Mobility and sex work: why, where, when? A typology of female-sex-worker mobility in Zimbabwe

Calum Davey[a,*], Jeffrey Dirawo[b], Phillis Mushati[b], Sitholubuhle Magutshwa[b], James R Hargreaves[a], Frances M Cowan[b,c]

[a] *London School of Hygiene and Tropical Medicine, 15 Tavistock Place, London, WC1H 9SH*
[b] *Centre for Sexual Health and HIV/AIDS Research, 9 Monmouth Rd, Avondale West, Harare*
[c] *Liverpool School of Tropical Medicine, Pembroke Place, Liverpool*

Acknowledgements

CD was funded by a Strategic Skills pre-doctoral Fellowship from the Medical Research Council, UK. JD, PM, SM, JH, and FC received no funding for this work. None of the authors have conflicts of interest to declare. We are very grateful to all of the women who gave their time to be interviewed and for their assistance in recruiting their peers.

*Corresponding Author
Email address: calum.davey@lshtm.ac.uk (Calum Davey)

Abstract

Sex worker mobility may have implications for health and access to care but has not been described in sub-Saharan Africa. We described sex-worker mobility in Zimbabwe and a mobility typology using data from 2,591 and 2,839 female sex workers in 14 sites from 2013 and 2016. We used latent class analysis to identify a typology of mobile sex workers. More women travelled for work in 2016 (59%) than in 2013 (27%), usually to find clients with more money (57% of the journeys), spending a median of 21 (2013) and 24 (2016) days away. A five-class mixture model best fit the data, with 39% in an infrequent work-mobility class, 21% in a domestic-high-mobility class, 16% in an international-high-mobility class, 16% in an infrequent opportunistic-non-work-mobility class, and 7% who travel with clients. More-mobile classes were better educated; risk behaviours differed by class. Mobility is increasing among sex worker in Zimbabwe, multi-faceted, and not explained by other vulnerabilities.

Keywords: Sex work; migration; mobility; healthcare; HIV

Background

Sex work takes different forms, but all involve the exchange of money or in-kind goods or services for sex (Raluca Buzdugan, Shiva S Halli, et al., 2009; Harcourt and Donovan, 2005). Sex workers often experience stigma (Scambler and Paoli, 2008), police harassment, arrest, client violence (Shannon and Csete, 2010), and worldwide are more likely to be infected with HIV than women of the same age (Baral et al., 2012). Public-health programmes that meet the needs of sex workers are essential.

Sex work 'typologies' have been useful for understanding variation among sex workers to improve the design of health programmes (Jain and Saggurti, 2012). Various typologies have been used, for example: place of solicitation (Sinha, 1997), place of sex (Raluca Buzdugan, Copas, et al., 2009), sex-work income (Hong et al., 2012), whether the primary purpose of the initial interaction is for exchanging sex (Harcourt and Donovan, 2005), and age (Delany-Moretlwe et al., 2015). Most of the research on sex work typologies has come from outside of sub-Saharan Africa.

Anecdotally, sex workers live highly-mobile lives, and there is evidence of high-levels of mobility in the literature (Goldenberg et al., 2014; Reed et al., 2012). Mobility has many dimensions: people may move as individuals or with a partner or family; over various geographic distances and may cross social or civic boundaries; with different motivations for moving; to destinations that differ in various ways; more or less often (frequency); for shorter or longer stays (duration); where they can use healthcare; and seasonally or periodically throughout the year or over a lifetime (Brown and Bell, 2004; Taylor et al., 2011). Mobile populations vary in the impact they have on the sending and receiving populations. A typology of sex-worker mobility that addresses the multiple-dimensional variation in mobility has not been developed.

There is limited quantitative literature on mobility of female sex workers in sub-Saharan Africa. Brothel-based sex work is less common in sub-Saharan Africa than in Asia, and therefore the extent of and reasons for moving may be different. The little evidence available suggests that mobility can be high and varies by context: surveys of female sex workers conducted in 1997-98 in four sub-Saharan cities found that the average proportion of time spent in the city in the past year ranged from 69% in Cotonou (N=433) to 95% in Ndola (N=332) (Morison et al., 2001). In the same years in Ghana, 17% of female sex workers surveyed (N=1,013) had ever worked outside of the country (Asamoah-Adu et al., 2001). In Kenya, 403 women working on a highway spent 25% of 28 nights away from home 'base', and fewer than 20% of women spent all 28 days at the same place (Ferguson and Morris, 2007).

Although often associated with higher risks of HIV (see for example Reed et al. (2011), Reed et al. (2012), and Halli et al. (2010)) and other health concerns (for example, depression in Patel et al. (2016)), mobility can be thought of as a form of capital (Hall, 2005), broadly defined as a 'means to combine goals in space' (Hooimeijer and Van der Knaap, 1994). Short-term mobility that does not involve a semi-permanent change of address (although it usually involves at least one overnight stay (Smith, 1989)) can be thought of as a complement of migration (Bell and Ward, 2000), that can offer a lower-cost substitute for permanent migration in some circumstances (Pollard, 1996). Immobility, or stability (Hanson, 2005), may itself be a problem for women who need to find clients with money. Exploring the mobility of sex workers places emphasis more 'on work than on sex' (Vanwesenbeeck, 2001), conceptualising sex work first in terms of labour, and second in terms of sexual risk: a focus that has been called for by sex-work organisations (RedTraSex, 2016).

We aimed to address the paucity of literature on mobility, especially in sub-Saharan Africa, by developing a typology of mobile sex workers. Using data from Zimbabwe collected in fourteen sites in 2013 and 2016, we explored the extent of sex worker mobility across a number of dimensions. We hypothesised that types of mobility would be characterisable from the data, and that mobility would be associated with sociodemographic characteristics, for example that younger women, aged less than 25, and older women aged more than 40 would be less likely to move for work because of weaker social networks and less inclination to travel, respectively. Viewed as a form of capital, we hypothesised that mobility would be associated with lower food insecurity, higher income per sex act, and more clients. With access to a larger client-base, we hypothesised that mobility would strengthen bargaining and be associated with higher reported condom-use with clients. We geolocated the names of places that sex workers reported visiting within the last year and described the the journeys, the sex workers' mobility, and the mobility from the places that the samples were drawn from. The results of this analysis may inform the design of services in the region, and mathematical models.

Methods

Setting and data collection

The data were collected as for the baseline and endline of the SAPPH-IRe trial, a cluster-randomised controlled trial of a complex intervention to reduce the proportion of the sex-worker population with an infectious level of HIV using a combination of: PrEP provision, immediate on-site ART initiation, and various forms of adherence support (Hargreaves et al., 2016). Details of the survey procedures have been published (Cowan et al., 2017); in short: for each of fourteen sites across the country, community mapping was conducted and between six and eight female sex workers purposively sampled as representatives of sub-networks identified at each site. After consenting-to and completing a face-to-face interview, and providing a blood spot, each woman was given two coupons to invite her peers to participate in the survey. Women who presented with coupons were asked for consent, interviewed, and asked for a blood spot, and given coupons two recruit two further female sex workers. This process of respondent-driven sampling (RDS, Heckathorn (1997)) continued until approximately 200 women were recruited in each site. Participants were compensated 5 USD for loss of earnings due to the interview, and a further 2 USD for each peer recruited. Women were eligible to participate if they had exchanged money for sex in the last 30 days, were at least 18 years old, and had lived in the site for the past six months.

Measures

Demographics and sex-work characteristics

The questionnaire included questions on: age, marital status (married, separated, divorced, widowed, never married), and highest education level completed (re-coded as 'no education', 'primary only', 'secondary or above'). The measures of food security differed between the two years, in 2013 women were coded as 'food insecure' if they reported any of: no to 'We can eat at least 2 meals a day', yes to 'sometimes we go to bed hungry', or yes to 'in the last week, have you had to go an entire day without eating because there was no food in your household?'; in 2016 women who reported yes to 'in the past four weeks, was there ever no food to eat of any kind in your house because of lack of resources to get food?' were coded as food insecure (drawn from Swindale and Bilinsky (2006)). Women were asked how old they were when they started sex work, how many clients they had in the last week (and whether this was more or less than average), and whether they consistently use condoms with clients (i.e. answered always to 'in the past month how often did you use condoms with your clients?' and no to 'think again about all the clients you had in the last month, have there been any times when you did not use condoms?').

The data were collected using RDS, therefore women had a non-random chance of being included in the survey. This could lead to bias, and the RDS-2 methodology inversely weights each individual according to their visibility in the network (Volz and Heckathorn, 2008). Women were asked how many other eligible women they knew, which was used to estimate their 'degree' in the network, i.e. the extent that each woman is connected to others. Often the number of eligible women each woman knows is used alone to weight the data (Malekinejad et al., 2008), however this can be noisy, clumped around commonly reported values (i.e. 10, 50), and contain outliers. Therefore, we used a method developed by McLaughlin et al. (2015) to impute 'visibility' of each woman in the network from the number of eligible women that each woman knows, the time of the interview relative to the start of survey, and position in the chain of referrals. The scores were normalised to allow comparisons across sites. The purposively-selected women in each site were excluded from the analysis. Weighted proportions, weighted means, and weighted medians were calculated using the RDS-2 method, weighted by the inverse of the imputed visibility score.

Mobility

Women were asked if they had worked in sex work anywhere other than the interview site in the past 12 months, using the same question in both surveys. They were then asked to recall the places that they had been in reverse-chronological order. Women were asked to name up-to five places were they had worked in sex work and the duration of stay at each place. In 2016, women were also asked for the names and duration of stay in up-to five places where they did not work in sex work; they were also asked for the total number of places they visited, the reasons for visiting each place reported, and whether they used healthcare services while visiting. Duration was recorded as an ordered categorical variable (e.g. 'less than a week', '1-2 weeks', etc.), and was re-coded as a continuous variable in days using the mid-point of each category so that it could be summarised across visits for each woman (e.g. the median duration); in 2013 two periods were missed from the questionnaire (see **Appendix 1**). The typical amount charged per sex act, which was also recorded as categorical, was re-coded as a continuous variable in dollars (e.g. 'up to $2' became $1, '$2-5' as $3.5, etc.). The data collection tools were piloted with sex-worker peer educators.

Places were recorded as reported by the participant with the name of the village, town or city, province, and country. These were written in the questionnaires by the interviewer, resulting in many spellings and misspellings. A full list of the places, as named, was generated for each site and researchers based in Harare identified the latitude and longitude coordinates of the places ('geocoded') by searching the Internet and using Google Maps (Google, 2018). When the name of a country was given without any town or city, the closest point on the Zimbabwean border from the site of interview was used.

Distances between places were estimated using the Google Maps API with the function *mapdist* in R package *ggmap* (Kahle and Wickham, 2013). The distances were for travel along a road, reported in kilometres and the hours it would have taken to travel by vehicle. Data on the province and country were downloaded from gadm.org (http://www.gadm.org/download) and spatially joined with the data using the *sp* package (Pebesma and Bivand, 2005). The code is available in **Appendix 5**.

Statistical analysis

Describing mobility. The journeys from the site of interview to the place visited were described in terms of the distance in kilometres, hours travelled, and the length of stay in days. The proportion of journeys between provinces, international, and between majority Shona-speaking eastern provinces and Ndbele-speaking western provinces was calculated. Places visited were categorised as 'town or city', 'growth point (official focal points for decentralisation of services into rural areas) or business centre', 'mine or farm', or 'other' – other types of place included villages, resorts, and highways. Summaries were reported using the median and inter-quartile range.

To estimate the number of unique places visited per woman, the latitude and longitude coordinates were rounded to one decimal place, which corresponds to approximately $10km^2$. Women were not asked if they returned to the site of interview between visits, therefore to summarise the distance travelled by each woman we calculated all of the possible journeys between the destinations and the site of interview (see **Figure 1**), calculated the median of the minimum and maximum distances that each woman could have travelled, and the median of 1000 medians from random draws of one journey per woman from all the possible journeys. We also presented the total distance when assuming that women returned to the site of interview between visits. Of the various features of mobility, we described: distance, frequency, duration, total time spent away, motivation (i.e. reason for moving), and whether or not reported using healthcare during the visit. We reported RDS-2 weighted medians and proportions.

