20	29–43	20	29–43
p25	23	32–46	23	32–46
p50	27	36–50	26	35–49
p75	30	39–53	29	38–52
p90	33	42–56	33	42–56
Panel B: Distribution of paternal ages				
p10	23	32–46	23	32–46
p25	26	35–49	26	35–49
p50	29	38–52	29	38–52
p75	33	42–56	33	42–56
p90	37	46–60	37	46–60

Notes: Distribution of parental age at birth and implied age range over the period when parent incomes are observed (1991–2001). Population estimates are based on Australian Bureau of Statistics (2017b).

Table B.2
Distribution of parent and child incomes.

	Full sample		Population	
	Mean	SD	Mean	SD
Panel A: Child generation (2014-15)				
$ income	94,600	81,500	112,500	75,800
log income	11.24	0.95	11.36	0.91
% with zero income	8.2		0.0	
% with spouse	69.5		73.5	
% with spouse with income	61.2		71.2	
Panel B: Child generation (2012-13)				
$ income	99,400	103,800	100,900	69,800
log income	11.28	0.91	11.25	0.92
% with zero income	5.8		0.0	
% with spouse	69.5		66.7	
% with spouse with income	63.9		65.5	
Panel C: Child generation (2010-11)				
$ income	93,500	111,700	96,600	87,400
log income	11.22	0.90	11.21	0.88
% with zero income	5.6		0.1	
% with spouse	69.5		66.7	
% with spouse with income	63.8		65.1	
Panel D: Parent generation (1993-94)				
$ income	77,600	80,100	82,400	56,600
log income	11.14	0.73	11.13	0.68
% with zero income	9.6		0.0	

Notes: Population estimates are based on analysis of confidentialised unit record files of the Survey of Income and Housing run by the Australian Bureau of Statistics. For the child generation, we display summary statistics for all three survey years in the window of observation (2011–2015); for parents we chose the earliest survey year in the window of observation (1991–2001), as in later years children are less likely to be observed in the household. More information on unit record data is available in Australian Bureau of Statistics, 2005, 2013, 2015, 2017.

Table B.3
National measures of intergenerational income mobility — father/mother and son/daugther combinations.

	IGE	Pearson correlation	Rank-based		
			Rank-rank slope	$E[r_{1i} \mid r_{0i} = 25]$	$E[r_{1i} \mid r_{0i} = 75]$
Father-son	0.157	0.135	0.171	55.3	63.8
	(0.002)	(0.002)	(0.001)	(0.1)	(0.1)
	414,100	414,100	440,500	440,500	440,500
Father-daughter	0.133	0.119	0.131	39	45.6
	(0.002)	(0.002)	(0.001)	(0.1)	(0.1)
	384,800	384,800	413,500	413,500	413,500
Mother-son	0.065	0.039	0.122	56.1	62.1
	(0.001)	(0.001)	(0.001)	(0.1)	(0.1)
	408,800	408,800	462,700	462,700	462,700
Mother-daughter	0.105	0.076	0.15	38.3	45.8
	(0.002)	(0.002)	(0.001)	(0.1)	(0.1)
	385,800	385,800	440,400	440,400	440,400

Notes: Presents estimates of five different measures of intergenerational persistence varying the sample only: having first calculated income ranks, log income and normalized log income, we: restrict estimation to children from the middle 80% of the parent income distribution; weight children by the inverse of the probability that a child at the same percentile rank in the child income distribution is linked to parents; and restrict estimation to children with the highest quality links, where the primary parent's predicted probability of being a parent is at least 0.95, and weight children by the inverse of the probability that a child at the same percentile rank in the child income distribution is linked to such parents.

Table B.4
National measures of intergenerational income mobility — robustness.

	IGE	Pearson correlation	Rank-based		
			Rank-rank slope	$E[r_{1i} \mid r_{0i} = 25]$	$E[r_{1i} \mid r_{0i} = 75]$
Middle 80%	0.241	0.195	0.181	46.2	55.2
	(0.002)	(0.002)	(0.001)	(0.0)	(0.0)
	747,000	747,000	772,600	772,600	772,600
Weighted	0.191	0.165	0.217	43.8	54.6
	(0.001)	(0.001)	(0.001)	(0.0)	(0.0)
	900,700	900,700	965,700	965,700	965,700
Highest quality links	0.195	0.161	0.208	44.0	54.4
	(0.002)	(0.001)	(0.001)	(0.0)	(0.0)
	807,700	807,700	855,100	855,100	855,100

Notes: Presents estimates of five different measures of intergenerational persistence varying the sample only: having first calculated income ranks, log income and normalized log income, we: restrict estimation to children from the middle 80% of the parent income distribution; weight children by the inverse of the probability that a child at the same percentile rank in the child income distribution is linked to parents; and restrict estimation to children with the highest quality links, where the primary parent's predicted probability of being a parent is at least 0.95, and weight children by the inverse of the probability that a child at the same percentile rank in the child income distribution is linked to such parents.

Table B.5
National measures of intergenerational income mobility — conservative.

