DEMOGRAPHIC RESEARCH
A peer-reviewed, open-access journal of population sciences

DEMOGRAPHIC RESEARCH

VOLUME 43, ARTICLE 1, PAGES 1–34
PUBLISHED 7 JULY 2020
https://www.demographic-research.org/Volumes/Vol43/1/
DOI: 10.4054/DemRes.2020.43.1

Research Article

Estimating annual homelessness

James O'Donnell

Contents

Demographic Research: Volume 43, Article 1
Research Article

Estimating annual homelessness

James O'Donnell[1]

Abstract

BACKGROUND
Homelessness is an important though exceedingly difficult phenomenon to measure and understand. The most common sources of data measure homelessness only on a given night or set of consecutive nights, contact with homelessness service providers, or past homeless episodes. We therefore lack an understanding of the wider impact and nature of homelessness in society.

OBJECTIVE
I set out to estimate the number of people who experience homelessness in a one-year period by duration and type of homelessness.

METHODS
A microsimulation model is used to recreate homeless episodes and impute those missed in common data sources. Model parameters are estimated using a combination of retrospective and longitudinal survey data from Australia. Administrative data from homelessness service providers are used to validate the estimates.

RESULTS
According to the results, 3.4 times as many people experienced homelessness in Australia in the 2013–2014 financial year than would have been counted on an average night. Almost one-third (32%) of episodes last for less than one month and the large majority involve 'couch surfing' or 'doubling up' with relatives or friends.

CONCLUSIONS
Homelessness and housing deprivation is more prevalent though more diverse and episodic than typically measured, affecting a large cross-section of the population and likely embedded within the dynamics of poverty and deprivation.

CONTRIBUTIONS
This research provides new estimates of the extent and duration of homelessness and housing deprivation that addresses existing data limitations and with implications for understanding the nature and impact of homelessness.

[1] School of Demography, Australian National University. Email: james.odonnell@anu.edu.au.

1. Introduction

Homelessness has traditionally been measured through point-in-time counts of people on a given night. Evidence of its relatively episodic nature (Culhane et al. 1994, 2007; Link et al. 1994, 1995), though, means that a substantially larger number of people experience homelessness over a given period, whether it be 1 year, 10 years, or a lifetime, than are counted on a single night. Research has struggled to quantify homelessness over longer periods. In recent decades, studies have utilised cross-sectional household surveys that ask about past experiences of homelessness but do not capture current experiences (e.g., Link et al. 1994) and administrative data from homelessness service providers, which only capture homelessness where individuals come into contact with the service system (e.g., Culhane et al. 1994). Longitudinal surveys of disadvantaged and homeless populations are becoming more common and provide better coverage of homelessness (e.g., Cobb-Clark et al. 2016). However, their targeted samples make it difficult to extrapolate findings to the general population, limiting their applicability to demographic estimation.

In this paper, I propose a method to estimate current homelessness from household surveys and so derive more complete estimates of annual homelessness. The method is demonstrated using the 2014 General Social Survey of Australia (ABS 2015). The research questions motivating this study are

1. How many people experienced homelessness in 2013–2014 in Australia?
2. How is homelessness experienced in terms of duration and form?

Responses to these questions provide information on the extent and duration of homelessness. Point-in-time counts and household surveys underestimate the true population exposed to homelessness to differing degrees, and so potentially understate the level of housing market volatility especially for low-income populations. Further, point-in-time estimates over-represent long-term and chronic homelessness (Metraux et al. 2001; Chamberlain and Johnson 2015), while retrospective household surveys tend to under-represent it. In both cases, this is because the chronic and long-term homeless are more likely to be homeless on any given night, giving them a higher probability of being included in point-in-time counts and excluded from household surveys. As a consequence, the two data sources produce different profiles of homeless durations, and to the extent that durations are positively associated with personal vulnerabilities (e.g., Wong and Piliavin 1997; Culhane and Kuhn 1998; Cobb-Clark et al. 2016), different sociodemographic profiles and drivers of homelessness and the people who experience it.

2. Background

Most studies estimate homelessness on the streets and in shelters at given points in time. These typically involve physical street searches by teams of volunteers (e.g., Rossi 1989), counts of people utilising homelessness services such as shelters and soup kitchens (e.g., Burt 1992; Firdion and Marpsat 2007), and/or secondary analysis of administrative data from service providers (e.g., Chamberlain 1999; Coumans et al. 2017). Researchers increasingly utilise a range of post-enumeration strategies to measure and adjust for undercounts, including capture-recapture and plant-capture (Coumans et al. 2017; Darcy and Jones 1975; D'Onise, Wang, and McDermott 2007; Hopper et al. 2008). Nevertheless, these represent estimates of homelessness that are static in time, sometimes giving the impression that homelessness itself is permanent and unchanging.

Evidence, though, suggests that homelessness can be highly episodic. This nature has been revealed through analyses of administrative data and targeted longitudinal sample surveys that compare prevalence rates in homeless shelters over different period lengths (Culhane et al. 1994), analyse the characteristics of people by their episode lengths (Chamberlain and Johnson 2013; Kuhn and Culhane 1998), and measure entries, exits, and returns to and from homelessness (Cobb-Clark et al. 2016; Culhane and Kuhn 1998; Metraux and Culhane 1999; Wong and Piliavin 1997). Since individuals and families experience homelessness episodically, moving in and out of homelessness through time, more people experience it than are counted on a single night. In the early 1990s, Link et al. (1994; 1995) conducted household surveys that asked respondents about past experiences of homelessness. These revealed substantially higher lifetime and five-year rates of homelessness than comparable point-in-time estimates. These studies have been replicated, with similar results, in the United States (Fusaro, Levy, and Schaefer 2018; Greenberg and Rosenheck 2010; Ringwalt et al. 1998; Tompsett et al. 2006; Tsai 2017), the United Kingdom (Bramley and Fitzpatrick 2018; Burrows 1997), Australia (ABS 2011, 2015; Chamberlain and Johnson 2015), Belgium, Germany, and Italy (Toro et al. 2007).

Estimation of annual rates of homelessness are restricted by the limitations of available data. A potentially large yet unmeasured source of bias in homelessness estimates from households surveys arises from the fact that survey sample frames are generally only of individuals living in private households, thus largely excluding those who are homeless at the time of the survey. The bias is relatively small in lifetime rates, but increases as the period of interest shortens to five years or one year. The bias is such that Shinn (2010: 20) argues this makes "surveys worthless for estimating current homelessness." Administrative data, on the other hand, are limited primarily by the fact that not all individuals experiencing homelessness interact with service providers.

While the ABS (2015), for example, estimated that 351,000 adults experienced a completed episode of homelessness in Australia in the year ending June 2014, only 93,000 adults were recorded in homelessness services data in the same financial year where they were recorded as homeless either at the start, end, or during their support period (AIHW 2018). Thus, while the greater prevalence of homelessness over increasing lengths of time is well understood, reliable estimates are difficult to derive.

3. Data

The Australian Bureau of Statistics (ABS 2015) conducted the fourth edition of the General Social Survey (GSS) from March to June 2014. The sample frame included all private dwellings in Australia, except those in very remote parts of Australia and discrete Indigenous communities. An initial sample of 18,574 dwellings were randomly selected from geographic areas, with low socioeconomic areas oversampled. Households residing in the sampled dwellings were contacted and an individual was randomly chosen to participate from within each. The final sample consisted of 12,932 people aged 15 years and over, giving a response rate of 80%. Interviews were conducted face to face. People who were experiencing homelessness at the time of the survey were excluded from the sample, as were people staying in non-private dwellings.

Data were analysed in the ABS DataLab, a remote access system designed to allow approved researchers access to unit record data. Outputs were checked and approved by ABS officers before release to protect against breaches of confidentiality. The dataset contains population weights scaled to match the total estimated in-scope population. A set of 60 Jackknife replicate weights were also provided to generate standard errors.

The Journeys Home study is used as an auxiliary data source. This was a six-wave survey conducted in Australia over a two-and-a-half-year period between 2011 and 2014 (Wooden et al. 2012; Scutella, Tseng, and Wooden 2017). The sample frame consisted of approximately 110,000 adults who had received an income support payment such as for unemployment, disability, or parenting support in May 2011 and who had been flagged in the income support system as homeless or at risk of homelessness or who were predicted by the survey design team to be vulnerable to homelessness. A multistage clustered survey design yielded a final sample of 1,682 adults at a response rate of 62%. The retention rate to wave 6 was 84%. Included in the survey was an accommodation calendar in which respondents listed the types of accommodation that they had stayed in each 10-day block since the previous survey (or the start of their current accommodation at wave 1). Thus, assuming accurate recall, the accommodation status of each respondent can be approximated for any given block during the survey period.