We summarised the proportion of female sex workers who travelled and how long they spent away at the for each interview site. We estimated the median overall days spent away as a weighted average of the median time away for women who moved and the zero time away for those who did not move. We were unable to explore the impact on the receiving communities, for example how the number of women visiting a place compared to the local number of sex workers.

Sex work typologies. Prior to our analysis we hypothesised a typology of mobility for sex work. We hypothesised that mobile women would exhibit mobility behaviour consistent with three types:

(A) women who work in truck stops and travel with clients, with longer distances travelled, including internationally, not staying for long periods, and reporting 'to get more clients' or 'travel with clients' as the reasons.

(B) women who move seasonally or because of special events to specific places, with domestic travel to one or two places, staying for short periods, and reporting to 'find clients with more money' as the only reason.

(C) women who move frequently over short distances to find new markets, with multiple journeys over short distances, staying for moderate lengths of time, and giving reasons 'to find more clients' or 'to make more money by being new'.

After listing the most common sets of reasons for moving, and inspecting correlations between different features of mobility, we used a latent class analysis (LCA) approach to classifying types of women using a mixture model. LCA is a form of unsupervised machine learning (Masyn, 2013). We used the *depmixS4* package, which allows for both binary and continuous manifest variables (Visser and Speekenbrink, 2010). We included the seven reasons that were reported for at least 5% of the journeys (the rest together as 'other'), and coded women by whether they ever gave each reason for any of the journeys reported. We included whether or not any travel was international, whether she ever travelled to a growth-point, or a mine, the median distance travelled, the median time away, and whether she lives in a town (as opposed to at a truck stop/colliery). We did not include the number of different places that women travelled to – this was very similar to the number of visits and there may have been under-reporting of repeat visits. The median-journey-time was log-transformed to account for skewed distributions; the median time away could not be transformed into a standard distribution and therefore was dichotomised into less than or more than two-weeks; the number of visits was modelled as Poisson (count); the other variables were modelled as binomial distributions (i.e. with a logit link function). To determine the number of classes, we ran the model for between two and 10 classes and plotted the Bayesian Information Criterion (BIC or Schwartz criterion) (Nylund et al., 2007) to identify a minimum using the 'scree-plot' method: better models were models at the minimum of the BIC curve and with the fewest possible classes. We ran each model 100 times with random starting values and plotted the 95% range to account for variation in the BIC between model runs, and to help determine whether we were finding local or global minima. We ran a series of diagnostics to evaluate the fit of models of classes around the minimum of the BIC plot and used Bootstrapping to compare the candidate models, described in **Appendix 3**.

To report the variability of mobility across sites, we plotted the density of the locations visited from each site in a 17x15 grid of 50x50km squares, organised by matched-pair for the trial and labelled with the 'type' of place (e.g. city, colliery, truck stop) and the proportion of the women who moved at each site who were members of each class identified in the LCA.

Associations between mobility and socioeconomic and behavioural risk factors. We explored associations between mobility for sex work and sociodemographic characteristics by describing women in each of the identified types of mobility, including those who had not worked elsewhere. We used descriptive analysis, calculating the RDS-2 adjusted proportions and medians.

Ethics

[REDACTED TO ANONYMISE AUTHORS]

Results

Demographics, sex work, mobility data

Demographic and sex-worker characteristics in both survey-rounds are shown in **Table I**. 2,591 women were recruited through RDS in 2013, and 2,839 in 2016, with similar age profiles. In both years, around a third had no education, primary education, and secondary education. Few women were married, the majority were divorced (1,652; 63% in 2013 and 1,823; 64% in 2016). Participants supported on average 1 child, and 45% and 42% were food insecure in 2013 and 2016, respectively. The median age that started sex work was 23 in 2013 and 24 in 2016. Women had median 5 clients in the past week, who were primarily solicited in bars, and 55% and 45% relied solely on selling sex for income in 2013 and 2016, respectively. In 2013, 65% used condoms consistently with clients, 58% in 2016. HIV prevalence was 60% in 2013 and 59% in 2016.

A higher proportion of female sex workers reported working elsewhere in the past 12 months in 2016 (59%) than in 2013 (27%). A similar proportion stayed elsewhere without working in 2016 (55%), such that 2,294 women (81%) reported any mobility in 2016. Of the mobile women, 681 (97%) in 2013 and 2,293 (99.96%) in 2016 reported the name of at least one place. The gazette of places worked or stayed contained 1,427 differently-spelled place-names, for 471 places. The coordinates were identified for 6,541 (99%) of the places named; there were 929 places geo-coded in 2013, and 5,612 in 2016. In the 2013 survey, 27 women were missing place names, and 5 in 2016.

Features of mobility

The journeys are described in the upper panel of **Table II**. There were 929 journeys in 2013, and 3,364 in 2016 that included sex work and 2,248 that did not include sex work. The median distance was higher for non-work mobility, although women stayed for less time: median (inter-quartile range) of 4 (4-21) days as compared to 18 (4-21) when working. Most journeys were within Zimbabwe, only between 10% and 17% were to another country; even fewer (less than 10%) travelled between linguistic areas within

Zimbabwe. For places where women worked in sex work, 164 different places were visited a median (IQR) 2 (1, 6) times each in 2013, 270 places a median 3 (1, 11) times each in 2016, and 262 places were visited without working in 2016, for 2 (1, 7) times each.

In 2016, women reported up to six reasons for moving for each journey reported. The most common reason to travel to places and work was to find clients with more money (57%), followed by to find more clients (39%). For moves without work, the most common reason was to visit family (72%). The results of listing the most common combinations of reasons given for travelling, and the correlations between the different features of mobility are reported in **Appendix 2**. In short, for travel that included sex work there were many combinations of reasons given for each journey, and the most common accounted for a small proportion of journeys. For moves without sex work, the most common combination was simply 'to visit family' and accounted for more than half of the total journeys. There were few strong correlations between features of mobility (**Appendix Figure A2**).

Only 12% and 8% of visits included the use of healthcare (**Table II**).

Women who made at least one journey are described in the middle panel of **Table II**. The mean (IQR) visits per person to places where they worked increased from 1 (1, 1) to 2 (1, 3) between 2013 and 2016. In 2013 and 2016, women were away 21 and 24 days for work, and 14 without working in 2016.

The median number of days in the past 12 months spent in the interview location in 2013 was 359, and 342 in 2016. 17 (1%) women reported spending more than the previous 12 months away working in 2016 (none in 2013), and 2 (0.1%) when not working. The median (IQR) total distances in kilometres to places from the site of interview (and back again) was 356 (180, 662), 464 (229, 967), and 443 (208, 850) for 2013, 2016 with and without working, respectively. The median figures were similar for the mean of the random draws from the possible routes, but with narrower distributions.

Mobility described at the level of the sites of interview and the places visited is shown in the bottom panel of **Table II**. There was not considerable variation between sites, although the variability in the duration away was higher for moves that involved sex work than for non-sex-work mobility.

Sex work-mobility types

The LCA mixture model converged for all 100 runs for each of 2 to 10 classes; the BIC plot is shown in **Appendix Figure A3.0**. After five classes there was a bi-modal distribution of the BIC for each number of classes. For subsequent analyses we generated model fits drawn from the part of the BIC distribution with the lower mode. Comparisons between the models are described in **Appendix 3**. The five-class model was preferred; the likelihood-ratio test found that the five-class model was superior to the four-class (LRT p-value=0.01) but did not show an advantage of choosing more complex models (LRT p-value=0.29). The probability profiles of the manifest variables are shown in **Figure 2**, with the probability of the class members exhibiting the binary variables is shown in grey-scale (black being 100%), and the journey times and number of visits are described with the means in each class. We had hypothesised that there would be three types of mobile sex worker and our analysis identified five.

We expected to find a group of mobile sex workers who moved with clients but did not stay for long period, and we found that this group was small (class 1, 7%). The second expected group was those who move domestically moderately frequently to find clients, which was similar to the largest class, class 5, with 39% of mobile women. The third was women who moved frequently over short distances staying for moderate lengths of time, closely resembling class 4 with 21% of the mobile women. These women were very likely to have visited a mine or farm for work. The two classes that we did not anticipate, therefore, were classes 2 and 3. Class 3 were women who reported working at places they had visited for reasons other than for work, and represented 16% of mobile women in the sample; class 2 were highly mobile and often travelled internationally over long distances and represented 16% of mobile female sex workers.

The five-class model was used to predict class membership, the demographics of the women in the classes are described in **Table III**, and by site in with density plots in **Figure 3**. Few destinations from any one site accounted for more than 20% of the journeys (purple on the figures). There was a strong correspondence between site and class, with 100% of the mobile women in a single class in 10 of the 14 sites, although no class was only present in one site.

Women in the internationally mobile classes were more highly educated, more likely to have been divorced, and less likely to have been widowed. There were only small differences overall in the median age of starting sex work, numbers of child dependents, and number of clients. Women who worked while traveling for other reasons (class 3) where more likely to be food insecure. Women who moved frequently over shorter distances (class 4) or with clients (class 1) were less consistent with condom use. Women who travelled frequently domestically (class 4) were less likely to report that sex work was their only source of income. Non-mobile women had been living longer at the site. HIV prevalence, visibility in the network, and the proportion who had stayed elsewhere in the last 12 months without working did not vary substantially between the classes.

Discussion

We have described the mobility of female sex workers in Zimbabwe, and identified five kinds of mobile sex worker. On average, women spent less than 10% of their time away from the interview site, however many travelled long distances, and stayed away for weeks at a time. Women of all ages and demographics reported moving for work. Contrary to our hypotheses, we did not find more mobile women had lower food insecurity, and that consistent condom use varied by the kind of mobility. We found that sex-worker mobility in Zimbabwe as increased dramatically between 2013 and 2016.

Our analysis faced a number of limitations. The first is that the inclusion criteria required that women had lived at the site for at least six months, potentially excluding some mobile women. The surveys took place over a short period and may have missed any women who were travelling during that time. It is likely that these issues led to an underestimate of the extent of mobility, and underestimated the proportions of the more mobile types of sex workers. The second is that the data were cross sectional, and we relied on recall over long periods to collect detailed data on mobility. Data on mobility from a cohort followed over time would improve the measures, however recruitment for cohort studies and follow-up of sex workers can be challenging (Ward and Day, 2006) which could reduce the representativeness of the study.

The analysis in this paper has a number of strengths. The places that women visited were geocoded directly, rather than, for example, using the centroids of provinces or other administrative areas. We took advantage of free and accessible software to calculate the distances in terms of travel by car along roads: in Zimbabwe, where intercity roads are scarce, the Euclidean distance may underestimate the journey distance. We investigated multiple dimensions of mobility, describing the features of mobility in detail at journey, women, and place levels before categorising participants using LCA. Distinguishing between moves for sex work and moves for other reasons resonates with the distinction between mobility for production, e.g. for work, and mobility for consumption, e.g. for leisure (Bell and Ward, 2000). However, many women engaged in sex work when they moved for other reasons, possibly reflecting the opportunistic and survival imperative behind much of sex work.

Measuring mobility has an inherent temporal problem: over what period are mobility events measured? Twelve months is relatively long, potentially affecting recall of visits which would underestimate mobility, while shorter periods might miss women who do not travel often. Richardson and Seethaler (2001) have suggested using just the one last trip, whenever it happened; however, had we used this we would not have been able to explore how mobility varied at the individual level. Our analysis was quantitative only, although we consulted with peer educators when developing the mobility tool for 2016. Qualitative data may have been informative for developing the typologies of sex worker by mobility. Finally, we did not have data on the unit of movement (i.e. whether women travelled with partners or children), or dates of travel that could have been used to explore seasonality and periodicity. We were not able to investigate the impact of the women visiting on the sex worker populations (Brown and Bell, 2004).