	IGE	Pearson correlation	Rank-based		
			Rank-rank slope	$E[r_{1i}\|r_{0i}=25]$	$E[r_{1i}\|r_{0i}=75]$
Panel A: Conservative rank-rank windows					
Weighted	0.188	0.173	0.232	43.4	55.0
	(0.001)	(0.001)	(0.001)	(0.0)	(0.0)
	924,000	924,000	966,500	966,500	966,500
Highest quality links	0.205	0.180	0.225	42.5	53.7
	(0.002)	(0.001)	(0.001)	(0.0)	(0.0)
	824,800	824,800	855,700	855,700	855,700
Panel B: Conservative IGE windows					
Weighted	0.198	0.169	0.227	43.5	54.9
	(0.001)	(0.001)	(0.001)	(0.0)	(0.0)
	912,600	912,600	966,200	966,200	966,200
Highest quality links	0.210	0.172	0.220	42.5	53.5
	(0.002)	(0.001)	(0.001)	(0.0)	(0.0)
	817,100	817,100	855,500	855,500	855,500
Panel C: 1978 birth cohort, parents 40–55 years at observation					
Weighted	0.225	0.167	0.191	44.5	54.1
	(0.005)	(0.004)	(0.003)	(0.1)	(0.1)
	97,800	97,800	102,200	102,200	102,200
Highest quality links	0.238	0.176	0.193	42.9	52.6
	(0.006)	(0.004)	(0.003)	(0.1)	(0.1)
	88,900	88,900	92,700	92,700	92,700

Notes: Presents estimates of five different measures of intergenerational persistence varying both the sample and the window over which income is observed. Panel A presents estimates based on windows of observation of: 25 years for parents, centred in 2003 and implying an average age of 51 years; and 9 years for children, centred at in 2011 and implying an average age of 31 years. Panel B presents estimates based on windows of observation of: 15 years for parents, centred in 1998 and implying an average age of 46 years; and 5 years for children, centred at in 2013 and implying an average age of 33 years. Panel C presents estimates based on the 1978 birth cohort only, with their income measured in the last three years of the panel, and restricting to parents aged 40–55 years during the window of observation of parental income and with at least five non-missing income observations.

Table B.6
Geographic mobility – percentage of children living in a different region or state to where they grew up.

	National	Across SA4		
		Least mobile	Median	Most mobile
% in a different region (SA4)	29.4	19.4	30.0	39.0
% in a different state	9.7	4.2	7.3	31.1

Notes: Shows the percentage of children who filed a tax return in 2015 from a different region (SA4) or state to that first recorded for their primary parent. Both national estimates and those for the typical (median) and least and most mobile SA4 on the given metric are shown. These measures of mobility will not capture those who have moved but did not file a tax return in 2015.

Table B.7
Explaining the evolution of regional estimates of intergenerational mobility – local-price-adjusted estimates of mobility.

	Window over which child income is measured					
	Five years			One year		
	(1)	(2)	(3)	(4)	(5)	(6)
Panel A: Rank-rank slope						
$[\mathrm{Corr}(y, \hat{y})]^2$	0.911	0.912	0.953	0.892	0.892	0.917
Coefficient (or average) on local		-0.15	0.46		0.04	0.21
labour market cycle proxy		(0.06)	(0.08)		(0.03)	(0.04)
Panel B: Conditional expected rank (P25)						
$[\mathrm{Corr}(y, \hat{y})]^2$	0.778	0.778	0.878	0.760	0.762	0.822
Coefficient (or average) on local		-1.9	-18.6		-9.5	-18.5
labour market cycle proxy		(5.7)	(8.2)		(3.0)	(4.0)
Panel C: Conditional expected rank (P75)						
$[\mathrm{Corr}(y, \hat{y})]^2$	0.757	0.758	0.870	0.724	0.724	0.795
Coefficient (or average) on local		-12.2	-3.6		-8.7	-11.3
labour market cycle proxy		(6.2)	(8.8)		(3.4)	(4.5)
N	1218	1218	1218	1392	1392	1392
No. of regressors	100	101	187	102	103	189

Notes: Replicates Table 4 from the body of the paper, but based on local-price-adjusted measures of intergenerational mobility. Presents model fit statistics and the coefficient or average of the coefficients on a proxy for the local labour market cycle for three models for the evolution of regional estimates of intergenerational mobility. Estimates are at an SA4 level (87 Australian regions) and centred around the years from 2000 through to 2013 (or 2015 when using a one year window). The first model (columns (1) and (4)) regresses the regional estimates on a nonlinear combination of regional and time fixed effects given by equations (8) and (9). The subsequent models add an additional explanatory variable — deviation over the five- or one-year period in the regional fraction unemployed from the national fraction unemployed (minus the mean over the full sample) — as a proxy for the local labour cycle. This is either included by itself (columns (2) and (5)) or interacted with regional fixed effects (columns (3) and (6)).

Table B.8
Regional estimates of intergenerational income mobility.