4. Defining and measuring homelessness

4.1 Defining homelessness

A clear definition of homelessness is critical for placing the results of this study in proper context. Most research over the last 40 years has focused on capturing people who are living on the streets or other public spaces and staying in homeless shelters. Over the years, though, many researchers and advocates have sought to consider a broader range of situations in which people might be considered homeless. This is typified by the European Typology of Homelessness and Housing Exclusion (ETHOS) developed by the European Federation of National Organisations Working with the Homeless and the European Observatory on Homelessness (Amore, Baker, and Howden-Chapman 2011). Under ETHOS, people living rough and staying in shelters are considered to be experiencing 'rooflessness,' which in addition to people staying in other types of supported accommodation and those due to be released from institutional facilities, including prisons, constitute 'homelessness.' People experiencing 'housing exclusion' under ETHOS include people staying in insecure and inadequate accommodation such as those staying temporarily doubled up with family and friends, in mobile homes, and in violent, insecure, or crowded dwellings.

In Australia, most data on homelessness have been collected under relatively broad definitions that tend to align to varying degrees with ETHOS. The ABS (2015) uses what is known as the 'statistical' definition for the GSS (ABS 2012a), while the designers of Journeys Home used the 'cultural' definition (Chamberlain and Mackenzie 1992). While the statistical and cultural definitions are conceptually different, the operational categories of homelessness are similar between the GSS and Journeys Home. Therefore, the following definition is used in this study to best align how homelessness is operationalised in the GSS and Journeys Home while remaining grounded in past and emerging international research:

> Situations where individuals – in their own view – do not have a permanent place to live due to reasons outside their control or personal choice. This is comprised of situations that can be described as 'literal' and 'hidden' homelessness.
>
> Literal homelessness includes situations where individuals
>
> 1. live on the streets, in cars, abandoned buildings, and other makeshift dwellings not usually suited to human habitation, or
> 2. stay in a homeless shelter or refuge.

Hidden homelessness includes situations where individuals temporarily stay

1. in temporary private sector accommodation, including single-room occupancies, boarding/lodging houses, and hostels;
2. in caravans/mobile homes; or
3. with other households, typically family or friends in a 'doubled-up' or 'couch surfing' arrangement.

Debate has long centred on whether hidden homelessness truly constitutes homelessness (e.g., Rossi 1989). Aside from being qualitatively different experiences from literal homelessness and requiring different policy responses, large populations live in mobile homes and with family and friends in permanent and stable arrangements, or do so because it suits their immediate interests and housing and life goals. Therefore, it is important to emphasise that people are counted as experiencing homelessness in this study only if they did not regard their accommodation as being a permanent place to live and were staying there for reasons outside of their control or choice. Furthermore, forms of hidden homelessness arguably exist on a spectrum of housing deprivation. Even if not considered as 'homelessness' in their own right, they are important to understand and measure as forms of disadvantage and potentially bridges between secure housing and literal homelessness (O'Donnell 2019; Watson 1984). Separate estimates of literal and hidden homelessness are provided in this paper to gauge the magnitude of each and provide a basis for international comparison.

4.2 Measuring homelessness

Homelessness is operationalised in the GSS through two questions. These are shown in Table 1. The first question asked respondents whether they had ever stayed in any of the accommodation types listed because they did not have a permanent place to live. The second asked for all the reasons they did not have a permanent place to live. Respondents are counted as having experienced homelessness if they had ever been without a permanent place and cited at least one of the reasons outside of the first six listed. Thus, this question (imperfectly) filters out of the homelessness count people who have only ever been without a place to live for reasons that are deemed to be related to a personal – and unconstrained – choice.

Homelessness was measured in the Journeys Home survey by asking respondents about their accommodation and tenure type at the time of the survey and since the previous survey. In this study, forms of accommodation are aggregated into one of the five types of literal and hidden homelessness. Survey records spent in secure housing are excluded from this study, leaving a subsample of 1,026 people who were recorded as homeless at some stage during the survey period. Information on the type, timing, and duration of homeless episodes are used as an auxiliary source to provide greater detail to the data generated from the GSS. As such, the resulting estimates rest largely on the GSS operationalisation of homelessness. Underlying the estimates is an assumption that differences in the way homelessness is defined and measured in the Journeys Home survey do not substantively affect the results.

Table 1: Homelessness questions in the GSS

Question	Response categories	
Have you ever experienced any of these things	1. Stayed with relatives	
because you did not have a permanent place to live?	2. Stayed at a friend's house	
	3. Stayed in a caravan (mobile home)	
More than one response is allowed.	4. Stayed at a boarding house/hostel	
	5. Stayed in a night shelter	
	6. Stayed in a shelter for the homeless	
	7. Stayed at a refuge (e.g., women's shelter)	
	8. Squatted in an abandoned building	
	9. Slept rough (include sleeping in cars, tents, etc.)	
	10. Stayed in a detention centre	
	11. Other (Please specify)	
	12. No	
What led to you being without a permanent place to	1. Travelling/on holiday	
live?	2. Work-related reason	
	3. House-sitting	
More than one response is allowed.	4. Saving money	Not homeless
	5. Just moved back/into town or city	
	6. Building or renovating home	
	7. Tight housing/rental market	
	8. Violence/abuse/neglect	
	9. Alcohol or drug use	
	10. Family/friend/relationship problems	
	11. Financial problems (e.g., not able to pay mortgage or rent)	
	12. Mental illness	
	13. Lost job	
	14. Gambling	
	15. Eviction	
	16. Natural disaster	
	17. Refugee	
	18. Damage to house (e.g., house fire)	
	19. Health issues	
	20. Other (Please specify)	

Source: ABS (2015).

Responses to the GSS homelessness questions are provided in Table 2. Based on the survey results, 13.4% of the population are estimated to have ever been without a permanent place to live due to a reason constituting homelessness. An estimated 1.9% of people living in private dwellings at the time of the survey experienced homelessness in the 12 months prior to the survey. The most commonly cited reason for homelessness was relationship problems. The most common accommodation type among those who have ever been homeless is staying with family or friends.

Table 2: Timing and length of most recent homelessness episodes, GSS 2014

	Weighted % of population	95% confidence interval Lower	Upper
Reasons for ever being without a permanent place to live			
Non-homelessness reason			
Travelling/on holiday	3.4	2.9	3.9
Work-related reason	2.8	2.3	3.2
House-sitting/saving money	2.8	2.4	3.2
Just moved back/into town or city	8.0	7.2	8.7
Building or renovating home	2.6	2.2	2.9
Homelessness reason			
Tight housing/rental market	2.0	1.6	2.3
Violence/abuse/neglect	1.4	1.1	1.7
Mental illness/alcohol or drug use	0.9	0.7	1.1
Family/friend/relationship problems	6.7	6.1	7.3
Financial problems/lost job	2.9	2.5	3.3
Other	3.9	3.3	4.5
Total homelessness	13.4	12.4	14.5
Accommodation types of those ever homeless			
Stayed with family/friends	11.9	10.9	12.9
Stayed in a homeless shelter or refuge	1.1	0.9	1.3
Slept rough/squatted	2.1	1.7	2.4
Stayed in a caravan/mobile home	2.5	2.1	2.9
Stayed at a boarding house/hostel	1.7	1.4	2.0
Other	0.5	0.3	0.7
Timing of most recent homelessness episode			
Less than 12 months ago	1.9	1.5	2.3
Less than 2 years ago	3.1	2.5	3.6
Less than 5 years ago	5.4	4.6	6.1
Less than 10 years ago	7.7	7.0	8.5
Ever	13.4	12.4	14.5
Length of most recent homelessness episode (last 10 years)			
1 day to less than 1 week	1.1	0.8	1.4
1 to less than 2 weeks	0.7	0.5	1.0
2 weeks to less than 1 month	1.0	0.7	1.3
1 to less than 1 month	1.0	0.7	1.3
2 to less than 3 months	0.9	0.7	1.2
3 to less than 6 months	1.3	1.1	1.6
6 to less than 12 months	1.2	1.0	1.5
12 months or more	1.1	0.8	1.4
Number of times ever without a permanent place to live			
One time	7.7	6.9	8.5
Two times	2.0	1.6	2.4
Three times	1.1	0.8	1.3
Four times	0.4	0.3	0.6
Five or more times	1.2	1.0	1.4

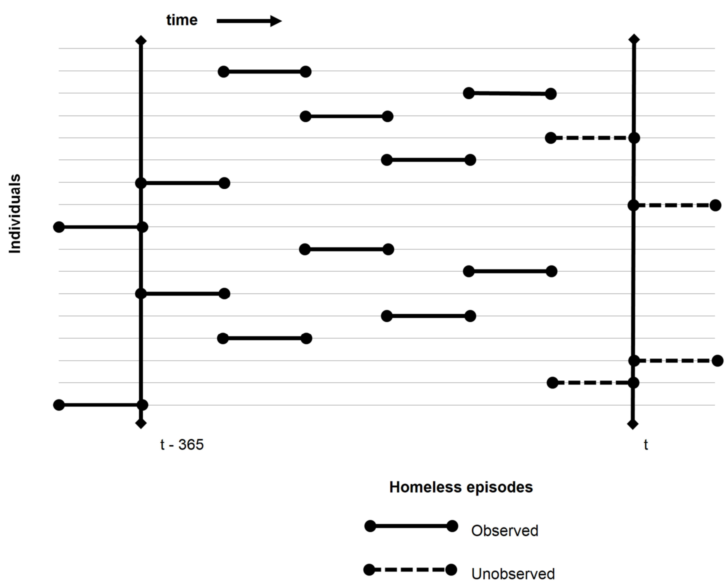
Source: ABS (2015).