Non-sex workers in Zimbabwe are also increasingly mobile (Agency and International, 2016) (see **Appendix 4** for a comparison between the *de jure* and *de facto* household members in the 2010 and 2015 DHS surveys). Around 10% of people in the 2015 DHS survey were staying somewhere other than their usual residence on the day that the survey data was collected from their household. This is higher than the corresponding figure for sex workers in 2016: the median total time away in our survey of sex workers in 2016 was 28 and therefore we would expect to find $(28/365) \cdot 100\% = 7.7\%$ women away from home if they were surveyed at home as for the DHS. However, since the RDS surveys were conducted over two weeks at each site, it is possible that mobile women would not be available to take part, as the DHS results emphasise. Women reported knowing a median of 9 others in 2016 who they could potentially recruit, however only 60% recruited two others for the survey in waves 0-4; it is likely that many factors would contribute to this, including mobility.

Our analysis identified five different types of mobility based on multiple features. Previous descriptions of sex worker mobility have focused on frequency (Reed et al., 2012), reasons for moving (especially whether physically forced, or tricked, or with own agency (Agustín, 2005; Steen et al., 2015)), type of place visited (Halli et al., 2010), or crossing national borders (Richter et al., 2014). Our analysis shows that these features overlap when defining sex worker mobility. For example, 21% of the mobile women were domestically mobile and travelled about two-and-a-half times per year over short distances, 16% rarely travelled explicitly because of sex work but travelled longer distances and worked while visiting nonetheless, and just 7% travelled with a client and stayed for less than two weeks.

Literature on mobility in low income settings has described ‘circulatory’ migration or mobility often as moves back and forth between two locations, or at least between places to an original ‘home’ (Chapman and Prothero, 1983; Zelinsky, 1971), however whether all mobile people have a specific location called ‘home’ has been questioned (Behr and Gober, 1982). In our analysis we relaxed the assumption that the women returned to one fixed place between visits, and although this did not have an enormous effect on the median distances covered, it did reduce long-distance outliers, which may be the journey patterns that were the most unlikely (because women might prefer more efficient routes connecting the visited places). In future research on mobility among sex workers, women should be asked about how they travelled from one place to another and whether it included a return 'home', perhaps by completing a timeline of locations and time spent in each place.

We did not find that many women reported moving because of harassment or breakdown in relationships. Most women in Zimbabwe work independent of pimps or other intermediaries (Wilson et al., 1989). However, given the high rate of food insecurity and that most of the women relied on sex work as their only source of income, mobility should be interpreted within a conceptual model that acknowledges that personal agency is constrained and influenced by, among other things, extreme poverty (Hagen-Zanker, 2008). More research, including qualitative research, is needed to investigate the health and well being impacts, both negative and positive, of mobility and stability among sex workers.

Further research is needed to understand more about why women move within a structural framework. Although we did not find that typologies of sex work were associated with key indicators of behavioural risk, future work should explore the implications of mobility for access to healthcare, and for adherence to treatment regimens such as for antiretroviral therapy.

Word count. 5,318

References

Agency, Z.N.S., International, I., 2016. Zimbabwe demographic and health survey 2015: Final report. Zimbabwe National Statistics Agency (ZIMSTAT); ICF International, Rockville, Maryland, USA.

Agustín, L.M., 2005. Migrants in the mistress's house: Other voices in the“ trafficking” debate. Social Politics: International Studies in Gender, State and Society 12, 96–117.

Asamoah-Adu, C., Khonde, N., Avorkliah, M., Bekoe, V., Alary, M., Mondor, M., Frost, E., Deceuninck, G., Asamoah-Adu, A., Pépin, J., 2001. HIV infection among sex workers in Accra: need to target new recruits entering the trade. JAIDS Journal of Acquired Immune Deficiency Syndromes 28, 358–366.

Baral, S., Beyrer, C., Muessig, K., Poteat, T., Wirtz, A.L., Decker, M.R., Sherman, S.G., Kerrigan, D., 2012. Burden of HIV among female sex workers in low-income and middle-income countries: A systematic review and meta-analysis. The Lancet Infectious Diseases 12, 538–549.

Behr, M., Gober, P., 1982. When a residence is not a house: Examining residence-based migration definitions. The Professional Geographer 34, 178–184.

Bell, M., Ward, G., 2000. Comparing temporary mobility with permanent migration. Tourism Geographies 2, 87–107.

Brown, D., Bell, M., 2004. Measuring temporary mobility: Dimensions and issues. International Union for the Scientific Study of Population Twenty-Fifth Conference.

Buzdugan, R., Copas, A., Moses, S., Blanchard, J., Isac, S., Ramesh, B.M., Washington, R., Halli, S.S., Cowan, F.M., 2009. Devising a female sex work typology using data from Karnataka, India. International journal of epidemiology 39, 439–448.

Buzdugan, R., Halli, S.S., Cowan, F.M., 2009. The female sex work typology in India in the context of HIV/AIDS. Tropical Medicine & International Health 14, 673–687.

Chapman, M., Prothero, R.M., 1983. Themes on circulation in the third world. International Migration Review 597–632.

Cowan, F.M., Davey, C.B., Fearon, E., Mushati, P., Dirawo, J., Cambiano, V., Mavedzenge, S.N., Hanisch, D., Wong-Gruenwald, R., Chemhuru, M., others, 2017. The HIV care cascade among female sex workers in Zimbabwe: results of a population-based survey from the Sisters Antiretroviral therapy Programme for Prevention of HIV, an Integrated Response (SAPPH-IRe) trial. JAIDS Journal of Acquired Immune Deficiency Syndromes 74, 375–382.

Delany-Moretlwe, S., Cowan, F.M., Busza, J., Bolton-Moore, C., Kelley, K., Fairlie, L., 2015. Providing comprehensive health services for young key populations: Needs, barriers and gaps. Journal of the International AIDS Society 18.

Ferguson, A.G., Morris, C.N., 2007. Mapping transactional sex on the Northern Corridor highway in Kenya. Health Place 13, 504–19. https://doi.org/10.1016/j.healthplace.2006.05.009

Goldenberg, S.M., Chettiar, J., Nguyen, P., Dobrer, S., Montaner, J., Shannon, K., 2014. Complexities of short-term mobility for sex work and migration among sex workers: Violence and sexual risks, barriers to care, and enhanced social and economic opportunities. Journal of Urban Health : Bulletin of the New York Academy of Medicine 91, 736–751. https://doi.org/10.1007/s11524-014-9888-1

Google, 2018. Https://www.google.co.uk/maps/.

Hagen-Zanker, J., 2008. Why do people migrate? A review of the theoretical literature. A Review of the Theoretical Literature (January 2008). Maastrcht Graduate School of Governance Working Paper.

Hall, C.M., 2005. Reconsidering the geography of tourism and contemporary mobility. Geographical Research 43, 125–139.

Halli, S.S., Buzdugan, R., Moses, S., Blanchard, J., Jain, A., Verma, R., Saggurti, N., 2010. High-risk sex among mobile female sex workers in the context of jatras (religious festivals) in karnataka, india. International Journal of STD & AIDS 21, 746–751. https://doi.org/10.1258/ijsa.2010.010192

Hanson, S., 2005. Perspectives on the geographic stability and mobility of people in cities. Proceedings of the National Academy of Sciences 102, 15301–15306.

Harcourt, C., Donovan, B., 2005. The many faces of sex work. Sexually Transmitted Infections 81, 201–206.

Hargreaves, J.R., Fearon, E., Davey, C., Phillips, A., Cambiano, V., Cowan, F.M., 2016. Statistical design and analysis plan for an impact evaluation of an HIV treatment and prevention intervention for female sex workers in Zimbabwe: a study protocol for a cluster randomised controlled trial. Trials 17, 6.

Heckathorn, D., 1997. Respondent-driven sampling: A new approach to the study of hidden populations. Social Problems 44, 174–199.

Hong, Y., Zhang, C., Li, X., Fang, X., Lin, X., Zhou, Y., Liu, W., 2012. HIV testing behaviors among female sex workers in Southwest China. AIDS and Behavior 16, 44–52.

Hooimeijer, P., Van der Knaap, G., 1994. From flows of people to networks of behaviour. Nederlandse Geografische Studies 173, 178–187.

Jain, A.K., Saggurti, N., 2012. The extent and nature of fluidity in typologies of female sex work in southern India: implications for HIV prevention programs. Journal of HIV/AIDS & social services 11, 169–191.

Kahle, D., Wickham, H., 2013. Ggmap: Spatial visualization with ggplot2. The R Journal 5, 144–161.

Malekinejad, M., Johnston, L.G., Kendall, C., Kerr, L.R.F.S., Rifkin, M.R., Rutherford, G.W., 2008. Using respondent-driven sampling methodology for HIV biological and behavioral surveillance in international settings: a systematic review. AIDS and Behavior 12, 105–130.

Masyn, K.E., 2013. Chapter 25 latent class analysis and finite mixture modeling. The Oxford handbook of quantitative methods 551.

McLaughlin, K.R., Handcock, M.S., Johnston, L.G., Japuki, X., Gexha-Bunjaku, D., Deva, E., others, 2015. Inference for the visibility distribution for respondent-driven sampling. American Statistical Association, Alexandria, VA.

Morison, L., Weiss, H.A., Buve, A., Carael, M., Abega, S.-C., Kaona, F., Kanhonou, L., Chege, J., Hayes, R.J., Heterogeneity of HIV Epidemics in African Cities, S.G. on, others, 2001. Commercial sex and the spread of HIV in four cities in sub-Saharan Africa. Aids 15, S61–S69.

Nylund, K.L., Asparouhov, T., Muthén, B.O., 2007. Deciding on the number of classes in latent class analysis and growth mixture modeling: A Monte Carlo simulation study. Structural equation modeling 14, 535–569.

Patel, S.K., Ganju, D., Prabhakar, P., Adhikary, R., 2016. Relationship between mobility, violence and major depression among female sex workers: A cross-sectional study in southern India. BMJ Open 6, e011439.

Pebesma, E.J., Bivand, R.S., 2005. Classes and methods for spatial data in R. R News 5, 9–13.

Pollard, H., 1996. Seasonal and permanent moves among the elderly, in: Newton, P., Bell, M. (Eds.), Population Shift: Mobility and Change in Australia. AGPS, pp. 378–91.

RedTraSex, 2016. RedTraSex statement at the high level.

Reed, E., Gupta, J., Biradavolu, M., Blankenship, K., 2012. Migration/mobility and risk factors for HIV among female sex workers in Andhra Pradesh, India: implications for HIV prevention. International journal of STD & AIDS 23, e7–e13.

Reed, E., Gupta, J., Biradavolu, M., Devireddy, V., Blankenship, K.M., 2011. The role of housing in determining HIV risk among female sex workers in Andhra Pradesh, India: considering women's life contexts. Social Science & Medicinee & medicine 72, 710–716.

Richardson, A., Seethaler, R., 2001. Estimating long distance travel behaviour from the most recent trip. TRB Transportation Research Circular E-C026—Personal Travel: The Long and Short of It.

Richter, M., Chersich, M.F., Vearey, J., Sartorius, B., Temmerman, M., Luchters, S., 2014. Migration status, work conditions and health utilization of female sex workers in three South African cities. Journal of immigrant and minority health 16, 7–17.

Scambler, G., Paoli, F., 2008. Health work, female sex workers and HIV/AIDS: Global and local dimensions of stigma and deviance as barriers to effective interventions. Social science & medicine 66, 1848–1862.

Shannon, K., Csete, J., 2010. Violence, condom negotiation, and HIV/STI risk among sex workers. JAMA 304, 573–574.

Sinha, M., 1997. A summary of the findings of the high risk behaviour study from 18 cities. National AIDS Control Organisation, Ministry of Health & Family Welfare.

Smith, S.K., 1989. Toward a methodology for estimating temporary residents. Journal of the American Statistical Association 84, 430–436.

Steen, R., Jana, S., Reza-Paul, S., Richter, M., 2015. Trafficking, sex work, and HIV: efforts to resolve conflicts. The Lancet 385, 94–96.

Swindale, A., Bilinsky, P., 2006. Development of a universally applicable household food insecurity measurement tool: Process, current status, and outstanding issues. The Journal of nutrition 136, 1449S–1452S.