SA4 code	SA4 name	Number of children	Mobility metric			
			IGE	Rank-rank slope	$E[r_{1i}\vert r_{0i}=25]$	$P[r_{1i}>80\vert r_{0i}\le 20]$
101	Capital Region	8800	0.173	0.198	45.3	10.9
			(0.015)	(0.011)	(0.4)	
102	Central Coast	13,800	0.164	0.182	43.8	11.0
			(0.011)	(0.009)	(0.3)	
103	Central West	10,500	0.196	0.208	46.4	12.3
			(0.013)	(0.010)	(0.4)	
104	Coffs Harbour - Grafton	6800	0.160	0.190	43.6	10.4
			(0.016)	(0.013)	(0.4)	
105	Far West and Orana	6600	0.210	0.260	42.7	10.7
			(0.017)	(0.012)	(0.4)	
106	Hunter Valley exc Newcastle	12,200	0.205	0.242	46.3	13.2
			(0.014)	(0.009)	(0.4)	
107	Illawarra	14,600	0.182	0.219	45.1	13.1
			(0.012)	(0.008)	(0.3)	
108	Mid North Coast	9600	0.178	0.234	42.6	9.2
			(0.013)	(0.011)	(0.3)	
109	Murray	6000	0.176	0.194	45.2	10.3
			(0.017)	(0.013)	(0.4)	
110	New England and North West	9800	0.182	0.219	43.9	9.2
			(0.013)	(0.010)	(0.3)	
111	Newcastle and Lake Macquarie	17,400	0.174	0.210	45.7	11.6
			(0.011)	(0.007)	(0.3)	
112	Richmond - Tweed	10,700	0.167	0.206	42.4	9.7
			(0.012)	(0.010)	(0.3)	
113	Riverina	8700	0.208	0.220	44.3	8.6
			(0.015)	(0.011)	(0.4)	
114	Southern Highlands and Shoalhaven	6200	0.150	0.182	43.8	12.6
			(0.016)	(0.013)	(0.4)	
115	Sydney - Baulkham Hills and Hawkesbury	10,200	0.173	0.177	49.2	18.3
			(0.013)	(0.010)	(0.5)	

(continued on next page)

Table B.8 (*continued*)

SA4 code	SA4 name	Number of children	Mobility metric			
			IGE	Rank-rank slope	$E[r_{1i}\|r_{0i}=25]$	$P[r_{1i}>80\|r_{0i}\leq 20]$
116	Sydney - Blacktown	14,000	0.191 (0.012)	0.240 (0.008)	42.3 (0.3)	9.7
117	Sydney - City and Inner South	5500	0.173 (0.017)	0.199 (0.013)	41.4 (0.4)	11.7
118	Sydney - Eastern Suburbs	7300	0.189 (0.014)	0.216 (0.012)	44.2 (0.5)	17.3
119	Sydney - Inner South West	21,700	0.206 (0.009)	0.257 (0.006)	42.7 (0.2)	12.9
120	Sydney - Inner West	8200	0.178 (0.014)	0.178 (0.011)	45.1 (0.4)	14.4
121	Sydney - North Sydney and Hornsby	13,800	0.175 (0.011)	0.193 (0.009)	47.3 (0.5)	19.7
122	Sydney - Northern Beaches	9900	0.151 (0.014)	0.168 (0.010)	46.1 (0.5)	16.1
123	Sydney - Outer South West	13,900	0.211 (0.013)	0.235 (0.008)	43.6 (0.3)	11.5
124	Sydney - Outer West and Blue Mountains	17,500	0.148 (0.011)	0.169 (0.008)	45.8 (0.3)	13.9
125	Sydney - Parramatta	15,900	0.229 (0.011)	0.270 (0.007)	43.3 (0.3)	11.2
126	Sydney - Ryde	6300	0.162 (0.017)	0.172 (0.013)	48.8 (0.6)	19.5
127	Sydney - South West	14,300	0.143 (0.010)	0.211 (0.008)	43.8 (0.3)	12.0
128	Sydney - Sutherland	11,300	0.196 (0.014)	0.188 (0.010)	48.5 (0.5)	16.7
201	Ballarat	7500	0.175 (0.016)	0.187 (0.012)	43.4 (0.4)	9.2
202	Bendigo	7200	0.154 (0.015)	0.201 (0.012)	43.4 (0.4)	9.2
203	Geelong	12,600	0.147 (0.012)	0.171 (0.009)	44.1 (0.3)	9.9
204	Hume	8600	0.168 (0.015)	0.188 (0.011)	43.8 (0.4)	9.4
205	Latrobe - Gippsland	14,100	0.147 (0.011)	0.184 (0.009)	45.5 (0.3)	12.6
206	Melbourne - Inner	11,600	0.176 (0.012)	0.213 (0.009)	40.4 (0.3)	9.9
207	Melbourne - Inner East	14,500	0.181 (0.011)	0.194 (0.008)	45.9 (0.4)	16.9
208	Melbourne - Inner South	13,700	0.177 (0.011)	0.197 (0.008)	44.1 (0.4)	13.0
209	Melbourne - North East	21,100	0.178 (0.010)	0.196 (0.007)	43.9 (0.3)	11.2
210	Melbourne - North West	14,700	0.188 (0.011)	0.205 (0.008)	43.1 (0.3)	10.8
211	Melbourne - Outer East	28,100	0.163 (0.009)	0.175 (0.006)	45.1 (0.2)	12.0
212	Melbourne - South East	26,200	0.155 (0.008)	0.195 (0.006)	43.6 (0.2)	11.6
213	Melbourne - West	22,500	0.141 (0.009)	0.189 (0.007)	43.0 (0.2)	9.9
214	Mornington Peninsula	12,200	0.186 (0.013)	0.192 (0.009)	42.4 (0.3)	9.9
215	North West	8000	0.195 (0.017)	0.229 (0.011)	43.5 (0.4)	8.6
216	Shepparton	7100	0.170 (0.016)	0.197 (0.012)	43.4 (0.4)	8.6
217	Warrnambool and South West	7100	0.164 (0.017)	0.174 (0.012)	45.3 (0.4)	9.8
301	Brisbane - East	9000	0.143 (0.015)	0.179 (0.011)	47.1 (0.4)	11.8
302	Brisbane - North	8000	0.203 (0.017)	0.220 (0.012)	47.3 (0.4)	15.0
303	Brisbane - South	11,800	0.179 (0.012)	0.222 (0.009)	46.6 (0.4)	14.0
304	Brisbane - West	7700	0.166 (0.016)	0.185 (0.012)	49.2 (0.6)	18.7
305	Brisbane Inner City	5300	0.183 (0.018)	0.201 (0.014)	45.8 (0.6)	15.6
306	Cairns	9500	0.213 (0.012)	0.254 (0.010)	42.5 (0.4)	11.4