5. Methods

5.1 Estimating ongoing homeless episodes

As discussed, people experiencing homelessness are typically missed from estimates derived from retrospective household surveys. In this section, I propose a method to address this deficit and provide a more complete estimate of past and current homelessness. The proposed method estimates ratios of homeless episodes missed by the survey to those captured. These ratios are referred to as k ratios. A hypothetical example is shown in Figure 1. Individuals on the vertical axis experience homeless episodes across a 12-month period (horizontal axis). They are selected to participate in the GSS at time t. If their homeless episode finishes by time t, they are eligible for inclusion in the sample and therefore represented in the results. Thus, the episode is 'observed' (episodes with a solid line in Figure 1). If the episode is ongoing at the time of sample selection, they are excluded from the sample frame and their episode is unrepresented and 'unobserved' (episodes with a dashed line). The ratio of unobserved to observed episodes is the k ratio. If this ratio can be calculated, it can be multiplied by and added to the observed number of episodes to estimate total homelessness across the year.

Figure 1: Homeless episodes for a hypothetical population



One means of calculating k ratios relies on survey information on the timing and length of past episodes. To explain, consider the proposition that the population who experience homelessness in the 12 months leading up to a household survey has a probability of being homeless at the time of sample selection. Conversely, this population has a probability of exiting homelessness into a private household in time to be eligible for sample selection. These probabilities are related to the timing and duration of episodes over the year. People who experience homelessness earlier in the year and for relatively short durations are more likely to exit homelessness in time to be represented in the survey, while those who experience longer episodes are less likely to be represented.

If episodes are evenly distributed in time through the year, probabilities of being homeless (or not) can be calculated mathematically. A person who experiences homelessness for one week, for example, has only a 1-in-52 chance of being homeless on a given night during the year, compared with a 40-in-52 chance for someone homeless for 40 weeks. Conversely, the chances of not being homeless on a given night is 51-in-52 for a person homeless for one week and 12-in-52 for a person homeless for 40 weeks. The odds of these probabilities give the k ratios. For persons experiencing homelessness for one week in a year, for example, dividing the probability of being homeless on a survey night (1/52 chance) by the probability of not being homeless (51/52 chance) gives a k ratio of 0.02. Thus, for every person who experienced homelessness for one week in the past 12 months, 0.02 people are currently experiencing an episode that will last for one week. Multiplying duration-specific k ratios by the survey estimate of the number of people previously homeless gives an estimate of the number of people currently homeless for that duration.

5.2 Estimating *k* ratios from retrospective survey data

Several aspects need to be taken into account before calculating k ratios. First, people can experience multiple episodes over a year, while surveys such as the GSS ask about the duration of the most recent episode, rather than the total number of episodes or time spent homeless over the year. Second, episode durations are reported in ranges rather than exact values. As shown in Table 2, there are eight reported duration ranges in the GSS from less than one week to 12 months and over. Third, episodes may have commenced in previous years as will certainly be the case for episodes longer than one year. Fourth, completed episodes may not be captured in the GSS where individuals die or emigrate before the survey period or are living in non-private dwellings (for example, prisons and hospitals) or in remote areas that are not covered by the survey.

Finally, trend or seasonal variation may affect the timing of episode commencement and duration. Fluctuations in housing markets, social and interpersonal support, and/or the weather, for example, could affect the number of entries to, exits from, and prevalence of homelessness at given points across the year. Thus, the relationship between the annual number of people homeless and the number homeless during the survey period will vary across the year.

The approach taken to dealing with issues of seasonality, coverage, and mortality in this study is to develop and run a microsimulation model. The idea is to generate a set of k ratios by simulating a large number of homeless episodes in each duration range (in this case, 50,000 in each of the eight ranges) and calculating how many occur prior to and overlap the date of sample selection for the GSS. A unique sample selection date is randomly chosen for each individual between 1 March and 30 June 2014 to mirror the GSS survey period. Homeless episodes are simulated by generating a start date and exact duration, from which an end date can be calculated (start date plus duration). Individuals can join a private household and be eligible for and represented in the GSS if their episode of homelessness finishes in the 12 months before the date of sample selection, but are ineligible if their episode overlaps this date. Dividing the number of episodes that overlap sample selection by the number that finished in the prior 12 months for each duration range gives a set of simulated k ratios.

Homeless episodes are simulated by first generating an exact duration. For example, episodes of 1 to 7 days are assigned a value of one, two, three, four, five, or six days. This assignment is achieved by constructing a homeless life table from the Journeys Home accommodation calendar. This is similar to a standard life table, except that in place of age, each row represents a single day of homelessness, and in place of mortality, life table decrements consist of exits from homelessness. The life table commences at day one, as all episodes are assumed to last at least one day, and ends with episodes lasting five years or longer. Probabilities of episodes concluding between day x and $x + 1$, q_x, are estimated by dividing the number of episodes that end on day x, d_x by the number of episodes that last to at least x days, H_x:

1. $q_x = \frac{d_x}{H_x}$.

Homelessness 'survivorship,' l_x, is estimated as

2. $l_1 = 1$
   $l_x = l_{x-1} \times (1 - q_{x-1})$.

In words, the proportion of people who continue to experience homelessness after x days, l_x, is equal to the number who were homeless the previous day minus those who exited homelessness over the day. Values of l_x are produced for all days from day 1 to day 1,825 (approximately five years) to give a homeless episode 'survivorship' function, a synthetic estimate of the proportion of people who remain homeless at the beginning of each day between day one and year five. Separate survivorship functions are estimated for episodes that do and do not involve literal homelessness.

Homeless episode survivorship functions are applied to the microsimulation to simulate the duration of episodes within each range. A random number is generated for each individual that ranges between the values of l_x for the minimum and maximum number of days in that duration range. Random numbers for episodes that last 1 to 2 months, for example, range between l_{30} and l_{61}. If the random number is less than or equal to l_x and greater than l_{x+1}, then the homeless episode is simulated to last for x days. Durations for the two shortest duration ranges (less than one week and one to two weeks) were not simulated in this way due to small sample sizes and the fact that Journeys Home respondents reported accommodation episodes in 10-day blocks. For these ranges, durations were simulated from uniform distributions. In other words, episodes of 1 to 7 days and 7 to 14 days are assigned any of the days within these ranges with equal probability.

The start date of homeless episodes are simulated to take account of seasonal variation and the possibility that people experience multiple episodes in a year or single episodes that last across multiple years. The date range in which episodes are allowed to commence is established by setting the maximum possible date ($date_{max}$) to the date of sample selection. The minimum date ($date_{min}$) is set to 365 days plus the simulated duration (l) of the episode prior to the date of sample selection (t):

3. $$date_{min} = t - 365 - l.$$

This allows for the fact that episodes may commence in previous years but must occur for at least some moment in time in the 12 months prior to the survey.

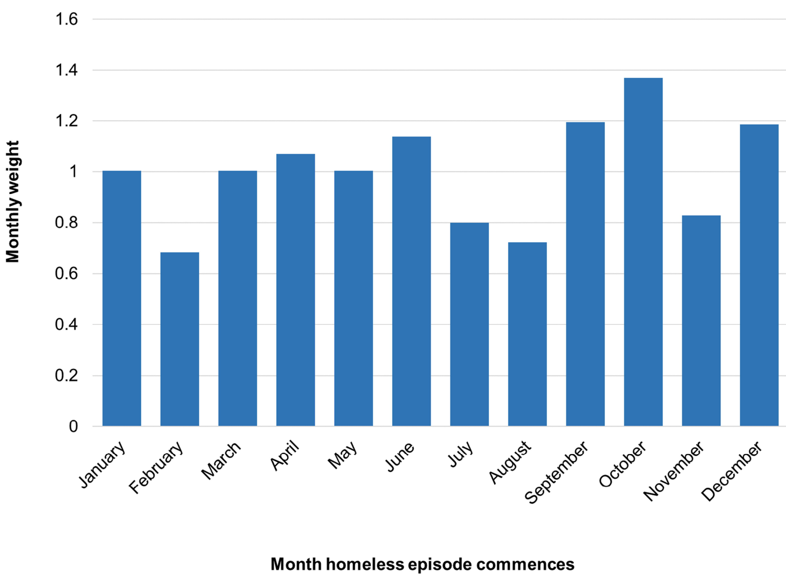
The probability of episodes beginning at any given date in the range is determined by applying a set of monthly weights. The weights allow the model to account for seasonal variation and multiple episodes. They are calculated from Journeys Home by first summing the numbers of homeless episodes that occurred in 2012–2013 and 2013–2014 by the month in which they commenced. Multiple episodes that occur in the year are subtracted to leave only individuals' most recent episodes (or the episodes that occurred closest to the end of the financial year) in the counts. The counts in each month are then divided by the average number of episodes commencing across the two years.