Taylor, B.S., Garduño, L.S., Reyes, E.V., Valiño, R., Rojas, R., Donastorg, Y., Brudney, K., Hirsch, J., 2011. HIV care for geographically mobile populations. Mount Sinai Journal of Medicine: A Journal of Translational and Personalized Medicine 78, 342–351.

Vanwesenbeeck, I., 2001. Another decade of social scientific work on sex work: A review of research 1990–2000. Annual review of sex research 12, 242–289.

Visser, I., Speekenbrink, M., 2010. depmixS4: An R package for Hidden Markov Models. Journal of Statistical Software 36, 1–21.

Volz, E., Heckathorn, D.D., 2008. Probability based estimation theory for respondent driven sampling. Journal of Official Statistics 24, 79–97.

Ward, H., Day, S., 2006. What happens to women who sell sex? Report of a unique occupational cohort. Sexually Transmitted Infections 82, 413–417.

Wilson, D., Chiroro, P., Lavelle, S., Mutero, C., 1989. Sex worker, client sex behaviour and condom use in Harare, Zimbabwe. AIDS care 1, 269–280.

Zelinsky, W., 1971. The hypothesis of the mobility transition. Geographical review 219–249.

Mobility and sex work: why, where, when? A typology of female-sex-worker mobility in Zimbabwe

Tables

22 October, 2018

Table I: demographics and mobility data

Table II: mobility at the journey, woman, and site levels

Table III: sociodemographic and behavioural risks of different types of movers

Table I: demographics and mobility data

	Round 1, 2013	Round 2, 2016
N	2,591	2,839
Age		
18-19	103; 4%	50; 2%
20-24	528; 20%	448; 15%
25-29	637; 24%	602; 21%
30-39	907; 36%	1,136; 41%
40+	416; 17%	603; 21%
Education level		
No education	815; 32%	881; 30%
Primary	913; 35%	993; 35%
Secondary	848; 33%	964; 35%
Marital status		
Married	6; 0%	41; 1%
Divorced	1,652; 63%	1,823; 64%
Widowed	459; 18%	552; 20%
Never married	474; 18%	422; 15%
Child support (IQR)	1 (1, 2)	1 (1, 2)
Food insecure	1,139; 45%	1,175; 42%
Age started SW (IQR)	23 (20, 28)	24 (20, 29)
Clients in last week (IQR)†	5 (3, 10)	5 (3, 10)
Consistently uses condoms	1,512; 65%	1,524; 58%
SW only source of income	1,423; 55%	1,292; 45%
USD per short sex act (IQR)	4 (4, 8)	4 (4, 4)
Community considers as SW		1,926; 69%
Place of solicitation		
Bar	1,862; 72%	1,733; 64%
Telephone	162; 7%	136; 5%
Street	442; 17%	494; 19%
Lodge	16; 1%	24; 1%
None of the above	106; 4%	299; 11%
HIV	1,531; 60%	1,671; 59%
Visibility (IQR)*	6 (5, 9)	6 (5, 8)
Mobility		
Years living at site	8 (3, 21)	9 (3, 23)
Worked elsewhere past 12 months	704; 27%	1,658; 59%
Stayed elsewhere past 12 months		1,559; 55%
Any mobility in past 12 months	704; 27%	2,294; 81%
At least one place named	681	2,293
Total places named	929	5,654
with geocodes	929	5,612
type of place identified	929	5,611
with distances from site	929	5,612

Table 1: Data were missing for: age started sex work (2 women in 2016), consistent condom use (246 women in 2013, 211 in 2016), the price of a sex act (17 women in 2013, 12 in 2016), place of solicitation (153 women in 2016), length of time at the interview site (15 women in 2013), HIV status (16 women in 2013), and whether worked elsewhere (1 women in 2013, 6 in 2016). †Women were asked whether this was more or less than average: in 2013 and 2016, 802/2575 and 920/2833 women said the last week was average, respectively. The majority in both years, 1,501 and 1,612, said this was less than usual. * Visibility reported unweighted.

Table II: mobility at the journey, woman, and site levels

	Round 1, 2013	Round 2, 2016	
		Work	Non-work
N journeys	929	3,364	2,248
Median distance, km (IQR)	246 (171-227)	246 (150-238)	314 (185-309)
Median travel-time, hours (IQR)	3.4 (2-3)	3.3 (2-3)	4.3 (3-4)
Median stay, days (IQR)	21 (4-61)	18 (4-21)	4 (4-21)
Intra-province	401; 43%	1,550; 46%	813; 36%
Inter-province	372; 40%	1,387; 41%	1,206; 54%
International	156; 17%	427; 13%	229; 10%
To different linguistic area	90; 10%	289; 9%	165; 7%
To town or city	551; 59%	1,811; 54%	1,302; 58%
To growth point	127; 14%	543; 16%	338; 15%
To farm or mine	143; 15%	599; 18%	289; 13%
To other	108; 12%	410; 12%	319; 14%
Number of places	164	270	262
Median visits per place	2 (1, 6)	3 (1, 11)	2 (1, 7)
Reasons for travelling			
Clients with more money		1,903; 57%	
More clients		1,313; 39%	
Earn more for being new		848; 25%	
Travelled with client		306; 9%	
Know FSW there		220; 7%	
Familiar place		97; 3%	14; 1%
Anonymity		154; 5%	6; 0%
Holiday or festival		26; 1%	71; 3%
Police harassment		10; 0%	1; 0%
Avoid regular clients		13; 0%	2; 0%
Avoid other FSW		54; 2%	5; 0%
Avoid boyfriend or husband		15; 0%	1; 0%
Avoid family		25; 1%	4; 0%
Be with boyfriend or husband		66; 2%	30; 1%
Visit family		197; 6%	1,614; 72%
Visit children		16; 0%	113; 5%
Work other than sex work		124; 4%	229; 10%
To use medical services		3; 0%	26; 1%
Other		3; 0%	26; 1%
Used healthcare		398; 12%	185; 8%
Individuals			
N Individuals	681	1,651	1,548
Median visits (IQR)	1 (1, 1)	2 (1, 3)	1 (1, 2)
Median different places (IQR)	1 (1, 1)	2 (1, 3)	1 (1, 2)
Median total time away (IQR)	21 (4, 61)	24 (7, 64)	14 (4, 28)
Visited town/city	448; 64%	1,165; 70%	1,003; 64%
Visited growth point	116; 16%	442; 27%	319; 20%
Visited mine/farm	123; 17%	448; 27%	261; 17%
Median to-from dist. (IQR)	356 (180, 662)	464 (229, 967)	443 (208, 850)
Median circuit dist (IQR)	357 (357, 361)	464 (460, 468)	424 (422, 426)
Sites			
N Sites	14	14	14
Median journeys/site (IQR)	68 (54, 74)	246 (170, 290)	161 (133, 180)
Median percentage who travel (IQR)	28 (22, 31)	59 (46, 65)	55 (47, 60)
Median duration away (IQR)	4 (4, 21)	6 (4, 21)	4 (4, 4)
Median distance (IQR)	123 (101, 160)	136 (90, 178)	182 (108, 223)

Table 2: Mobility of sex workers described at the level of the journey, woman, and site of interview. Proportions and medians with Inter-quartile range are reported. There were 7 journeys without any reasons given.

Table II: sociodemographic and behavioural risks of different types of movers

	Mobility classes					
	Class 1	Class 2	Class 3	Class 4	Class 5	**Not-mobile**
N	114	263	264	343	624	1175
Age						
18-19	2; 1%	5; 2%	4; 1%	4; 1%	10; 1%	23; 2%
20-24	18; 16%	41; 14%	52; 19%	55; 15%	103; 16%	166; 14%
25-29	33; 29%	83; 32%	56; 21%	69; 21%	140; 21%	205; 17%
30-39	46; 39%	95; 36%	120; 47%	145; 43%	254; 41%	458; 40%
40+	15; 14%	39; 16%	32; 12%	70; 20%	117; 20%	323; 27%
Education level						
No education	36; 31%	62; 23%	96; 34%	125; 35%	142; 22%	404; 33%
Primary	47; 44%	94; 35%	86; 34%	113; 32%	238; 39%	388; 34%
Secondary	31; 25%	107; 42%	82; 33%	105; 32%	243; 39%	383; 33%
Marital status						
Married	0; 0%	2; 1%	0; 0%	6; 2%	9; 1%	23; 2%
Divorced	76; 67%	195; 74%	183; 69%	212; 62%	407; 65%	710; 60%
Widowed	21; 19%	31; 12%	50; 20%	71; 20%	109; 18%	259; 22%
Never married	17; 14%	35; 13%	30; 11%	54; 16%	99; 16%	183; 16%
Child support (IQR)	1 (1, 2)	1 (1, 2)	1 (1, 1)	1 (1, 2)	1 (1, 2)	1 (1, 2)
Food insecure	40; 35%	99; 38%	143; 56%	151; 46%	245; 40%	469; 40%
Age started SW (IQR)	24 (20, 29)	23 (20, 27)	24 (20, 28)	23 (19, 29)	23 (19, 28)	25 (20, 31)
Clients in last week (IQR)	5 (3, 10)	6 (3, 12)	6 (4, 10)	5 (3, 10)	5 (3, 10)	4 (2, 8)
Consistently uses condoms	48; 45%	169; 67%	125; 56%	140; 46%	345; 57%	665; 62%
SW only source of income	63; 55%	138; 51%	170; 66%	123; 36%	256; 40%	512; 44%
USD per short sex act (IQR)	4 (4, 4)	4 (4, 4)	4 (4, 4)	4 (4, 4)	4 (4, 8)	4 (4, 8)
Community consider as SW	82; 73%	163; 61%	199; 77%	245; 73%	445; 72%	749; 65%
Place of solicitation						
Bar	79; 73%	157; 65%	197; 77%	187; 60%	397; 67%	684; 60%
Telephone	4; 4%	17; 7%	13; 5%	9; 3%	32; 6%	57; 5%
Street	14; 14%	45; 19%	39; 14%	81; 26%	73; 13%	230; 21%
Lodge	0; 0%	1; 1%	2; 1%	5; 2%	7; 1%	9; 1%
None of the above	10; 10%	19; 9%	8; 3%	34; 10%	75; 13%	151; 13%
Years living at site (IQR)	5 (3, 9)	7 (3, 19)	7 (3, 22)	9 (4, 24)	7 (3, 20)	11 (4, 26)
HIV	79; 70%	144; 54%	175; 67%	190; 53%	341; 56%	714; 61%
Visibility	7 (6, 9)	6 (4, 9)	7 (6, 9)	6 (5, 8)	6 (4, 8)	6 (5, 8)
Stayed elsewhere	69; 61%	158; 60%	164; 62%	190; 55%	315; 50%	636; 54%

Table 3: There were 50 women excluded (missing data on at least one of the manifest variables), and 6 women missing data on working elsewhere. Class 1 travelled with clients. Class 2 travelled often over long distances and internationally. Class 3 worked when they travelled for other reasons. Class 4 travelled frequently over short distances for moderate periods. Class 5 moved domestically infrequently.