(*continued on next page*)

Table B.8 (*continued*)

SA4 code	SA4 name	Number of children	Mobility metric			
			IGE	Rank-rank slope	$E[r_{1i}\|r_{0i}=25]$	$P[r_{1i}>80\|r_{0i}\leq 20]$
307	Darling Downs - Maranoa	5700	0.165	0.229	47.1	12.4
			(0.016)	(0.014)	(0.4)	
308	Fitzroy	10,800	0.200	0.220	50.9	18.8
			(0.015)	(0.010)	(0.4)	
309	Gold Coast	13,800	0.163	0.182	43.4	11.2
			(0.011)	(0.009)	(0.3)	
310	Ipswich	11,300	0.162	0.216	44.6	10.8
			(0.013)	(0.010)	(0.3)	
311	Logan - Beaudesert	15,000	0.174	0.218	44.0	9.9
			(0.011)	(0.008)	(0.3)	
312	Mackay	8200	0.172	0.184	53.4	20.3
			(0.015)	(0.011)	(0.5)	
313	Moreton Bay - North	8500	0.157	0.214	44.4	9.9
			(0.014)	(0.011)	(0.3)	
314	Moreton Bay - South	7100	0.177	0.182	48.1	13.6
			(0.020)	(0.013)	(0.5)	
315	Queensland - Outback	3800	0.270	0.319	42.7	11.3
			(0.021)	(0.016)	(0.6)	
316	Sunshine Coast	10,600	0.143	0.175	44.6	12.3
			(0.014)	(0.011)	(0.3)	
317	Toowoomba	7300	0.161	0.191	47.7	13.4
			(0.016)	(0.012)	(0.4)	
318	Townsville	10,300	0.210	0.225	46.9	14.1
			(0.015)	(0.010)	(0.4)	
319	Wide Bay	12,300	0.176	0.226	47.0	12.4
			(0.011)	(0.010)	(0.3)	
401	Adelaide - Central and Hills	11,400	0.180	0.187	45.3	12.9
			(0.013)	(0.009)	(0.4)	
402	Adelaide - North	19,700	0.184	0.233	42.0	8.2
			(0.009)	(0.007)	(0.2)	
403	Adelaide - South	17,700	0.177	0.204	43.9	10.7
			(0.010)	(0.008)	(0.3)	
404	Adelaide - West	9000	0.174	0.240	41.8	9.5
			(0.013)	(0.010)	(0.3)	
405	Barossa - Yorke - Mid North	4800	0.145	0.213	43.7	8.0
			(0.018)	(0.015)	(0.5)	
406	South Australia - Outback	4800	0.224	0.227	45.1	11.4
			(0.026)	(0.014)	(0.5)	
407	South Australia - South East	7600	0.191	0.215	42.0	7.6
			(0.017)	(0.012)	(0.4)	
501	Bunbury	6800	0.190	0.230	51.0	17.2
			(0.016)	(0.013)	(0.5)	
502	Mandurah	2500	0.154	0.227	52.5	21.8
			(0.023)	(0.022)	(0.8)	
503	Perth - Inner	5000	0.159	0.183	51.1	22.5
			(0.016)	(0.014)	(0.7)	
504	Perth - North East	8700	0.158	0.188	50.6	20.2
			(0.015)	(0.012)	(0.4)	
505	Perth - North West	20,000	0.182	0.202	50.6	19.3
			(0.009)	(0.007)	(0.3)	
506	Perth - South East	19,000	0.175	0.192	50.4	19.0
			(0.010)	(0.008)	(0.3)	
507	Perth - South West	13,000	0.181	0.204	50.9	18.5
			(0.011)	(0.009)	(0.4)	
508	Western Australia - Outback	9700	0.258	0.273	46.9	15.5
			(0.015)	(0.011)	(0.5)	
509	Western Australia - Wheat Belt	6200	0.204	0.241	48.7	16.9
			(0.018)	(0.014)	(0.5)	
601	Hobart	11,800	0.229	0.259	40.5	8.1
			(0.015)	(0.009)	(0.3)	
602	Launceston and North East	7700	0.161	0.222	42.6	8.2
			(0.014)	(0.011)	(0.4)	
603	South East	1400	0.192	0.184	43.0	6.8
			(0.036)	(0.029)	(0.8)	
604	West and North West	7300	0.158	0.181	44.6	9.6
			(0.016)	(0.012)	(0.4)	
701	Darwin	5300	0.181	0.199	47.0	16.4
			(0.019)	(0.013)	(0.6)	
702	Northern Territory - Outback	2500	0.267	0.283	40.5	7.2
			(0.027)	(0.019)	(0.8)	
801	Australian Capital Territory	18,000	0.184	0.191	48.5	16.7
			(0.011)	(0.008)	(0.4)	

Notes: Presents estimates of intergenerational mobility for those born in Australia in the 1978-82 financial years. Parent household total pretax incomes are measured from 1991 to 2001, while the total household incomes of the adult children are measured from 2011 to 2015. Sample sizes rounded to the nearest 100 and standard errors in parentheses.