The resulting monthly weights are applied to the microsimulation. They determine the probability of an episode beginning in any particular month within the date range

$date_{min}$ to $date_{max}$. For example, an episode that has 13 months in the date range has a probability of commencing in any particular month equal to the weight divided by 13 and standardised so that the sum of weights across the date range equals one. Start dates are assigned by drawing a random number between zero and one: For example, if the probability of an episode commencing in September 2013 is equal to 8% and the random number for a particular episode is 0.4, then that episode is scheduled to commence on 15 September 2013, and so on.

Monthly weights are shown in Figure 2. Note that these weights attempt to adjust for monthly variation in homelessness occurrence but assume the underlying trend in homelessness rates over the reference period is zero. The weights suggest that entries to homelessness may be more common towards the end of the year (the Australian spring), though the differences may be driven by stochastic variation or other factors unique to the Journeys Home survey. Nevertheless, weights are used in their raw form as they provide the best evidence of seasonality.

Figure 2: Monthly homelessness incidence weights



Source: Author's estimates from Wooden et al. (2012) and Scutella, Tseng, and Wooden (2017).

Other parameters and population characteristics of the microsimulation are established through a combination of data sources. These parameters are age (in five-year age groups), sex, a probability of dying in the year leading up to the GSS, and a probability of being out of scope at the time of GSS. Individuals are simulated to die where a randomly drawn number is less than or equal to their age and sex-specific mortality probability, extracted from national Life Tables for 2013–2015 (ABS 2016), and be out of scope where a random number is less than or equal to the proportion of the total population by age who are not in the scope of the GSS. This conservatively assumes rates of homelessness among those who die or are out of scope are the same as for the total population. Separate estimates of homelessness among these groups are provided in the Results section to gauge the possible impact of this assumption.

The final step in simulating the k ratios and estimating the number of missed homeless episodes is to weight the microsimulation results by age, sex, and duration. The microsimulation produces (1) completed homeless episodes in the 12 months prior to the GSS, (2) episodes ongoing at the time of the GSS, and (3) episodes that are otherwise out of scope and missed by the GSS. The number of episodes in each of these three categories is summed by age, sex, and duration range. Dividing the number of ongoing episodes by the number of completed episodes gives an estimated k ratio by age, sex, and duration range. Likewise, the ratio of out-of-scope to completed episodes is also calculated. Both sets of ratios are multiplied by the predicted number of people who experienced a completed homeless episode in 2013–2014 (see the Appendix for how these predictions are derived) to give estimates of the number of people experiencing homelessness at the time of the GSS and the number of out-of-scope episodes by age, sex, and duration range. Adding the number of completed, ongoing, and out-of-scope episodes thus gives an estimate of the total number of people who experienced homelessness in 2013–2014.

5.3 Estimating different forms of homelessness

Estimation of different forms of homelessness is difficult due to sample size issues and confidentiality restrictions in the GSS. In particular, the redacted unit record file made available for public use does not contain the accommodation type experienced during respondents' most recent homelessness episode. It includes only the accommodation type ever experienced during any episode without a permanent place to live. Interestingly, the majority of respondents who have been homeless in the last 10 years reported being without a permanent place only once in their lives. Table 3 presents survey estimates of the population experiencing homelessness in the last 10 years (excluding the missing homeless) by the accommodation type ever stayed in and the number of lifetime episodes without a permanent place to live.

Accommodation types can be computed directly from the available information for people who have experienced only one episode. This is because any accommodation ever experienced must relate to their most recent (and only) experience. For respondents with multiple past experiences, accommodation in their most recent episode is simulated in a separate model by assigning them accommodation types for each episode they have ever experienced. The proportion of simulated individuals who experience each accommodation type in their most recent episode of homelessness is then applied to the main part of the microsimulation as a set of probabilities to assign accommodation types to each episode. See the Appendix for more information on how these calculations are performed.

Table 3: Accommodation types ever experienced by number of lifetime homeless episodes

	Number of episodes without a permanent place to live				
Ever stayed in accommodation type	1	2	3	4	5+
	Estimated population (000s) homeless in last 10 years (std. error)				
Stayed with family/friends	2,731 (93)	775 (51)	340 (35)	134 (22)	333 (31)
Stayed in a caravan/mobile home	266 (36)	164 (23)	74 (14)	42 (11)	125 (16)
Stayed at a boarding house/hostel	156 (22)	89 (17)	72 (13)	27 (7)	125 (18)
Stayed in a shelter/refuge	47 (11)	37 (9)	32 (9)	14 (6)	92 (13)
Slept rough/squatted	137 (21)	61 (11)	70 (13)	25 (7)	181 (18)
Total homeless in last 10 years	3,027 (98)	819 (54)	367 (35)	146 (22)	358 (31)

Note: This table is based on calculations on the raw survey data, so it excludes estimates for the missing homeless.
Source: ABS (2015).

6. Results

The results suggest that just over 500,000 adults experienced literal or hidden homelessness in Australia in 2013–2014. The standard error (s.e.) around this estimate is 65,700. This equates to 2.7% of the total population. Of these, 52,300 (s.e. 12,000), or 0.28% of the population, are estimated to have experienced literal homelessness. Key results are presented in Table 4. The number previously homeless in 2013–2014 (column 2) is derived from those captured in the GSS results, estimated at 33,100 (s.e. 6,800) people for literal and 351,000 (s.e. 41,000) for total homelessness. The currently homeless (column 4) is the number of adults who experienced homelessness at the time of the GSS. An estimated 148,400 people (s.e. 34,300) were experiencing homelessness at the time of the GSS, of whom 18,000 (s.e. 5,800) had experienced literal homelessness. The 148,400 estimate also approximates a point-in-time or nightly estimate of homelessness, equating to 0.79% of the adult population. The estimated number of people who experienced homelessness and were out of scope at the time of the GSS (column 6 of Table 4) is relatively small (though subject to considerable uncertainty).

Table 4: Microsimulation estimates of adult annual homelessness, Australia, 2013–2014

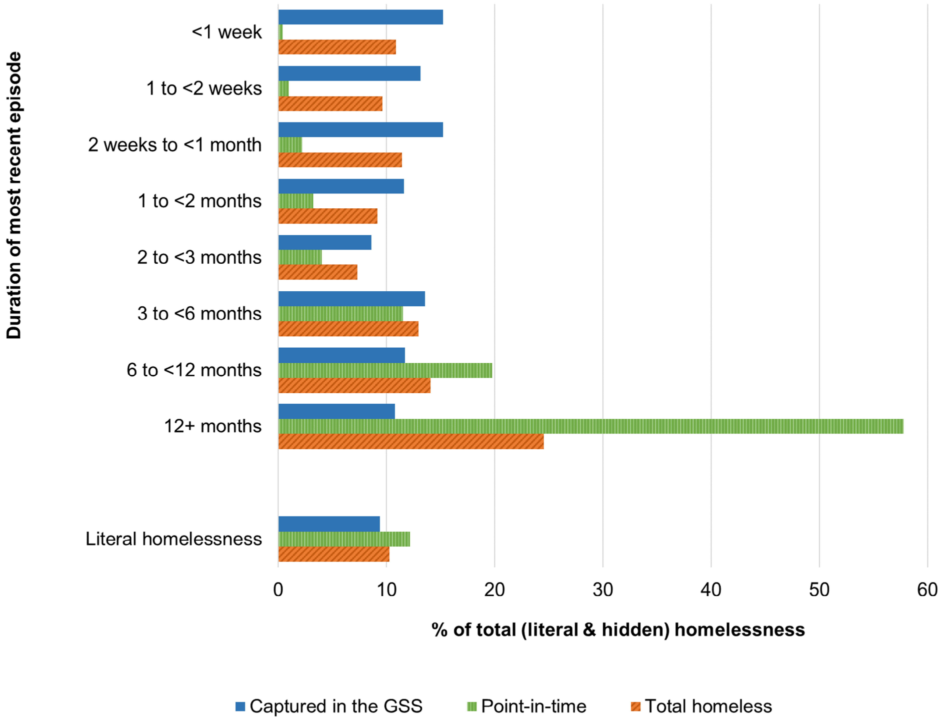
Duration of most recent episode	Homeless last 12 months (000s)	*k* ratio	Currently homeless (000s)	Out-of-scope homeless (000s)	Total homeless (000s)	Std. error (000s)
Literal homelessness						
<1 week	3.70	0.013	0.05	0.11	3.85	1.6
1 to <2 weeks	3.75	0.026	0.10	0.09	3.94	1.5
2 weeks to <1 month	3.88	0.055	0.22	0.10	4.20	1.8
1 to <2 months	3.27	0.133	0.44	0.13	3.83	1.8
2 to <3 months	5.09	0.190	0.97	0.22	6.28	2.9
3 to <6 months	4.94	0.388	1.92	0.24	7.10	2.4
6 to <12 months	4.05	0.693	2.81	0.10	6.95	2.5
12+ months	4.38	2.639	11.6	0.25	16.19	6.5
Total	33.1	0.546	18.0	1.24	52.3	12.0
Literal and hidden homelessness						
<1 week	53.5	0.011	0.60	1.29	55.4	16.1
1 to <2 weeks	46.2	0.032	1.46	1.26	48.9	18.0
2 wks to <1 month	53.5	0.061	3.28	1.25	58.1	16.0
1 to <2 months	40.9	0.118	4.83	1.02	46.7	18.7
2 to <3 months	30.3	0.198	6.01	0.79	37.1	10.5
3 to <6 months	47.7	0.358	17.1	1.22	66.0	14.1
6 to <12 months	41.1	0.714	29.4	0.92	71.4	21.5
12+ months	37.8	2.269	85.8	1.12	124.8	45.3
Total	351.0	0.423	148.4	8.88	508.3	65.7

Source: Author's estimates from ABS (2015); Wooden et al. (2012); and Scutella, Tseng, and Wooden (2017).