Mobility and sex work: why, where, when? A typology of female-sex-worker mobility in Zimbabwe

Figures

22 October, 2018

Figure 1: Circuits

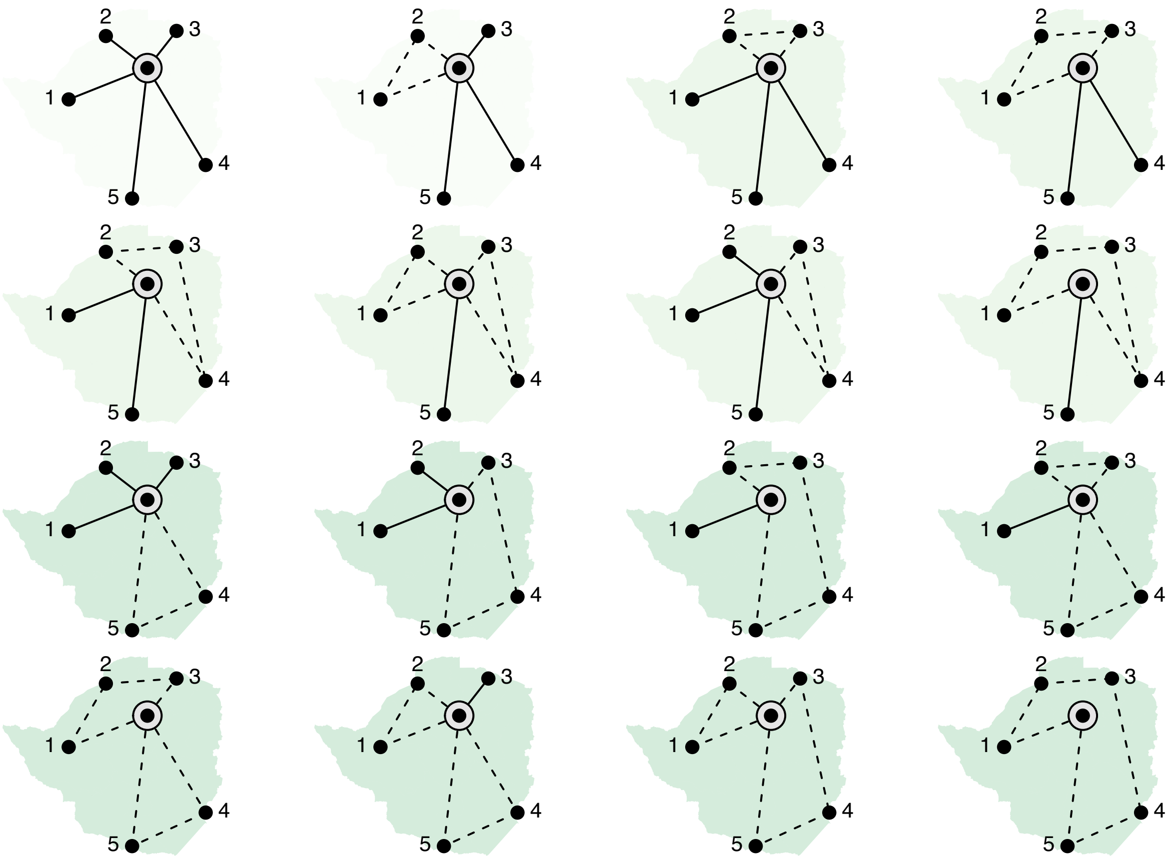


Figure 1

6

Figure 2: latent class analysis results

	Class 1 (114, 7%)	Class 2 (263, 16%)	Class 3 (264, 16%)	Class 4 (343, 21%)	Class 5 (624, 39%)
Visited growth point/business centre	19%	16%	10%	24%	39%
Visited mine/farm	2%	14%	14%	100%	0%
Clients with more money	12%	86%	16%	86%	81%
More clients	1%	68%	6%	62%	64%
Earn more for being new	2%	47%	10%	43%	40%
Travelled with client	100%	15%	3%	4%	3%
Know FSW there	5%	16%	19%	10%	9%
Anonymity	0%	14%	3%	7%	8%
Visit family	4%	10%	32%	6%	3%
0ther	9%	32%	49%	14%	12%
Lives in a town	64%	75%	64%	90%	64%
Median stay over two weeks	14%	71%	62%	40%	56%
Travelled internationally	34%	93%	17%	2%	4%
Median journey time (hours)	2.45	3.58	2.04	1.21	1.64
Number of visits in 12 months	1.81	2.51	1.51	2.44	1.74

Figure 2

Figure 3: density of destinations from sites

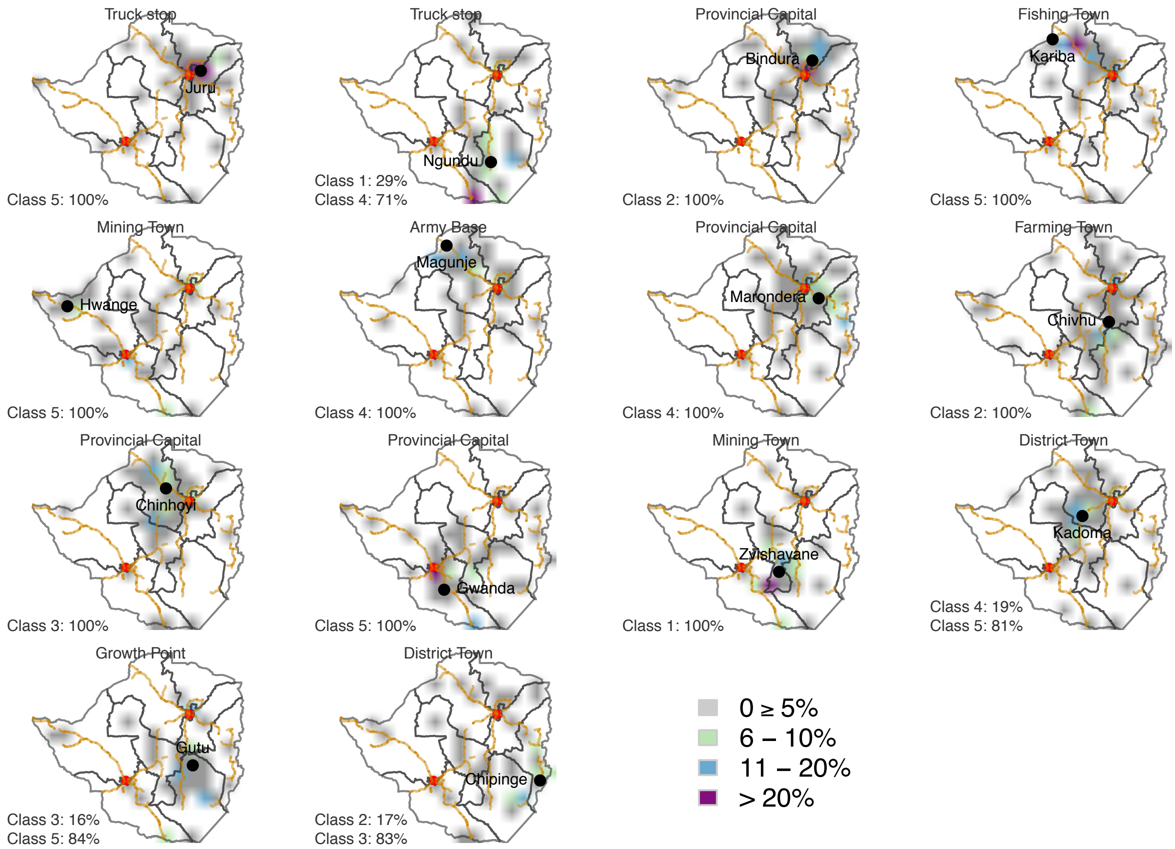


Figure 3

Captions

Figure 1

Illustration of the possible routes between places visited. Solid lines represent there-and-back journeys, dashed-lines represent one-way journeys. The first cell (top left) shows the there-and-back journey that is the only possible option with just one place reported. The second shows the second possible journey if two places are reported. Cells three and four show the two additional journeys possible with three places reported. Line two shows the four additional journeys possible with four reported places, and the bottom two rows show the eight additional journeys possible when five places were reported. Women always start from and return to the interview site and can choose to return or move to the next places from each place visited, therefore for *j* places there are *j*-1 binary decisions, and hence 2^{j-1} journeys.

Figure 2

Results of the five-class latent class model, with the classes ordered by size from left to right. For the binary manifest variables, the posterior probabilities are shown with a continuous colour-scale from white (zero probability) to black (100\% probability). The exponentiated intercepts of the median journey time and number of visits models are shown, i.e. the mean for the class.

Figure 3

The density of reported visits are shown for each interview site (labelled), with the type of place above each plot, using a 17x15 grid of 255 50x50km squares. The proportion of the mobile women in each class from the six-class LCA model are shown on the bottom-left of each figure.

Appendix 1: recoding of duration-away variable

Duration in days	2013 coding	2016 coding
0	—	Did not spend the night
3.5	Less than a week	Less than a week
17.5	—	1-2 weeks
21	2-4 weeks	2-4 weeks
60.8	1-3 months	1-3 months
152.1	—	4-6months
273.8	>6months	>6months

Appendix 2: configurations of reasons for moving

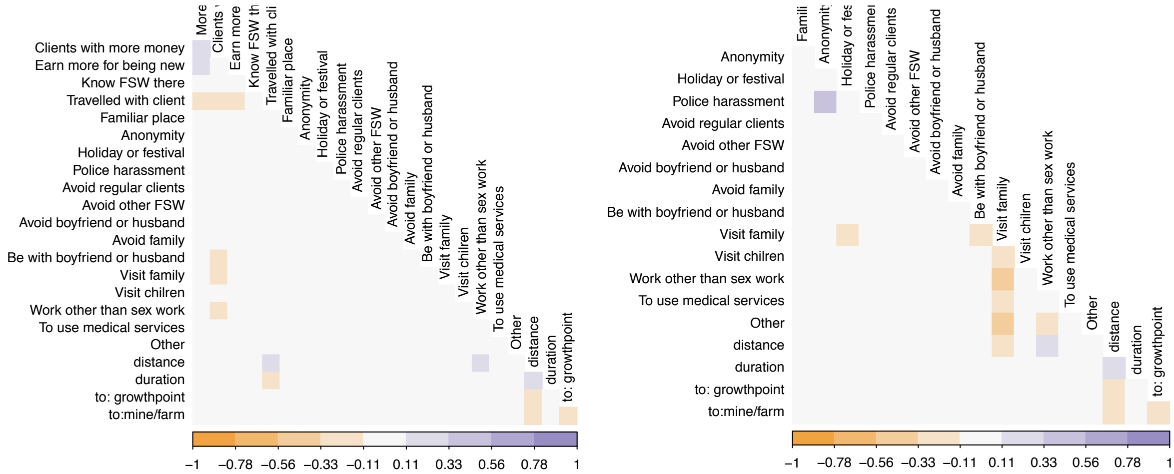
Women could report multiple reasons for each journey reported. Therefore, we explored the distribution of the reasons. We plotted the correlation at the journey-level between reasons, distance, and time spent away, separated by whether the visit included sex work. We listed the configurations of reasons by frequency, and reported the twenty most common configurations. The rows show the no / yes (0/1) response for each reason for travelling, in the following order: Clients with more money*, More clients*, Earn more for being new*, Travelled with client*, Know FSW there*, Familiar place, Anonymity, Holiday or festival, Police harassment, Avoid regular clients, Avoid other FSW, Avoid boyfriend or husband, Avoid family, Be with boyfriend or husband, Visit family, Visit children, Work other than sex work, To use medical services, Other. The reasons marked with an asterisk (*) were not offered for moves that did not included working in sex work (e.g. for the second table below).

	cases
0-1-0-0-0-0-0-0-0-0-0-0-0-0-0-0-0-0-0	656
1-1-0-0-0-0-0-0-0-0-0-0-0-0-0-0-0-0-0	514
0-0-0-0-1-0-0-0-0-0-0-0-0-0-0-0-0-0-0	267
1-1-1-0-0-0-0-0-0-0-0-0-0-0-0-0-0-0-0	239
0-0-1-0-0-0-0-0-0-0-0-0-0-0-0-0-0-0-0	213
1-0-0-0-0-0-0-0-0-0-0-0-0-0-0-0-0-0-0	200
0-1-1-0-0-0-0-0-0-0-0-0-0-0-0-0-0-0-0	149
0-0-0-0-0-0-0-0-0-0-0-0-0-0-1-0-0-0-0	132
1-0-1-0-0-0-0-0-0-0-0-0-0-0-0-0-0-0-0	98
0-0-0-1-0-0-0-0-0-0-0-0-0-0-0-0-0-0-0	94
0-0-0-0-0-0-0-0-0-0-0-0-0-0-0-0-1-0-0	71
0-0-0-0-0-0-0-0-0-0-0-0-0-1-0-0-0-0-0	60
0-1-0-1-0-0-0-0-0-0-0-0-0-0-0-0-0-0-0	41
0-0-0-0-0-0-0-0-0-0-0-0-0-0-0-0-0-0-1	35
1-1-0-0-0-0-1-0-0-0-0-0-0-0-0-0-0-0-0	33
0-0-0-0-0-0-0-0-0-0-1-0-0-0-0-0-0-0-0	31
0-0-0-0-0-0-1-0-0-0-0-0-0-0-0-0-0-0-0	29
0-0-0-0-0-1-0-0-0-0-0-0-0-0-0-0-0-0-0	23
1-1-0-1-0-0-0-0-0-0-0-0-0-0-0-0-0-0-0	20
0-1-0-0-0-0-1-0-0-0-0-0-0-0-0-0-0-0-0	18
Remaining cases	441
Remaining reasons	145
Total	3364