Appendix C. Robustness of key measures to treatment of missing values

In this Appendix, we examine the sensitivity of our national measures of intergenerational mobility to the treatment of missing values — years in which child or parent incomes are not observed. Recent work has noted the sensitivity of the intergenerational elasticity to such assumptions, citing it as partial justification for adopting new measures of intergenerational mobility. For example, (Chetty et al., 2014) promotes the rank-rank slope while Mitnik et al. (2015) proposes a new elasticity measure, which we also present here. There are many potential ways to treat missing values in income data. We consider the following:

- Imputing $1 to all missing values
- Imputing $1,000 to all missing values
- Imputing $10,000 to all missing values
- Dropping annual missing values
- Dropping lifetime missing values

Recent concerns have been in the context of missing values for child income, so we begin by applying these transformations to missing values in the child and child's spouse income histories (Panel A). However, we go on to apply the same transformations to missing values in the parents' income histories (Panel B) and all income histories (Panel C). In all cases we apply the same treatments to negative income, though negative incomes are sufficiently rare that this does not influence the conclusions drawn.

Table C.1 presents the results from this exercise. As expected, the intergenerational elasticity and correlation are much more sensitive to the treatment of zeroes than the rank-based mobility measures, across all panels. However, when only concerned with child household missing values, this sensitivity is greatly reduced when imputations are restricted to more plausible values. For example, it is difficult to imagine a situation in which it is appropriate to impute an income of $1 — reasonable imputed values for earnings capacity, subsistence income or similar would likely be much higher. Strikingly, Panels B and C show that the elasticity estimates are even more sensitive when missing values in parent income are also treated. This is true even for the IGE measure proposed by Mitnik et al. (2015), which is robust to the treatment of missing values in the child generation.

Table C.1 provides an important caveat on the IGE and Pearson correlation measures. That said, the range of values apparent still mark Australia out as a particularly mobile advanced economy. The exercises also highlight the potential importance of missing values in parent income histories, as well as those in child income histories. An assessment of the most appropriate treatment of these missing values is well beyond the scope of this paper. Such an exercise would need to include an assessment of the underlying processes generating missing values. For example, missing values

Table C.1
National measures of intergenerational income mobility.

	IGE	IGE-Mitnik	Pearson correlation	Rank-based		
				Rank-rank slope	$E[r_{1i} \mid r_{0i} = 25]$	$E[r_{1i} \mid r_{0i} = 75]$
Impute $1	0.26	0.17	0.12	0.22	45.0	55.7
	(0.00)	(0.00)	(0.00)	(0.00)	(0.0)	(0.0)
Impute $1k	0.20	0.17	0.17	0.22	45.0	55.7
	(0.00)	(0.00)	(0.00)	(0.00)	(0.0)	(0.0)
Impute $10k	0.17	0.16	0.18	0.22	45.0	55.7
	(0.00)	(0.00)	(0.00)	(0.00)	(0.0)	(0.0)
Drop annual	0.17	0.16	0.17	0.22	44.9	55.8
	(0.00)	(0.00)	(0.00)	(0.00)	(0.0)	(0.0)
Drop lifetime	0.19	0.16	0.16	0.22	44.9	55.8
	(0.00)	(0.00)	(0.00)	(0.00)	(0.0)	(0.0)
Panel B: Parent missing values						
Impute $1	0.07	0.07	0.13	0.22	45.0	55.8
	(0.00)	(0.00)	(0.00)	(0.00)	(0.0)	(0.0)
Impute $1k	0.17	0.15	0.15	0.22	45.0	55.8
	(0.00)	(0.00)	(0.00)	(0.00)	(0.0)	(0.0)
Impute $10k	0.26	0.22	0.16	0.21	45.1	55.7
	(0.00)	(0.00)	(0.00)	(0.00)	(0.0)	(0.0)
Drop annual	0.22	0.20	0.14	0.18	46.4	55.5
	(0.00)	(0.00)	(0.00)	(0.00)	(0.0)	(0.0)
Drop lifetime	0.19	0.17	0.16	0.20	46.0	55.8
	(0.00)	(0.00)	(0.00)	(0.00)	(0.0)	(0.0)
Panel C: Child and parent missing values						
Impute $1	0.12	0.07	0.16	0.22	45.0	55.7
	(0.00)	(0.00)	(0.00)	(0.00)	(0.0)	(0.0)
Impute $1k	0.19	0.15	0.23	0.22	45.0	55.7
	(0.00)	(0.00)	(0.00)	(0.00)	(0.0)	(0.0)
Impute $10k	0.24	0.21	0.22	0.21	45.0	55.7
	(0.00)	(0.00)	(0.00)	(0.00)	(0.0)	(0.0)
Drop annual	0.21	0.19	0.15	0.19	46.1	55.7
	(0.00)	(0.00)	(0.00)	(0.00)	(0.0)	(0.0)
Drop lifetime	0.19	0.16	0.16	0.20	45.8	55.9
	(0.00)	(0.00)	(0.00)	(0.00)	(0.0)	(0.0)

Notes: Presents estimates of intergenerational mobility for those born in Australia in the 1978-82 financial years. Parent household total pretax incomes are measured from 1991 to 2001, while the total household incomes of the adult children are measured from 2011 to 2015. Each row applies a different treatment to missing values, either imputing a value to them, dropping them or treating them as zeroes and dropping only lifetime zeroes. Standard errors in parentheses.

arising from unemployment, caring responsibilities or emigration may differ substantially in the information they carry about expected lifetime incomes.