Simulated k ratios are shown in column 3 of Table 4. As can be seen, k ratios increase sharply with episode duration, from 0.01 for episodes less than one week to 2.27 for episodes 12 months and longer. Thus, for every person captured in the GSS who had been homeless for a year or more in 2013–2014, 2.27 people are estimated to have been in the midst of a long-term episode at the time of the GSS. The ratio is higher for people who experience literal homelessness, suggesting these people experience longer episodes on average. The estimated k ratio across all durations indicates that for every person captured in the GSS, 0.42 people were experiencing homelessness at the time. To put another way and before accounting for the out-of-scope episodes, total annual homelessness is predicted to be 42% higher than captured in a retrospective survey and 3.4 times higher than a point-in-time estimate.

The variation in k ratios across episode durations has strong implications for the measured severity of homelessness. According to the results, homeless episodes captured in the GSS are substantially shorter on average than all episodes that occurred over the year. Conversely, a point-in-time count during this time is predicted to capture a disproportionately large number of people who will be experiencing very long episodes. This is shown in Figure 3. By these estimates, 25% of the population who experienced homelessness in 2013–2014 were homeless for 12 months or longer, compared with 58% of the point-in-time population and 11% of the population captured in the GSS. At the other end of the scale, just 4% of a point-in-time count would be in the midst of an episode lasting less than one month, compared with 44% of those captured in the GSS and 32% of total annual homelessness. Consequently, the GSS may somewhat understate the population who experience literal homelessness over the year. Conversely, the seemingly more episodic nature of hidden homelessness may mean that less hidden homelessness is captured in a point-in-time count than would be counted in an annual estimate.

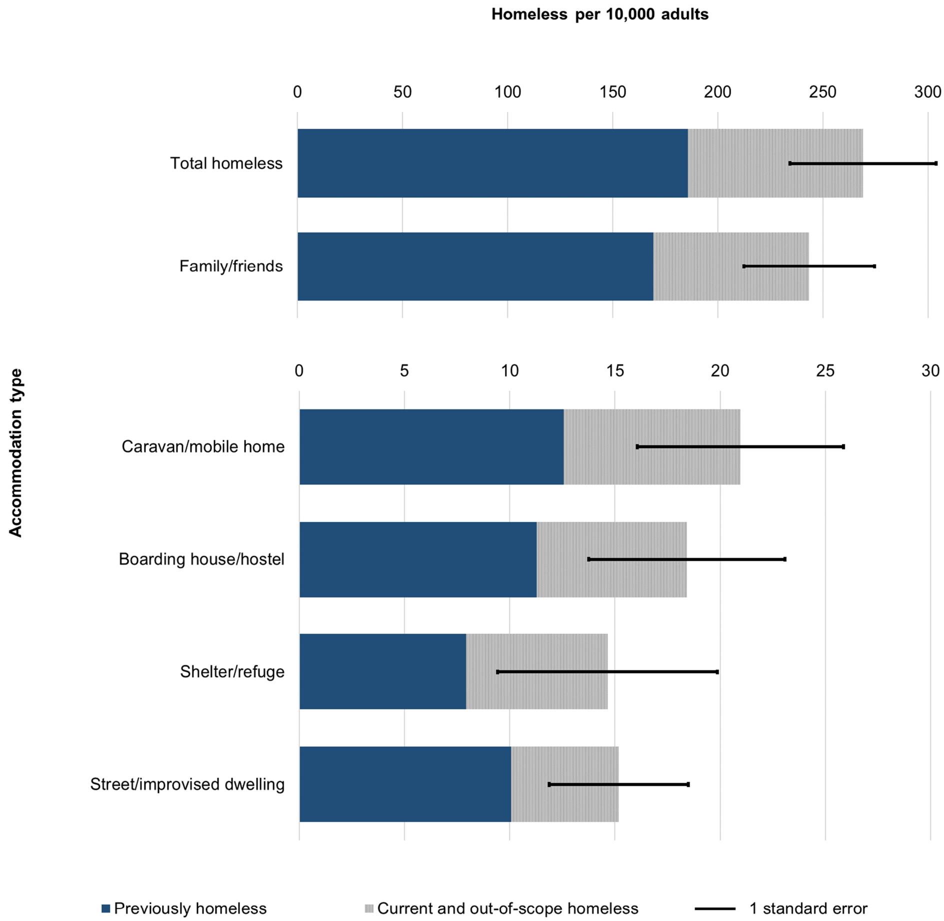
Figure 3: Predicted homeless episode durations, 2013–2014



Source: Author's estimates from ABS (2015); Wooden et al. (2012); and Scutella, Tseng, and Wooden (2017).

Estimated annual rates of homelessness in each accommodation category are shown in Figure 4. Strikingly, 459,900 (s.e. 58,900) people, or 243 in 10,000 adults in the general population, are predicted to have stayed with family or friends while homeless in 2013–2014. This equates to 90% of the total homeless estimate. Estimates for the other accommodation types are substantially smaller and have relatively large standard errors. Interestingly, though, the most common of these is staying in caravans and mobile homes (21 per 10,000 adults; s.e. 4.9) and boarding houses and hostels (18 per 10,000; s.e. 4.7). The rates for homelessness in shelters/refuges and on the street/improvised dwelling are 15 per 10,000 people each (s.e. 5.2 and 3.3, respectively). The proportion of nightly to total homelessness is relatively high for shelters/refuges (0.43), caravans/mobile homes (0.38), and boarding houses/hostels (0.37) compared with total homelessness (0.29), suggesting that homelessness in these categories is likely to be over-represented in point-in-time counts and under-represented in retrospective measures.

Figure 4: Annual homelessness rates by accommodation type, 2013–2014



Source: Author's estimates from ABS (2015); Wooden et al. (2012); and Scutella, Tseng, and Wooden (2017).

7. Validating and comparing the estimates

One way to validate the estimates is to compare them against administrative data from Australia's homelessness services system. The GSS asks whether respondents sought assistance from housing and homelessness services during their most recent episode:

> *"Did you seek assistance from services such as these?*
> *Housing service providers*
> *Crisis accommodation/supported accommodation for the homeless (e.g., shelter, women's refuge etc.)"* ABS (2015: 100)

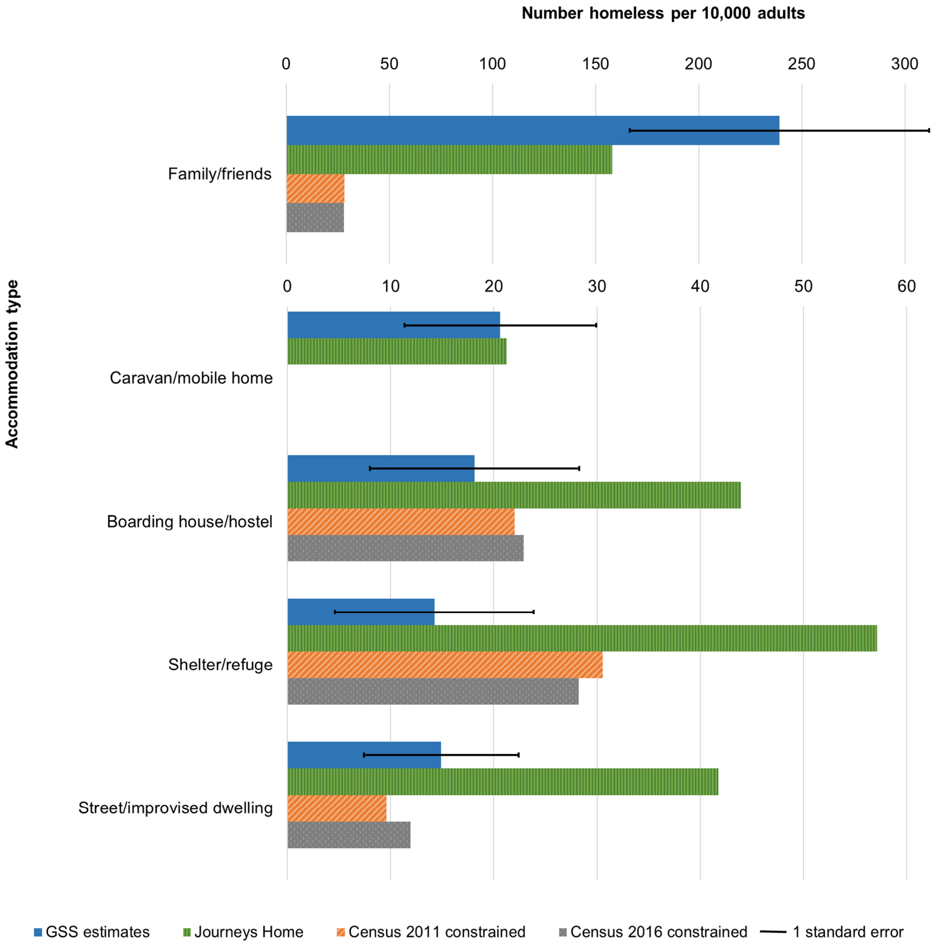
Other types of services provided in response categories include mental health services, church or community organisations, hospitals, and the police. Note that this asks whether respondents sought help from housing or homelessness services. It is not known how these responses concord with actual presentations to services which are likely to be the product of referrals from other services, outreach in which services providers seek out clients, and family and group presentations, in addition to self-referrals. Respondents were allowed to provide multiple responses. Of the annual homeless population, 20% (s.e. 3.7%) are predicted to have sought assistance from housing and homelessness services in 2013–2014. This amounts to 103,000 adults (s.e. 21,900).