For the 3,364 moves that included sex work, the most common combination accounted for 656 of the journeys, and the second 514, leaving 2,194 journeys with other combinations of reasons. In contrast, for moves that did not include sex work, the most common combination of reasons constituted more than half of the total journeys (1,533/2,248; 68%).

	cases
0-0-0-0-0-0-0-0-0-1-0-0-0-0	1533
0-0-0-0-0-0-0-0-0-0-0-1-0-0	222
0-0-0-0-0-0-0-0-0-0-0-0-0-1	218
0-0-0-0-0-0-0-0-0-0-1-0-0-0	71
0-0-1-0-0-0-0-0-0-0-0-0-0-0	61
0-0-0-0-0-0-0-0-0-1-1-0-0-0	37
0-0-0-0-0-0-0-0-1-0-0-0-0-0	27
0-0-0-0-0-0-0-0-0-0-0-0-1-0	20
0-0-0-0-0-0-0-0-0-1-0-0-0-1	12
0-0-1-0-0-0-0-0-0-1-0-0-0-0	7
1-0-0-0-0-0-0-0-0-0-0-0-0-0	7
0-0-0-0-0-1-0-0-0-0-0-0-0-0	5
0-0-0-0-0-0-0-0-0-1-0-0-1-0	2
0-0-0-0-0-0-0-0-1-1-0-0-0-0	2
0-0-0-0-0-0-0-1-0-0-0-0-0-0	2
0-0-0-0-1-0-0-0-0-0-0-0-0-0	2
0-1-0-0-0-0-0-0-0-0-0-0-0-0	2
0-1-0-0-0-0-0-0-0-1-0-0-0-0	2
1-0-0-0-0-0-0-0-0-1-0-0-0-0	2
0-0-0-0-0-0-0-0-0-0-1-0-1-0	1
Remaining cases	13
Remaining reasons	14
Total	2248

Fig. A2: correlations between reasons for moving



We found little evidence of correlations between the reasons given for moving (see Appendix **Figure A2**) For moves that included sex work, there was a weak positive correlation between moving for more money and more clients, and weak negative correlation between moving for more money and towards relationships, work other than sex work, or travel with a client. For journeys that did not involve sex work, there was positive correlation between police harassment and anonymity; and negative correlations between visiting family and work other then sex work and 'other'.

Appendix 3: LCA diagnostics

Summary

We ran a series of diagnositics to evaluate the fit of models of classes around the minimum of the BIC plot, adapted from Garrett & Zeger (2000). We calculated the difference between the observed and expected frequencies of each binary variable, and the median of the continuous variables. We computed the bivariate residuals comparing the observed and expected relationships between each of the binary variables. We plotted the observed and expected number of visits reported, log of the median distance travelled, and the frequency of the 50 most common configurations of the reasons for travelling.

To compare the most persemoneous model with models with more classes, we used a bootstrap log-likelihood-ratio test (McLachlan, 1987). This test is more robust than tests that make asymtotic assumptions (Reynolds & Templin, 2004). We followed the steps in Tekle, Gudicha & Vermunt (2016): the most parsimonious model was used to predict values for the manifest variable; both of the models being compared were then fitted to this predicted data; the log-likelihood ratio was computed; this was repeated 500 times to produce a distribution of the log-likelihood ratio under the null condition, and the observed log-likelihood ratio was compared to this distribution, with the *p*-value calculated as the proportion of the distribution higher than the observed value. Finally, we presented the probabilities for each maniest variable in each class, and considered the substantive usefulness of additional classes.

Variables in the LCA

Variable	Data type	Data transformation / link function
More clients	Binary	Logit
Clients with more money	Binary	Logit
Earn more for being new	Binary	Logit
Know FSW there	Binary	Logit
Travelled with client	Binary	Logit
Anonymity	Binary	Logit
Visit family	Binary	Logit
Other	Binary	Logit
Visited growth point	Binary	Logit
Visited mine or farm	Binary	Logit
Lives in a town	Binary	Logit
Median stay over two weeks	Binary	Logit
Travelled internationally	Binary	Logit
Median journey time (hours)	>6months	Log
Number of visits in 12 months	>6months	Poisson

References

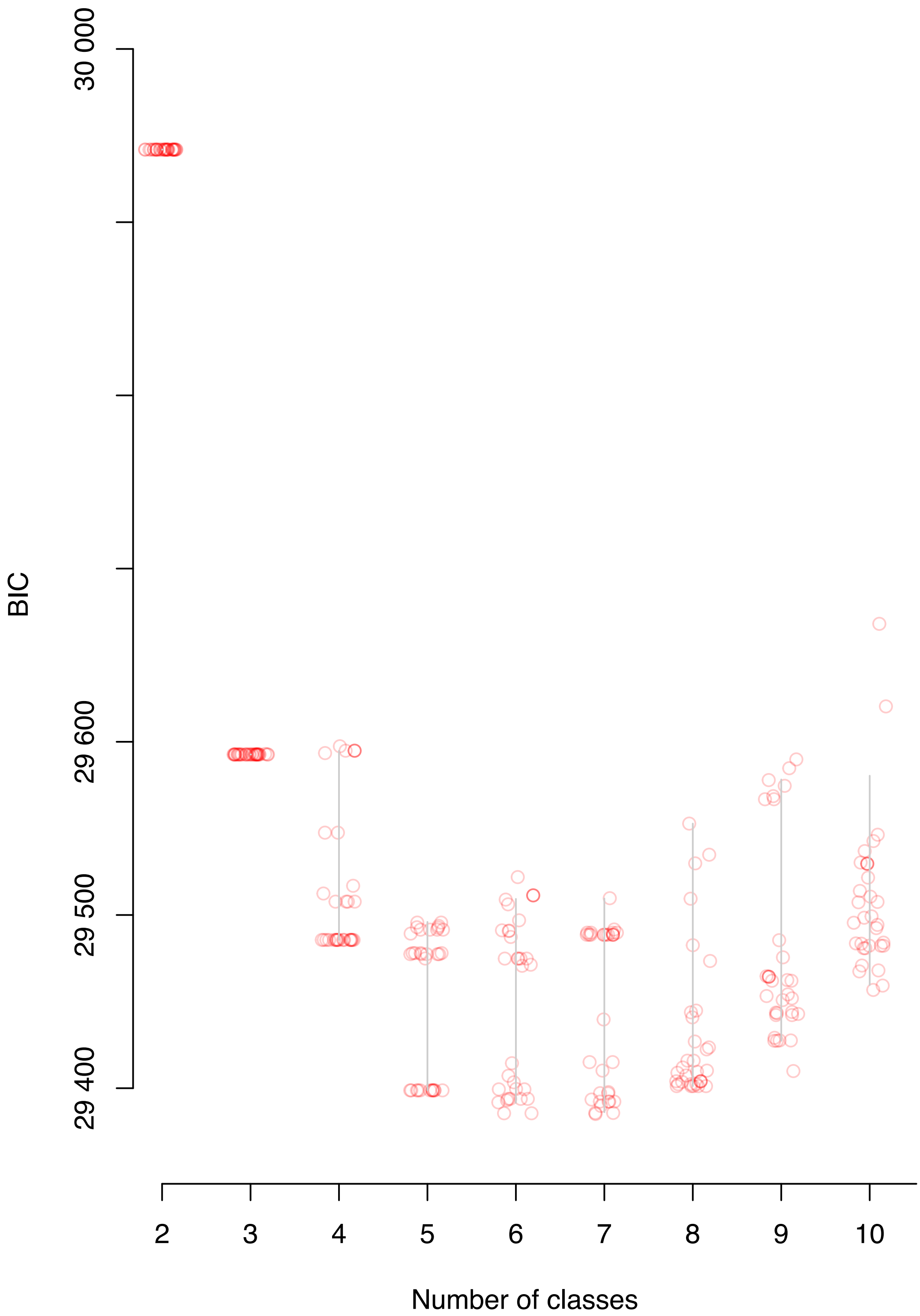
Garrett, E.S. & Zeger, S.L. (2000) Latent class model diagnosis. *Biometrics*. 56 (4), 1055–1067.

McLachlan, G.J. (1987) On bootstrapping the likelihood ratio test stastistic for the number of components in a normal mixture. *Applied statistics*. 318–324.

Reynolds, J.H. & Templin, W.D. (2004) Comparing mixture estimates by parametric bootstrapping likelihood ratios. *Journal of Agricultural, Biological, and Environmental Statistics*. 9 (1), 57.

Tekle, F.B., Gudicha, D.W. & Vermunt, J.K. (2016) Power analysis for the bootstrap likelihood ratio test for the number of classes in latent class models. *Advances in Data Analysis and Classification*. 10 (2), 209–224.

Fig. A3.0: BIC



The BIQ figure (**Figure A3.0**) shows that between five and seven classes best fit the data. The grey vertical lines line show the 95% range of the results of 100 runs. The plot shows that from five classes, or even from four, there is a bi-modal distribution in the BIC centered around 29 400 and 29 500. All three of these models produced predicted proportions of the binary variables that were very similar to the observed data (see **Figures A3.1, A3.2, and A3.3**). The five-class model produced some large residuals in the bivariate residuals analysis, especially for the 'travelled with a client' reason for moving, for six and seven classes the residuals were all moderate or small (see **Figures A3.4, A3.5, and A3.6**). The correlations between observed and expected numbers of visits, log median distance, and combinations of reasons were similar for all three models (see **Figures A3.7, A3.8, and A3.8**).

In the bootstrap-log-likelihood ratio tests, there was evidence for a difference between the four-classs and the five-class models, but not between the five-class and the six-class models (see **Figures A3.10 and A3.11**).

Contrasting the class sizes predicted from each model (see **Figures A3.12, A3.13, and A3.14**), and the probability profiles, revealed that there were few substantive differences between the five and six class models. Class 5 in the five-class model decomposed into classes 2 and 6 in the six-class model. Increasing the number of classes from four to five revealed the small class of women who often travelled with clients, who were otherwise classed with the women who worked while moving for other reasons (class 3).