Appendix D. Construction of potential correlates of mobility

This Appendix describes the source and construction of the potential correlates of mobility explored in Section VI. The choice of correlates is inspired by those in the study by Chetty et al. (2014) but is not intended to be exhaustive. Further, some correlates are either less readily available or less applicable in the Australian setting. For example, Australia is a much more centralized federation – states and territories do not levy income taxes and the federal government distributes revenue from the major consumption tax according to a formula designed to equalize fiscal capacity. Thus while the quality of local public services likely does influence mobility across Australian regions, the link to state and territory fiscal settings may be more tenuous.

Table D.1 lists the correlates, their definitions and sources. The two most complicated correlates to construct are the ethnic and income segregation indices. We follow the approaches in Chetty et al. (2014). We measure ethnic segregation using a Theil Index H using data on the ancestry reported within the Statistical Area 2 making up the Statistical Area 4 from the 2011 Census of Population and Housing. There are an average a little over 20 of these areas within each of the SA4. Ancestry is considered at the highest level of classification provided, which results in nine ethnic groups. Ethnic diversity within a given SA2 j is measured using an entropy index as follows, where ϕ_{ej} is the fraction of individuals in SA2 j in a given ethnic group e:

$$E_j = \Sigma_e \phi_{ej} \log_2 \frac{1}{\phi_{ej}}$$

Dropping the j subscript we can also define entropy E at the SA4 level – this we use as a measure of ethnic diversity. Finally, we can measure the degree of ethnic segregation in the SA4 as:

$$H = \sum_j \left[\frac{\text{pop}_j}{\text{pop}_{total}} \frac{E - E_j}{E} \right]$$

where pop_j and pop_{total} are the populations of SA2 j and SA4 respectively. This index is maximized at 1 when there is no ethnic diversity in the individual SA2, but some in the larger SA4 that is therefore completely segregated. It is minimized at 0 when the ethnic diversity in the SA2 equals that in the SA4.

For income segregation we use data on income reported (in one of seventeen categories) within the Statistical Area 2 making up the Statistical Area 4 from the 2011 Census of Population and Housing. A slightly different approach is taken given the ordered nature of the income categories. For each of the income categories c we can calculate a two-group Theil index $H(c)$ using the formulae above for the segregation of those at or below the midpoint of the income category. It then becomes simply a matter of aggregating these indices into a single metric. Following Reardon (2011) we take a weighted sum that weights each index according to the fraction of the local population in the given income category and the entropy observed E(c) at the SA4-level at that point of the income distribution:

$$\text{income segregation} = 2\log(2) \sum_c \frac{\text{pop}_c}{\text{pop}_{total}} E(c) H(c)$$

Despite capturing segregation along different lines and having different definitions, the two measures of segregation are quite highly correlated, with a correlation of over 0.7 across the SA4. In the text we examine only bivariate correlations. In Table D.2 we examine the multivariate correlations between regional measures of intergenerational mobility and explanatory factors. The associations between the fraction Indigenous and intergenerational persistence, and higher incomes and higher conditional expected ranks, remain. However, due to the small number of regions and large number of correlates we do not put much weight on this analysis.

Table D.1
Definition and source of correlates.

Category	Correlate	Definition	Source
Diversity and distance	Fraction Indigenous	Fraction of people identifying as Aboriginal and/or Torres Strait Islander	2011 Census
	Ethnic diversity	See text	2011 Census
	Ethnic segregation	See text	2011 Census
	Income segregation	See text	2011 Census
	Mean commute	Mean commuting distance (km)	2016 Census
Labour market	Mean income	Mean total income (2012-13 taxpayers)	ABS small area income statistics
	Unemployment rate	Fraction of 29–33 year olds unemployed	2011 Census
	NILF rate	Fraction of 29–33 year olds not in the labour force (NILF)	2011 Census
Education	School attendance	Fraction of 15 year olds in full-time study	2011 Census
	Tertiary attendance	Fraction of 21 year olds in full-time study	2011 Census
Inequality	Gini	Gini coefficient of total income (2012-13 taxpayers)	ABS small area income statistics
	Top 1% share	Top 1% share of total income (2012-13 taxpayers)	ABS small area income statistics
Social	Fraction volunteers	Fraction of people volunteering	2011 Census
	Fraction religious	Fraction of people with religion	2011 Census
	Fraction married	Fraction of 59–63 year olds married	2011 Census

Notes: More details on the sources can be found in Australian Bureau of Statistics (2012) (2011 Census); Australian Bureau of Statistics (2017a) (2016 Census) and Australian Bureau of Statistics (2016) (ABS small area income statistics). Census data was extracted using the Australian Bureau of Statistics TableBuilder product. The unemployment rate and NILF rate are for the age group containing the birth cohort of interest, while the fraction married is an age group roughly corresponding to the parent generation.

Table D.2
Regression table of associations between mobility and potential correlates.