This prediction is compared to the equivalent figure in the Specialist Homelessness Service Collection (SHSC), the national data collection for government-funded homelessness service providers in Australia (AIHW 2018). In 2013–2014, 197,400 adults sought or received services. Of these, 93,400 were recorded as either homeless at first or final presentation or at some stage during a support period. There were another 33,900 adults whose homelessness status was not recorded. This may have been because an initial contact was not maintained or followed up. Their homelessness status is imputed by assuming they have the same rate of homelessness as those who did not receive support services but provided enough information to have a status inferred. This rate is approximately 12%, which gives an estimate of an extra 4,000 adults and 97,300 people in total who sought assistance for homelessness in 2013–2014. Assuming this represents 20% of total homelessness, this implies that 480,400 people (s.e. 17,600) were homeless in 2013–2014. This is approximately 27,900 people fewer than the model estimate (508,300 people), or 5.5%. While this difference may reflect a model deficiency, it may also be the result of sampling error ($t = 0.26$; $p = 0.60$) or differences in the measurement of homelessness.

Estimates for the different accommodation types are validated against other data sources. In 2013–2014, 59,400 adults were recorded in the SHSC as having received accommodation support (AIHW 2018) 2.1 times larger than the number predicted by the model to have stayed in homeless shelters and refuges (27,700, s.e. 9,900). A large portion of this difference is likely explained by the broader range of accommodation provided in Australia's homelessness services system. While the GSS focuses on accommodation in shelters and refuges, and only where respondents consider themselves to be without a permanent place to live, the homelessness services system includes longer-term and semi-permanent accommodation, including transitional housing. Model results are also compared against census counts and the Journeys Home survey. This is shown in Figure 5. To provide a reasonable basis for comparison between the annual model estimates and the point-in-time census counts, rates of homelessness from the 2011 and 2016 censuses are multiplied by the inverse of the k ratios to give an annualised census estimate. The Journeys Home estimates are derived by calculating the proportion of homeless episodes that involved each accommodation type across the 2012–2013 and 2013–2014 financial years.

The microsimulation model predicts a substantially larger population stay with family and friends than the two most recent census estimates. Although this may result from model overestimates, it is also likely to reflect a longstanding issue in estimating this form of homelessness in the Australian Census (Chamberlain 1999). The weighted Journeys Home estimates also suggest it is much more prevalent than the census suggests, though not as much as in the GSS model. The microsimulation estimate of street homelessness is somewhat above the census constrained estimates but substantially smaller than the Journeys Home constrained estimates. The model also appears to underestimate the population staying in shelters/refuges and boarding houses/hostels. In saying that, note that (a) Journeys Home represents a highly disadvantaged population, so it will perhaps overestimate more severe forms of homelessness; (b) as with the SHSC, the census shelter count includes all forms of accommodation support, not just shelters and refuges; and (c) unlike in the census and Journeys Home, GSS respondents are asked to self-report staying in different accommodation types specifically because they did not have a permanent place to live. In other words, people may stay in different accommodation forms without necessarily considering it temporary. Nevertheless, these comparisons suggest a high degree of uncertainty worthy of further research.

Figure 5: Comparison of GSS model predictions to census and Journeys Home estimates



Source: Author's calculations from ABS (2012b, 2015, 2018); Wooden et al. (2012); Scutella, Tseng, and Wooden (2017).

8. Conclusion

Homelessness is difficult to measure. Just as with point-in-time estimates, annual predictions are subject to substantial uncertainty. While the results of this study are intended to produce a more accurate estimate of the population who experience homelessness over a year, the indirect approach potentially introduces new sources of error while amplifying existing ones. Standard errors give a sense of the sampling error, however the potential for nonsampling error appears large. Possible sources of error relate to respondent recall, sample design, and consistency in the definition and measurement of homelessness across data sources. Of particular note, implicit assumptions relating to the GSS sampling strategy may have been violated. For example, the model assumes no one currently experiencing homeless has been inadvertently captured in the survey. Filtering the currently homeless out of the survey is probably straightforward for the street homeless though perhaps more difficult for those staying with family or friends or in longer term supported accommodation that is difficult to distinguish from private housing. Failure to exclude these groups will lead to double counting of homelessness in this model. The results also appear sensitive to the estimation of long-term homelessness (given the large value of k for episodes lasting one year or longer), the majority of which is unobserved in the GSS and therefore largely unknown and highly uncertain.

The combining of datasets is a key advantage of this study in imputing information on missing homelessness, validating the results, and generally addressing the limitations and coverage issues inherent in each. However, the accuracy and validity of the results rest, if not on identical operational definitions of homelessness across sources, at least on an assumption that differences in how homelessness is defined and measured do not affect the results. Importantly, the survey, administrative, and census datasets used in this study all construct homelessness from the same types of accommodation. However, differences in measurement no doubt arise given the GSS asks respondents to self-report past experiences with specifically worded questions, where homelessness in Journeys Home and the census is based on the collection, processing, and interpretation of housing and accommodation-related data by a team of researchers, and the SHSC measurement is based on administrative processing of data collected and reported by service providers and intake systems. Indeed, a likely explanation for why estimates differ across sources may derive from the fact that individuals perceive their housing conditions differently from researchers and service providers. As a result, although the microsimulation was designed to minimise reliance on consistent definitions and measurements, the ability to compare and validate the results against other data sources is heavily compromised. Future data collection in this area ought to consider how to better integrate the design, timing, and measurement aspects of point-in-time, household survey, and administrative data instruments.

Limitations notwithstanding, the annual rates estimated in this study are valuable. The comparison to administrative data suggests the magnitudes of these estimates are plausible. Even if not highly precise, they point to certain truths about homelessness that are sometimes suppressed in traditional measures. First, homelessness is more prevalent than typically measured, affecting a larger cross-section of the population. Second, homelessness is more diverse in length and severity with a larger population experiencing temporary and episodic homelessness than point-in-time estimates indicate. Third, various forms of housing exclusion and deprivation, including staying with family/friends and in various forms of marginal and sub-market accommodation are common and likely substantially more so than street and sheltered homelessness. Whether or not these are considered forms of homelessness, it is clear that doubling up with family and friends in particular is highly prevalent among people who consider themselves to be without a permanent residence. Further, the finding that a large proportion of the population who experience literal or hidden homelessness stay with family and friends suggests that experiences of street and sheltered homelessness are commonly accompanied by periods of doubling up. This provides circumstantial evidence of the important links and pathways between doubling up and literal homelessness.

These points offer theoretical and practical insights into the nature of homelessness. Consideration of the greater prevalence and diversity of homelessness gives rise to the hypothesis that social and economic structures expose a larger population to housing instability – if not outright (and literal) homelessness – than generally understood. Indeed, given the sometimes volatile and precarious nature of employment and housing for households at the lower end of the income spectrum (Desmond, Gershenson, and Kiviat 2015; Iceland and Bauman 2007; McKernan and Ratcliffe 2005; Phinney et al. 2007), housing instability may be deeply entwined in the experience and dynamics of poverty and financial hardship. The experience for many individuals and families, however, is relatively transitory, perhaps owing to the personal, interpersonal, and institutional resources they utilise to avoid or escape the most severe and long-lasting consequences of homelessness, including chronic 'rough sleeping' (Piliavin et al. 1993; Wong and Piliavin 1997; Shinn et al. 1998; O'Donnell 2019). In particular, people with interpersonal support networks appear to draw on these supports to a very large extent in managing housing loss. For many, this may be a positive mechanism, providing a stepping stone back into stable housing. For others though, it may be a pathway to deeper housing deprivation and homelessness.

9. Acknowledgements

This paper uses unit record data from the ABS General Social Survey and the Journeys Home: Longitudinal Study of Factors Affecting Housing Stability. Journeys Home was initiated and funded by the Australian Government Department of Social Services

(DSS). The Department of Employment provided information for use in Journeys Home, and the survey was managed by the Melbourne Institute of Applied Economic and Social Research (Melbourne Institute).