Observed and predicted proportions

Fig. A3.1: Five classes

	Observed	Expected	Difference
med.over2weeks	0.5	0.5	0
international	0.2	0.2	0
town	0.7	0.7	0
growthpoint	0.3	0.2	0
minefarm	0.3	0.2	0
r_1	0.5	0.5	0
r_2	0.7	0.7	0
r_3	0.3	0.3	0
r_4	0.1	0.1	0
r_5	0.1	0.1	0
r_7	0.1	0.1	0
r_15	0.1	0.1	0
r_other	0.2	0.2	0
visits	2.0	2.0	0
med.time	1.7	1.7	0

Fig. A3.2: Six classes

	Observed	Expected	Difference
med.over2weeks	0.5	0.6	0.0
international	0.2	0.2	0.0
town	0.7	0.7	0.0
growthpoint	0.3	0.3	0.0
minefarm	0.3	0.3	0.0
r_1	0.5	0.5	0.0
r_2	0.7	0.6	0.0
r_3	0.3	0.4	0.0
r_4	0.1	0.1	0.0
r_5	0.1	0.1	0.0
r_7	0.1	0.1	0.0
r_15	0.1	0.1	0.0
r_other	0.2	0.2	0.0
visits	2.0	2.0	0.0
med.time	1.7	1.8	-0.1

Fig. A3.3: Seven classes

	Observed	Expected	Difference
med.over2weeks	0.5	0.6	0.0
international	0.2	0.2	0.0
town	0.7	0.7	0.0
growthpoint	0.3	0.3	0.0
minefarm	0.3	0.2	0.0
r_1	0.5	0.5	0.0
r_2	0.7	0.7	0.0
r_3	0.3	0.3	0.0
r_4	0.1	0.1	0.0
r_5	0.1	0.1	0.0
r_7	0.1	0.1	0.0
r_15	0.1	0.1	0.0
r_other	0.2	0.2	0.0
visits	2.0	2.0	0.0
med.time	1.7	1.8	-0.1

Bivariate residuals

88 ### Fig. A3.4: Five classes

	international	town	growthpoint	minefarm	r_1	r_2	r_3	r_4	r_5	r_7	r_15	r_other
med.over2weeks	0.3	8.9	4.4	2.7	4.7	1.9	0.3	0.2	7.4	0.9	4.3	3.7
international		9.6	14.6	2.5	0.9	0.3	4.7	0.2	0.4	1.2	2.6	5.2
town			5.8	11.6	3.1	4.7	12.8	2.1	2.2	2.5	5.8	14.0
growthpoint				7.1	6.5	4.0	7.6	4.1	8.8	6.3	9.3	17.7
minefarm					2.6	2.8	3.7	3.8	2.8	4.5	5.4	5.5
r_1						0.6	2.0	0.7	0.7	6.5	3.3	3.8
r_2							10.6	0.9	2.1	1.1	3.1	3.5
r_3								3.0	4.5	1.9	2.5	3.6
r_4									0.4	0.9	2.5	5.9
r_5										0.9	3.7	4.6
r_7											3.2	4.2
r_15												10.8

Fig. A3.5: Six classes

	international	town	growthpoint	minefarm	r_1	r_2	r_3	r_4	r_5	r_7	r_15	r_other
med.over2weeks	6.7	4.7	8.7	6.2	7.3	9.0	5.6	6.2	8.8	7.5	19.5	5.9
international		1.3	8.3	1.0	1.2	1.5	3.5	0.2	0.9	6.2	4.6	0.6
town			11.9	1.4	1.4	1.1	1.6	0.3	1.1	4.6	2.8	1.8
growthpoint				4.1	3.4	5.9	3.8	3.0	6.6	5.9	4.9	9.2
minefarm					1.4	3.7	7.3	1.3	2.1	3.6	1.3	4.4
r_1						6.4	4.7	0.2	0.2	2.9	1.5	0.8
r_2							4.2	1.2	2.1	4.2	3.6	2.1
r_3								4.3	1.0	4.4	1.8	1.4
r_4									0.9	3.2	0.8	1.7
r_5										4.3	0.9	7.1
r_7											7.6	3.2
r_15												1.5

Fig. A3.6: Seven classes

	international	town	growthpoint	minefarm	r_1	r_2	r_3	r_4	r_5	r_7	r_15	r_other
med.over2weeks	2.7	8.8	8.0	1.6	6.5	1.9	2.1	4.7	9.0	1.2	3.5	1.6
international		4.3	9.9	0.7	2.5	4.9	0.1	3.5	2.2	0.3	2.8	0.2
town			3.8	6.6	2.6	0.6	7.2	1.9	0.8	1.2	8.3	5.8
growthpoint				1.6	4.2	0.4	0.2	0.2	1.3	2.3	2.8	1.1
minefarm					3.3	2.2	1.2	2.6	2.1	1.0	11.4	0.7
r_1						7.4	2.8	6.4	2.6	5.5	7.1	10.9
r_2							27.2	1.2	1.7	0.4	4.0	3.2
r_3								2.9	1.6	8.5	2.4	3.9
r_4									0.8	1.0	3.4	1.3
r_5										0.9	3.5	0.8
r_7											6.2	0.5
r_15												8.6

Observed vs expected plots (visits, log median distance, cases)

Fig. A3.7: Five classes

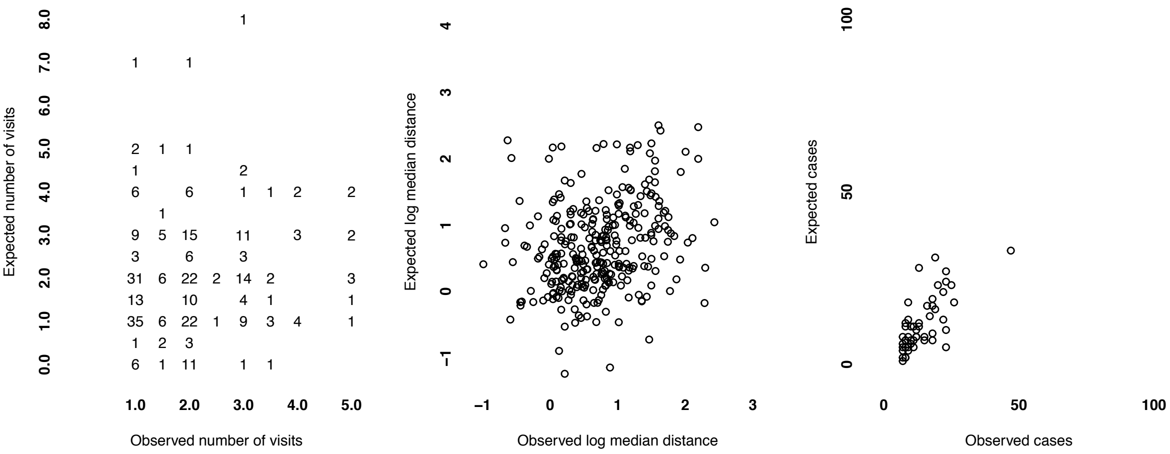


Fig. A3.8: Six classes

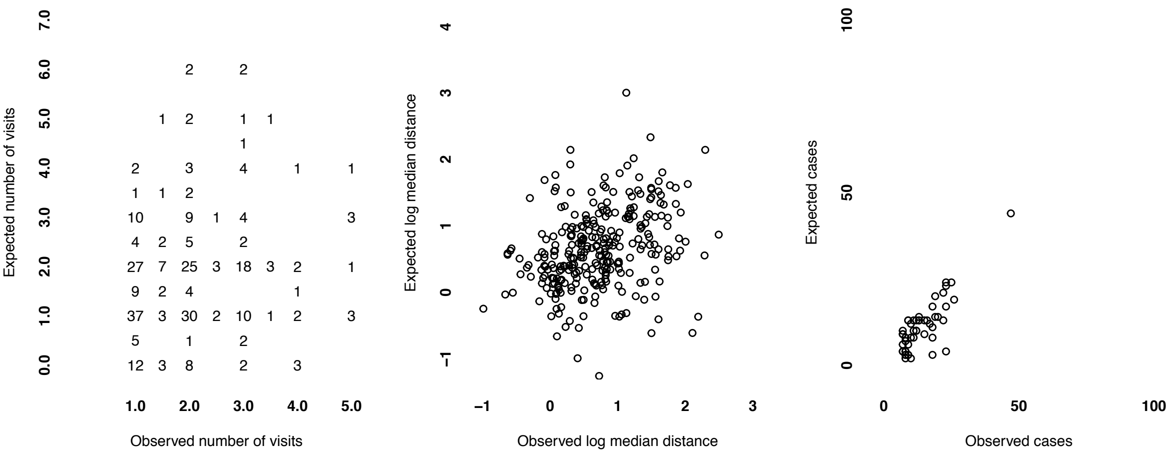
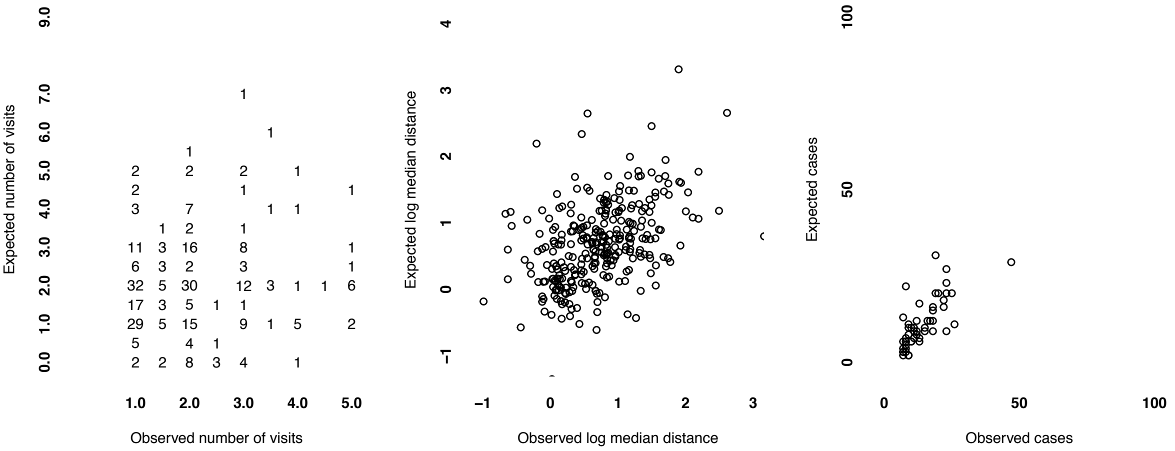
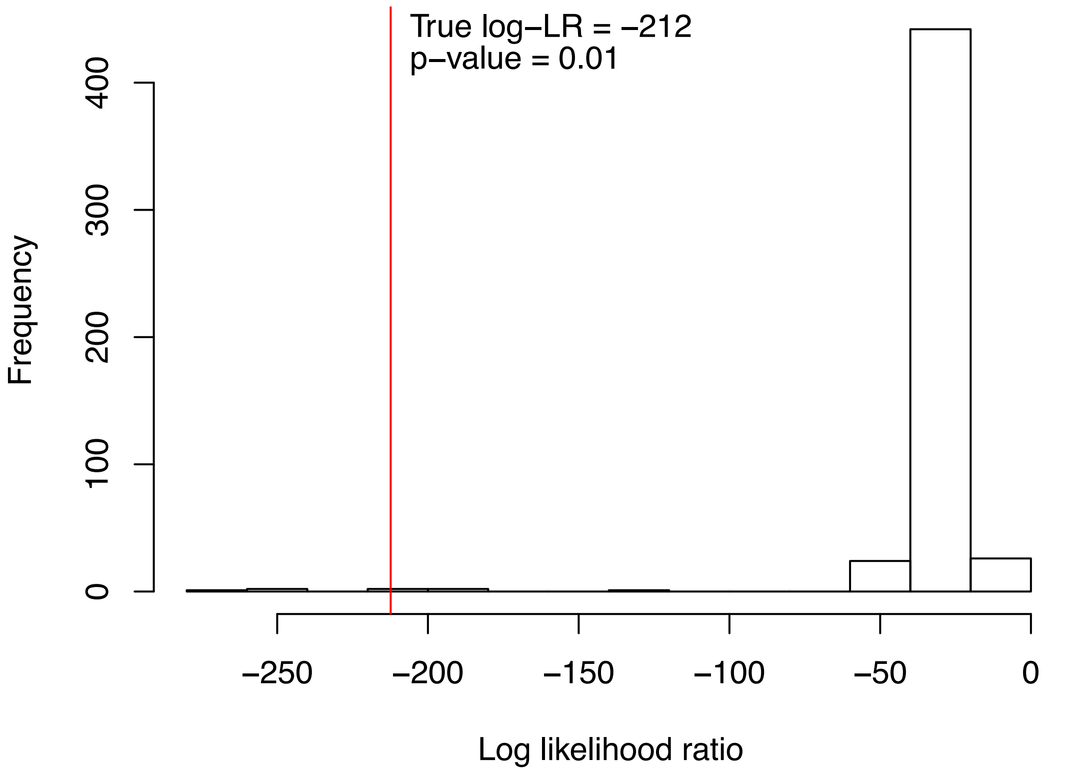