	IGE	Rank-rank slope	$E[r_{1i}\|r_{0i} = 25]$	$E[r_{1i}\|r_{0i} = 75]$	$P[r_{1i} > 80\|r_{0i} \leq 20]$
Fraction Indigenous	1.70***	0.64	-1.07**	-0.75*	-0.61
	(0.61)	(0.65)	(0.42)	(0.40)	(0.52)
Ethnic diversity	0.35	0.60	-1.24**	-0.96**	0.29
	(0.68)	(0.72)	(0.47)	(0.44)	(0.58)
Ethnic segregation	0.06	-0.13	-0.16	-0.20	-0.82
	(0.59)	(0.63)	(0.41)	(0.38)	(0.50)
Income segregation	-0.16	-0.22	-0.33	-0.43	-0.38
	(0.51)	(0.55)	(0.36)	(0.33)	(0.44)
Mean commute	-0.46	-0.53	-0.50*	-0.76***	-0.29
	(0.36)	(0.39)	(0.25)	(0.24)	(0.31)
Mean income	0.23	0.45	3.04***	3.29***	3.83***
	(0.56)	(0.60)	(0.39)	(0.36)	(0.48)
Unemployment rate	-1.30	0.03	-0.48	-0.43	0.05
	(1.37)	(1.46)	(0.95)	(0.89)	(1.17)
NILF rate	0.06	1.52	-0.09	0.65	-0.00
	(1.24)	(1.32)	(0.86)	(0.81)	(1.06)
School attendance	0.21	-1.73	-2.71***	-3.54***	-3.04***
	(1.26)	(1.34)	(0.87)	(0.82)	(1.07)
Tertiary attendance	-0.19	0.15	0.61	0.67	0.47
	(0.71)	(0.75)	(0.49)	(0.46)	(0.60)
Gini	0.50	0.66	1.83***	2.16***	2.07***
	(0.71)	(0.76)	(0.50)	(0.46)	(0.61)
Top 1% share	-0.91	-1.38	-3.42***	-4.11***	-2.93***
	(0.81)	(0.87)	(0.56)	(0.53)	(0.69)
Fraction volunteers	-0.01	0.21	-0.64	-0.53	-0.71
	(0.60)	(0.64)	(0.41)	(0.39)	(0.51)
Fraction religious	0.19	0.60	0.21	0.51*	0.10
	(0.40)	(0.43)	(0.28)	(0.26)	(0.34)
Fraction married	-0.41	-1.17	2.69***	2.11***	2.47***
	(0.84)	(0.90)	(0.58)	(0.55)	(0.72)
N	87	87	87	87	87
R^2	0.49	0.53	0.80	0.82	0.82

Notes: Based on a regression of the given intergenerational mobility metric and the full set of covariates measured in 87 Australian regions. The covariates have all been normalised to have standard deviation one. Standard errors are in parentheses. Significance levels indicated as follows: * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$.

References

Acciari, P., Polo, A., Violante, G.L., 2019. And yet, it moves? Intergenerational mobility in Italy. Technical Report. CEPR Discussion Paper No. DP13646.

Australian Bureau of Statistics, 1991. Census of Population and Housing, 1991. Cat. No. 2101.0. ABS, Canberra.

Australian Bureau of Statistics, 2005. Survey of Income and Housing - Confidentialised Unit Record File, Technical Paper (Reissue), 1994-95. Cat. No. 6541.0. ABS, Canberra.

Australian Bureau of Statistics, 2010. Family Characteristics Survey, Australia, 2009-10. Cat. No. 4442.0. ABS, Canberra.

Australian Bureau of Statistics, 2011. Australian Statistical Geography Standard (ASGS): volume 1 – main structure and greater capital city statistical areas, July 2011. Cat. No. 1270.0.55.001. ABS, Canberra.

Australian Bureau of Statistics, 2012. 2011 Census of Population and Housing. ABS, Canberra.

Australian Bureau of Statistics, 2013. Microdata: household expenditure, income and housing, 2011-12. Cat. No. 6541.0.30.001. ABS, Canberra.

Australian Bureau of Statistics, 2015. Microdata: Household Expenditure, Income and Housing, 2013-14. Cat. No. 6541.0.30.001. ABS, Canberra.

Australian Bureau of Statistics, 2016. Estimates of personal income for small areas, 2012-13. Cat. No. 6524.0.55.002. ABS, Canberra.

Australian Bureau of Statistics, 2017. 2016 Census of Population and Housing. ABS, Canberra.

Australian Bureau of Statistics, 2017. Births, Australia, 2016. Cat. No. 3301.0. ABS, Canberra.

Australian Bureau of Statistics, 2017. Consumer price index, Australia, March 2017. Cat. No. 6401.0. ABS, Canberra.

Australian Bureau of Statistics, 2017. Labour force, Australia, detailed, November 2017. Cat. No. 6291.0.55.001. ABS, Canberra.

Australian Bureau of Statistics, 2017. Microdata: household expenditure, income and housing, 2015-16. Cat. No. 6540.0. ABS, Canberra.

Björklund, A., Roine, J., Waldenström, D., 2012. Intergenerational top income mobility in Sweden: capitalist dynasties in the land of equal opportunity? J. Public Econ. 96 (5–6), 474–484.

Black, S.E., Devereux, P.J., 2011. Recent developments in intergenerational mobility. In: Handbook of Labor Economics, 4. Elsevier, pp. 1487–1541.