The findings reported in this paper are those of the author and should not be attributed to the ABS, DSS, the Department of Employment, or the Melbourne Institute.

References

ABS (2011). *General social survey: Summary results, Australia, 2010*. Cat no. 4159.0. Canberra: Australian Bureau of Statistics.

ABS (2012a). *Information paper – A statistical definition of homelessness*. Cat no. 4922.0. Canberra: Australian Bureau of Statistics.

ABS (2012b). *Census of population and housing: Estimating homelessness, 2011*. Cat no. 2049. Canberra: Australian Bureau of Statistics.

ABS (2015). *General social survey: Summary results, Australia, 2014*. Cat no. 4159.0. Canberra: Australian Bureau of Statistics.

ABS (2016). *Life tables, states, territories and Australia, 2013–2015*. Cat. no. 3302.0.55.001. Canberra: Australian Bureau of Statistics.

ABS (2018). *Census of population and housing: Estimating homelessness, 2016*. Cat no. 2049.0. Canberra: Australian Bureau of Statistics.

AIHW (2018). *Specialist Homelessness Services Collection*. Canberra: Australian Institute of Health and Welfare.

Amore, K., Baker, M., and Howden-Chapman, P. (2011). The ETHOS definition and classification of homelessness: An analysis. *European Journal of Homelessness* 5(2): 19–37.

Bramley, G. and Fitzpatrick, S. (2018). Homelessness in the UK: Who is most at risk? *Housing Studies* 33(1): 96–116. doi:10.1080/02673037.2017.1344957.

Burrows, R. (1997). The social distribution of the experience of homelessness. In: Burrows, R., Pleace, N., and Quilgars, D. (eds.). *Homelessness and social policy*. London: Routledge: 50–68. doi:10.4324/9780203443323_chapter_4.

Burt, M.R. (1992). *Over the edge: The growth of homelessness in the 1980s*. New York, New York: Russell Sage Foundation.

Chamberlain, C. (1999). *Counting the homeless: Implications for policy development*. Cat no. 2041.0. Canberra: Australian Bureau of Statistics.

Chamberlain, C. and Johnson, G. (2013). Pathways into adult homelessness. *Journal of Sociology* 49(1): 60–77. doi:10.1177/1440783311422458.

Chamberlain, C. and Johnson, G. (2015). How many Australians have slept rough? *Australian Journal of Social Issues* 50(4): 439–456. doi:10.1002/j.1839-4655.2015.tb00359.x.

Chamberlain, C. and Mackenzie, D. (1992). Understanding contemporary homelessness: Issues of definition and meaning. *Australian Journal of Social Issues* 27(4): 274–297. doi:10.1002/j.1839-4655.1992.tb00911.x.

Cobb-Clark, D.A., Herault, N., Scutella, R., and Tseng, Y. (2016). A journey home: What drives how long people are homeless? *Journal of Urban Economics* 91(January 2016): 57–72. doi:10.1016/j.jue.2015.11.005.

Coumans, A.M., Cruyff, M., Van der Heijden, P.G.M., Wolf, J., and Schmeets, H. (2017). Estimating homelessness in the Netherlands using a capture-recapture approach. *Social Indicators Research* 130(1): 189–212. doi:10.1007/s11205-015-1171-7.

Culhane, D.P., Dejowski, E.F., Ibanez, J., Needham, E., and Macchia, I. (1994). Public shelter admission rates in Philadelphia and New York City: The implications of turnover for sheltered population counts. *Housing Policy Debate* 5(2): 107–140. doi:10.1080/10511482.1994.9521155.

Culhane, D. and Kuhn, R. (1998). Patterns and determinants of public shelter utilization among homeless adults in New York City and Philadelphia. *Journal of Policy Analysis and Management* 17(1): 23–43. doi:10.1002/(SICI)1520-6688(199824)17:1<23::AID-PAM2>3.0.CO;2-J.

Culhane, D.P., Metraux, S., Park, J.M., Schretzman, M., and Valente, J. (2007). Testing a typology of family homelessness based on patterns of public shelter utilization in four US jurisdictions: Implication for policy and program planning. *Housing Policy Debate* 18(1): 1–28. doi:10.1080/10511482.2007.9521591.

Darcy, L. and Jones, D.L. (1975). The size of the homeless men population of Sydney. *Australian Journal of Social Issues* 10(3): 208–215. doi:10.1002/j.1839-4655.1975.tb00551.x.

Desmond, M., Gershenson, C., and Kiviat, B. (2015). Forced relocation and residential instability among urban renters. *Social Service Review* 89(2): 227–262. doi:10.1086/681091.

D'Onise, K., Wang, Y., and McDermott, R. (2007). The importance of numbers: Using capture-recapture to make the homeless count in Adelaide. *Australian Journal of Primary Health* 13(1): 89–96. doi:10.1071/PY07012.

Firdion, J-M. and Marpsat, M. (2007). A research program on homelessness in France. *Journal of Social Issues* 63(3): 597–588. doi:10.1111/j.1540-4560.2007.00524.x.

Fusaro, V. A., Levy, H.G., and Shaefer, H.L. (2018). Racial and ethnic disparities in the lifetime prevalence of homelessness in the United States. *Demography* 55(6): 2119–2128. doi:10.1007/s13524-018-0717-0.

Greenberg, G.A. and Rosenheck, R.A. (2010). Correlates of past homelessness in the National Epidemiological Survey on Alcohol and Related Conditions. *Administration and Policy in Mental Health and Mental Health Services Research* 37(4): 357–366. doi:10.1007/s10488-009-0243-x.

Hopper, K., Shinn, M., Laska, E., Meisner, M., and Wanderling, J. (2008). Estimating numbers of unsheltered homeless people through plant-capture and postcount survey methods. *American Journal of Public Health* 98(8): 1438–1442. doi:10.2105/AJPH.2005.083600.

Iceland, J. and Bauman, K.J. (2007). Income poverty and material hardship: How strong is the association? *The Journal of Socio-Economics* 36(3): 376–396. doi:10.1016/j.socec.2006.12.003.

Kuhn, R. and Culhane, D.P. (1998). Applying cluster analysis to test a typology of homelessness by pattern of shelter utilization: Results from the analysis of administrative data. *American Journal of Community Psychology* 26(2): 207–232. doi:10.1023/A:1022176402357.

Link, B.G., Phelan, J., Bresnahan, M., Stueve, A., Moore, R.E., and Susser, E. (1995). Lifetime and five-year prevalence of homelessness in the United States: New evidence on an old debate. *American Journal of Orthopsychiatry* 65(3): 347–354. doi:10.1037/h0079653.

Link, B.G., Susser, E., Stueve, A., Phelan, J., Moore, R.E., and Struening, E. (1994). Lifetime and five-year prevalence of homelessness in the United States. *American Journal of Public Health* 84(12): 1907–1912. doi:10.2105/AJPH.84.12.1907.

Lomax, N. and Norman, P. (2016). Estimating population attribute values in a table: 'Get me started in' iterative proportional fitting. *The Professional Geographer* 68(3): 451–461. doi:10.1080/00330124.2015.1099449.

McKernan, S. and Ratcliffe, C. (2005). Events that trigger poverty entries and exits. *Social Science Quarterly* 86(s1): 1146–1169. doi:10.1111/j.0038-4941.2005.00340.x.

Metraux, S. and Culhane, D.P. (1999). Family dynamics, housing and recurring homelessness among women in New York City homeless shelters. *Journal of Family Issues* 20(3): 371–396. doi:10.1177/019251399020003004.

Metraux, S., Culhane, D., Raphael, S., White, M., Pearson, C., Hirsch, E., Ferrell, P., Rice, S., Ritter, B., and Cleghorn, J.S. (2001). Assessing homeless population size through the use of emergency and transitional shelter services in 1998: Results from the analysis of administrative data from nine US jurisdictions. *Public Health Reports* 116(4): 344–352. doi:10.1016/S0033-3549(04)50056-0.

O'Donnell, J. (2019). Does social housing reduce homelessness? A multistate analysis of housing and homelessness pathways. *Housing Studies* doi:10.1080/02673037.2018.1549318.

Phinney, R., Danziger, S., Pollack, H.A., and Seefeldt, K. (2007). Housing instability among current and former welfare recipients. *American Journal of Public Health* 97(5): 832–837. doi:10.2105/AJPH.2005.082677.

Piliavin, I., Sosin, M., Westerfelt, A.H., and Matsueda, R.L. (1993). The duration of homeless careers: An exploratory study. *Social Service Review* 67(4): 576–598. doi:10.1086/604012.

Ringwalt, C.L., Green, J.M., Robertson, M., and McPheeters M. (1998). The prevalence of homelessness among adolescents in the United States. *American Journal of Public Health* 88(9): 1325–1329. doi:10.2105/AJPH.88.9.1325.

Rossi, P.H. (1989). *Down and out in America: The origins of homelessness*. Chicago: University of Chicago Press. doi:10.7208/chicago/9780226162324.001.0001.