Böhlmark, A., Lindquist, M.J., 2006. Life-cycle variations in the association between current and lifetime income: replication and extension for Sweden. J Labor Econ 24 (4), 879–896.

Boserup, S.H., Kopczuk, W., Kreiner, C.T., 2013. Intergenerational wealth mobility: evidence from Danish wealth records of three generations. Univ. of Copenhagen mimeo.

Bratberg, E., Davis, J., Mazumder, B., Nybom, M., Schnitzlein, D.D., Vaage, K., 2017. A comparison of intergenerational mobility curves in Germany, Norway, Sweden, and the US. Scand. J. Econ. 119 (1), 72–101.

Bell, B., Blundell, J., Machin, S., 2018. Where is the land of hope and glory? The geography of intergenerational mobility in England and Wales. Discussion Paper No. 1,591, Centre for Economic Performance, LSE.

Bütikofer, A., Dalla-Zuanna, A., Salvanes, K. G., 2018. Breaking the links: natural resource booms and intergenerational mobility. Discussion Paper Series in Economics 19/2018, Norwegian School of Economics, Department of Economics.

Carter, A., Breunig, R., 2019. Do earned income tax credits for older workers prolong labour market participation and boost earned income? Evidence from Australia's mature age worker tax offset. Econ. Record 95 (309), 200–226.

Chadwick, L., Solon, G., 2002. Intergenerational income mobility among daughters. Am. Econ. Rev. 92 (1), 335–344.

Chetty, R., Hendren, N., 2018. The impacts of neighborhoods on intergenerational mobility I: childhood exposure effects. Q. J. Econ. 133 (3), 1107–1162.

Chetty, R., Hendren, N., Kline, P., Saez, E., 2014. Where is the land of opportunity? The geography of intergenerational mobility in the united states. Q J Econ 129 (4), 1553–1623.

Cobb-Clark, D. A., Dahmann, S., Salamanca, N., Zhu, A., 2017. Intergenerational disadvantage: learning about equal opportunity from social assistance receipt. Life Course Centre Working Paper No. 2017-17.

Connolly, M., Corak, M., Haeck, C., 2019. Intergenerational mobility between and within Canada and the United States. J. Labor Econ. 37 (S2), S595–S641.

Corak, M., 2019. The Canadian geography of intergenerational income mobility. Econ. J..

Corak, M., Heisz, A., 1999. The intergenerational earnings and income mobility of Canadian men: evidence from longitudinal income tax data. J. Hum. Resour. 34 (3), 504–533.

Dahl, M.W., DeLeire, T., 2008. The association between children's earnings and fathers' lifetime earnings: Estimates using administrative data. University of Wisconsin-Madison, Institute for Research on Poverty.

Davis, J., Mazumder, B., 2018. Racial and ethnic differences in the geography of intergenerational mobility. SSRN Working Paper.
Deutscher, N., 2020. Place, peers and the teenage years: long-run neighborhood effects in Australia. Am. Econ. J. 12 (2), 220–249.
Güell, M., Pellizzari, M., Pica, G., Rodríguez Mora, J.V., 2018. Correlating social mobility and economic outcomes. Econ. J. 128 (612), F353–F403.
Haider, S., Solon, G., 2006. Life-cycle variation in the association between current and lifetime earnings. Am. Econ. Rev. 96 (4), 1308–1320.
Heidrich, S., 2017. Intergenerational mobility in Sweden: a regional perspective. J Popul Econ 30 (4), 1241–1280.
Leigh, A., 2007. Intergenerational mobility in Australia. BE J. Econ. Anal. Policy 7 (2).
Mazumder, B., 2005. Fortunate sons: new estimates of intergenerational mobility in the United States using social security earnings data. Rev. Econ. Stat. 87 (2), 235–255.
Mazumder, B., 2016. Estimating the intergenerational elasticity and rank association in the United States: overcoming the current limitations of tax data. Res. Labor Econ. 43, 83–129.
Mendolia, S., Siminski, P., 2016. New estimates of intergenerational mobility in Australia. Econ. Record 92 (298), 361–373.
Mitnik, P., Bryant, V., Weber, M., Grusky, D.B., 2015. New estimates of intergenerational income mobility using administrative data. Statistics of Income, Internal Revenue Service. mimeo (in preparation).
Mogstad, M., Romano, J. P., Shaikh, A., Wilhelm, D., 2020. Inference for ranks with applications to mobility across neighborhoods and academic achievement across countries. National Bureau of Economic Research Working Paper No. 26,883.
Murray, C., Clark, R.G., Mendolia, S., Siminski, P., 2018. Direct measures of intergenerational income mobility for Australia. Econ. Record 94 (307), 445–468.
Nybom, M., Stuhler, J., 2016. Heterogeneous income profiles and lifecycle bias in intergenerational mobility estimation. J. Hum. Resour. 51 (1), 239–268.
Nybom, M., Stuhler, J., 2017. Biases in standard measures of intergenerational income dependence. J. Hum. Resour. 52 (3), 800–825.
Reardon, S. F., 2011. Measures of income segregation. Stanford Center for Education Policy Analysis Working Paper.
Solon, G., 1992. Intergenerational income mobility in the United States. Am Econ Rev 393–408.
Solon, G., 1999. Intergenerational Mobility in the Labor Market. In: Handbook of labor economics, 3. Elsevier, pp. 1761–1800.