Scutella, R., Tseng, Y., and Wooden, M. (2017). Journeys Home: Tracking the most vulnerable. *Longitudinal and Life Course Studies* 8(3): 302–318. doi:10.14301/llcs.v8i2.460.

Shinn, M. (2010). Homelessness, poverty and social exclusion in the United States and Europe. *European Journal of Homelessness* 4(1): 19–44.

Statacorp (2015) *Stata Statistical Software: Release 14*. College Station: Statacorp LLC.

Tompsett, C.J., Toro, P.A., Guzicki, M., Manrique, M., and Zatakia, J. (2006). Homelessness in the United States: Assessing changes in prevalence and public opinion, 1993–2001. *American Journal of Community Psychology* 37(1–2): 29–46. doi:10.1007/s10464-005-9007-2.

Toro, P.A., Tompsett, C.J., Lombardo, S., Philippot, P., Nachtergael, H., Galand, B., Schlienz, N., Stammel, N., Yabar, Y., Blume, M., MacKay, L., and Harvey, K. (2007). Homelessness in Europe and the United States: A comparison of prevalence and public opinion. *Journal of Social Issues* 63(3): 505–524. doi:10.1111/j.1540-4560.2007.00521.x.

Tsai, J. (2017). Lifetime and 1-year prevalence of homelessness in the US population: Results from the National Epidemiologic Survey on Alcohol and Related Conditions-III. *Journal of Public Health* 40(1): 65–74. doi:10.1093/pubmed/fdx034.

Watson, S. (1984). Definitions of homelessness: A feminist perspective. *Critical Social Policy* 4(11): 60–73. doi:10.1177/026101838400401106.

Wong, Y.I. and Piliavin, I. (1997). A dynamic analysis of homeless-domicile transitions. *Social Problems* 44(3): 408–423. doi:10.1525/sp.1997.44.3.03x0123s.

Wooden, M., Bevitt, A., Chigavazira, A., Greer, N., Johnson, G., Killackey, E., Moschion, J., Scutella, R., Tseng, Y., and Watson, N. (2012). Introducing 'Journeys Home'. *The Australian Economic Review* 45(3): 368–378. doi:10.1111/j.1467-8462.2012.00690.x.

Appendix: Estimating homelessness in the General Social Survey

A set of logistic regression models is run on the Australian GSS data to estimate the number of people who experienced homelessness in the previous 12 months by the duration of their episode, the forms of accommodation (for example, street and shelter) they experienced, and whether they sought help. The models are simultaneously estimated on the GSS data using Generalised Structural Equation Models in Stata 14 (Statacorp 2015). There are five dependent variables: (a) $when$: whether homelessness was experienced homelessness in the previous 12 months; (b) $length$: the duration of respondents' most recent episode among those who had been homeless in the previous 10 years; (c) $episodes$: the number of episodes they have ever been without a permanent place to live; (d) $accom$: the forms of accommodation they have experienced when without a place to live; and (e) $help$: whether they sought help from a housing or homelessness service provider during their most recent episode.

The dependent variables are regressed on the age and sex of the respondent. Age is modelled using a cubic spline with knots at 25, 35, 45, and 55 years. Knots were selected by testing different combinations and comparing their fit and p-values. The regression equations take the form

4. $$\log\left(\frac{p(when_i)}{p(when_0)}\right) = \beta_{0f} + \beta_1.age + \beta_2.age^2 + \beta_3.age^3 + \beta_4.\max(0, age - a_k)^3 + \beta_5 female$$

5. $$\log\left(\frac{p(length_d)}{p(length_8)}\right) = \beta_{0f} + \beta_1.age + \beta_2.age^2 + \beta_3.age^3 + \beta_4.\max(0, age - a_k)^3 + \beta_5 female + \beta_6 when_i$$

6. $$\log\left(\frac{p(episodes^e)}{p(episodes^1)}\right) = \beta_0 + \beta_1 age + \beta_2 age^2 + \beta_3 age^3 + \beta_4.\max(0, age - a_k)^3 + \beta_5 female + \beta_6 when_i + \beta_7 length_d$$

7. $$\log\left(\frac{p(accom^s)}{1-p(accom^s)}\right) = \beta_0 + \beta_1 age + \beta_2 age^2 + \beta_3 age^3 + \beta_4.\max(0, age - a_k)^3 + \beta_1 when_i + \beta_2 length_d + \beta_3 episodes_e$$

8. $$\log\left(\frac{p_{help}}{1-p_{help}}\right) = \beta_0 + \beta_1.age + \beta_2.age^2 + \beta_3.age^3 + \beta_4.\max(0, age - a_i)^3 + \beta_5.female + \beta_6.age.female + \beta_7.when_i + \beta_8.length_d,$$

where age is the single year age of the respondent at the time of the survey, a_k are the four age knots (20, 30, 40, and 50 years), and $female$ is a dummy variable indicating the sex of the respondent. Equation 4 is run on the full GSS sample and the subsequent equations on the subsample who had been homeless in the previous 10 years. Equations 4, 5, and 6 are run with multinomial logistic regression, Equation 8 with binary logistic regression, and Equation 7 with a set of binary logistic regressions indicating whether respondents had ever stayed in each of the five accommodation types (e.g., street, shelter).

Respondent ages are adjusted in Equations 5 through 8 to reflect respondents' approximate age at the time of their most recent homelessness episode – that is, their age at the time of the survey minus six months for episodes in the last 12 months, three years for episodes 1 to 5 years ago, and 7.5 years ago for episodes 5 to 10 years ago. Interactions are included between each of the age variables and $female$ to produce separate splines for males and females. Interactions between $when$ and the age and sex variables are also tested in the second equation, though dropped if their p-values were greater than 0.15.

Predicted probabilities are multiplied by the GSS in-scope population by age and sex to produce homelessness population estimates. Estimates are constrained to population totals by Iterative Proportional Fitting (IPF). For example, homelessness estimates by age, x, and sex, f, in the previous 12 months (h_{af}^{12}) are first constrained to the total homeless population in the last 12 months ($h^{12} = 351{,}000$):

9. $$h_{x,f}^{12} = \frac{h_{x,f}^{12}}{\sum_x \sum_f h_{x,f}^{12}} \times h^{12}.$$

Estimates by age, sex, and length, d, of homelessness in the last 12 months, h_{xf}^{12d}, are then iteratively constrained to h_{xf}^{12} and the total homeless population in the last 12 months by length of episode, h^{12d}:

10. $$h_{x,f}^{12,d} = \frac{h_{x,f}^{12d}}{\sum_d h_{x,f}^{12d}} \times h_{x,f}^{12}, \text{ and}$$
11. $$h_{x,f}^{12,d} = \frac{h_{xf}^{12d}}{\sum_x \sum_f h_{x,f}^{12d}} \times h^{12,d}.$$

These calculations are repeated until h_{xf}^{12d} converge on a stable set of values. For more information on IPF, see Lomax and Norman (2016). Estimates of h_{xf}^{12d} are used to weight simulated episodes so that their sum is representative of homelessness nationally. A small number of those aged 15 years who were predicted to have experienced homelessness as 14-year-olds are dropped from the model to focus on the adult homeless population (defined in the GSS and in this paper as 15 years and over). The model is rerun using each of the 60 Jackknife replicate weights provided in the

GSS. This produces an additional 60 set of results from which standard errors can be estimated.

The approach to estimating different forms of homelessness starts by predicting the number of episodes without a permanent place to live that respondents have ever experienced (Equation 6) and all the accommodation types they stayed in during these episodes (Equation 7). The next step is to approximate probabilities of having stayed in each accommodation type in the most recent episode. To estimate probabilities for the population with two or more lifetime episodes, stays in $accom^s$ are simulated for each episode. The simulation is performed by using the probability of ever staying in $accom^s$ among those with one lifetime episode ($p\left(accom_{x,f}^{d,e=1,s}\right)$) – given the individual's age, sex, and duration of most recent episode – as an initial proxy for the probability of staying in $accom^s$ in each episode. A random number, R, is drawn for each individual with two or more lifetime episodes. If R is less than or equal to $p\left(accom_{x,f}^{d,e=1,s}\right)$, then the individual is assigned to $accom^s$ in their first episode – noting that individuals can be assigned to multiple accommodation types for each episode. Whether individual i stays in $accom^s$ in their second episode is simulated with the same probability:

$$12.\quad accom_i^{e=2,d,s} = \begin{cases} 1, & R \leq p\left(accom_{x,f}^{family} \mid episodes = 1, length = d\right) \\ 0, & otherwise \end{cases}.$$

From these simulated episodes, the number of individuals with two past episodes who stayed in $accom^s$ in their second and most recent episode can be calculated. This population is then weighted so that the number of individuals who stay in $accom^s$ in either or both of their two episodes is equal to the predicted population, $h_{x.f}^{12,d,e=2,s}$, calculated from Equation 7. The procedure is repeated for the population with three past episodes by simulating a third episode and so on, up to a maximum of five episodes. The weighted population who experienced different forms of homelessness in their most recent episode by age, sex, and length of homelessness is then converted into a probability and applied to the main part of the microsimulation.