Fig. A3.9: Seven classes

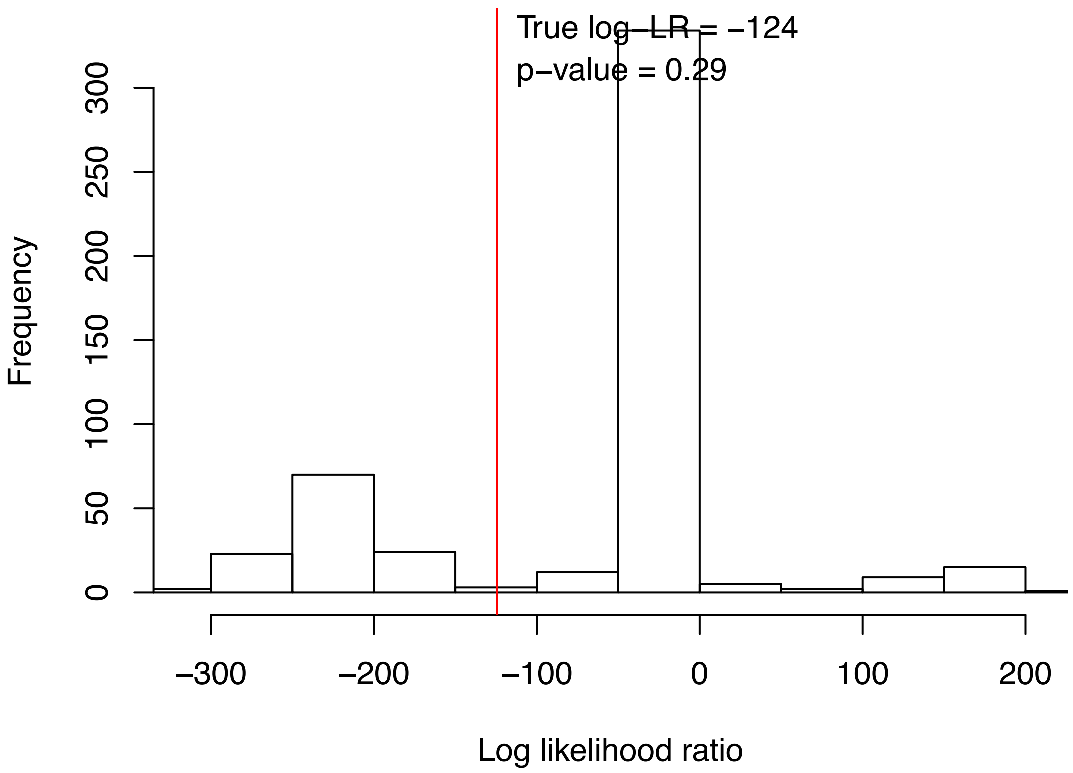


Bootstrap log-likelihood ratio tests

102 **Fig. A3.10: 4-class and 5-class models**



104 **Fig. A3.11: 5-class and 6-class models**



Classes and item probabilities

Fig. A3.12: Four-class

	Class 1 (280, 17%)	Class 2 (344, 21%)	Class 3 (347, 22%)	Class 4 (637, 40%)
Visited growth point/business centre	16%	24%	11%	39%
Visited mine/farm	14%	100%	11%	0%
Clients with more money	84%	85%	12%	80%
More clients	66%	61%	2%	63%
Earn more for being new	45%	43%	6%	39%
Travelled with client	15%	3%	30%	4%
Know FSW there	15%	10%	15%	9%
Anonymity	13%	7%	2%	8%
Visit family	10%	6%	25%	4%
0ther	31%	15%	38%	14%
Lives in a town	75%	89%	63%	64%
Median stay over two weeks	70%	41%	47%	57%
Travelled internationally	95%	1%	21%	2%
Median journey time (hours)	3.43	1.22	2.18	1.64
Number of visits in 12 months	2.48	2.43	1.56	1.75

Fig. A3.12: Five-class – repeated from Figure 2 in manuscript

	Class 1 (114, 7%)	Class 2 (263, 16%)	Class 3 (264, 16%)	Class 4 (343, 21%)	Class 5 (624, 39%)
Visited growth point/business centre	19%	16%	10%	24%	39%
Visited mine/farm	2%	14%	14%	100%	0%
Clients with more money	12%	86%	16%	86%	81%
More clients	1%	68%	6%	62%	64%
Earn more for being new	2%	47%	10%	43%	40%
Travelled with client	100%	15%	3%	4%	3%
Know FSW there	5%	16%	19%	10%	9%
Anonymity	0%	14%	3%	7%	8%
Visit family	4%	10%	32%	6%	3%
Other	9%	32%	49%	14%	12%
Lives in a town	64%	75%	64%	90%	64%
Median stay over two weeks	14%	71%	62%	40%	56%
Travelled internationally	34%	93%	17%	2%	4%
Median journey time (hours)	2.45	3.58	2.04	1.21	1.64
Number of visits in 12 months	1.81	2.51	1.51	2.44	1.74

Fig. A3.13: Six-class

	Class 1 (118, 7%)	Class 2 (156, 10%)	Class 3 (240, 15%)	Class 4 (267, 17%)	Class 5 (342, 21%)	Class 6 (485, 30%)
Visited growth point/business centre	16%	63%	17%	9%	24%	30%
Visited mine/farm	4%	0%	15%	14%	100%	0%
Clients with more money	10%	73%	87%	16%	86%	85%
More clients	1%	60%	70%	7%	61%	65%
Earn more for being new	2%	57%	50%	11%	43%	31%
Travelled with client	100%	7%	16%	0%	4%	1%
Know FSW there	7%	14%	17%	20%	10%	5%
Anonymity	0%	10%	15%	3%	7%	7%
Visit family	4%	4%	11%	33%	6%	2%
Other	11%	17%	33%	50%	14%	10%
Lives in a town	65%	35%	74%	65%	90%	76%
Median stay over two weeks	19%	30%	70%	63%	40%	66%
Travelled internationally	33%	0%	87%	18%	2%	9%
Median journey time (hours)	2.4	0.99	3.71	2.2	1.2	1.91
Number of visits in 12 months	1.78	1.88	2.6	1.52	2.44	1.65

Fig. A3.14: Seven-class

	Class 1 (68, 4%)	Class 2 (119, 7%)	Class 3 (149, 9%)	Class 4 (194, 12%)	Class 5 (245, 15%)	Class 6 (273, 17%)	Class 7 (560, 35%)
Visited growth point/business centre	79%	20%	63%	4%	9%	10%	29%
Visited mine/farm	45%	4%	35%	8%	14%	100%	0%
Clients with more money	93%	14%	82%	82%	9%	86%	81%
More clients	100%	2%	65%	65%	5%	56%	59%
Earn more for being new	79%	2%	42%	48%	7%	38%	36%
Travelled with client	15%	100%	9%	15%	2%	2%	0%
Know FSW there	35%	5%	15%	14%	21%	8%	5%
Anonymity	14%	0%	8%	17%	3%	7%	7%
Visit family	11%	4%	9%	11%	36%	5%	1%
0ther	32%	10%	31%	34%	53%	9%	8%
Lives in a town	43%	63%	97%	70%	61%	90%	63%
Median stay over two weeks	33%	15%	73%	72%	61%	36%	56%
Travelled internationally	9%	32%	28%	95%	13%	0%	7%
Median journey time (hours)	1.09	2.3	2.26	4.26	2.21	1.1	1.56
Number of visits in 12 months	3.11	1.82	2.92	2.29	1.5	2.16	1.5

Appendix 4: mobility in Zimbabwe DHS

```
dhs2011 <- read.dta(paste0(data.z,
                           '/ZW_2010-11_DHS/','zwhr62dt/','ZWHR62FL.DTA'))
dhs2015 <- read.dta(paste0(data.z,
                           '/ZW_2015_DHS/','zwhr71dt/','ZWHR71FL.DTA'))

# Total number of de jure household members gives the number of household members
# that usually live in the household -
# Total number of de facto household members gives the number of household members
#that slept in the household the previous night, including visitors.
# Standard recode manual for DHS 6: https://dhsprogram.com/pubs
#/pdf/DHSG4/Recode6_DHS_22March2013_DHSG4.pdf
diff11 <- as.data.frame(dhs2011$hv012 - dhs2011$hv013)
diff15 <- as.data.frame(dhs2015$hv012 - dhs2015$hv013)

# Count the households with people away
away11 <- car::recode(diff11[,1], "lo:0=0;1:hi=1")
away15 <- car::recode(diff15[,1], "lo:0=0;1:hi=1")

# Count the households with extra people staying
stay11 <- car::recode(diff11[,1], "lo:-1=1;0:hi=0")
stay15 <- car::recode(diff15[,1], "lo:-1=1;0:hi=0")

# Find the proportions
t4 <- cbind(rbind(weighted.mean(away11, (dhs2011$hv005/100000)),
            weighted.mean(stay11, (dhs2011$hv005/100000))),
            rbind(weighted.mean(away15, (dhs2015$hv005/100000)),
            weighted.mean(stay15, (dhs2015$hv005/100000))))
colnames(t4) <- c('2010', '2015')
rownames(t4) <- c('Proportion away', 'Proportion staying')

print(xtable(t4, align = c('l','c','c')), comment = FALSE, booktabs = T,
      sanitize.text.function = subheadings, size="\\fontsize{10pt}{10pt}\\selectfont")
```

	2010	2015
Proportion away	0.09	0.12
Proportion staying	0.07	0.09

120 **Appendix 5: geocoding code**

```
# Load geocoding
gazette           <- read.xlsx(
  "../../Data/Zimbabwe/Locations/Locations_03_12_2017.xlsx", sheet=1)
gazette$Longitude <- as.numeric(as.character(gazette$Longitude))
gazette$Latitude  <- as.numeric(as.character(gazette$Latitude))
# Load place types
types <- read.xlsx(
  "../../Data/Zimbabwe/Locations/Zim_places_types_13_MAY_2018.xlsx")
types$Longitude <- as.numeric(as.character(types$Longitude))
types$Latitude  <- as.numeric(as.character(types$Latitude))
types2 <- types %>%
filter(!is.na(Latitude)) %>%
group_by(Latitude, Longitude) %>%
summarise(cat = paste0(unique(category), collapse = ','))
gazette <- merge(gazette, types2,
                 by=c('Longitude', 'Latitude'), all.x=T)
# Remove duplicates and rows with missing location
gazette <- subset(gazette, !is.na(gazette$Longitude)
           & duplicated(gazette$names)==FALSE)
           [,c("names", "cat", "Longitude", "Latitude")]
# Load the locations of the sites
sites.t           <- read.xlsx(
  "../../Data/Zimbabwe/Locations/TrialSites.xlsx", sheet=1)
sites.t           <- sites.t[,2:4]
# Merge the Lat / Long gazette from the site list and gazette
temp1 <- merge(temp1, sites.t, by.x="site_name", by.y="Site", all.x=T)
temp1 <- merge(temp1, gazette, by.x="citytown.v", by.y='names', all.x=T)
# Tidy up columns and names
temp1  <- temp1[c(vars, names(temp1[(length(temp1)-4):length(temp1)]))]
colnames(temp1) <- c(vars, 'lat0', 'lon0', "category", "lon1", "lat1")
names(temp1)[names(temp1)=="site_name"] <- "citytown"

# Add unique identifoer
temp1$uid       <- paste0(temp1$id, '_', temp1$round)
temp1$visit.num <- as.numeric(temp1$visit.num)
temp1$citytown.v[temp1$citytown.v=="."] <- ""
temp1$citytown.v[is.na(temp1$citytown.v) &
                  (temp1$worked.elsewhere==1 | temp1$stayed.elsewhere==1)] <- ""
temp1$citytown.v[temp1$worked.elsewhere!=1 & temp1$visit.num<2] <- NA
```

```r
temp1$citytown.v[temp1$stayed.elsewhere!=1  & temp1$visit.num>2] <- NA
temp1 <- temp1[order(temp1$round, temp1$id, temp1$visit.num),]

# Round the coordinates
r <- 5 #sets the precision of the coordinates
temp1$lat0 <- as.numeric(round(temp1$lat0, r))
temp1$lon0 <- as.numeric(round(temp1$lon0, r))
temp1$lat1 <- as.numeric(round(temp1$lat1, r))
temp1$lon1 <- as.numeric(round(temp1$lon1, r))

# Spatial-join the district and provinces of the sites
#and the places visited  (DIVA-GIS, Gadm)
zim1 <- readRDS(
  "../../Data/Zimbabwe/Maps/ZWE_adm1.rds")
zim2 <- readRDS(
  "../../Data/Zimbabwe/Maps/ZWE_adm2.rds")
ssa <- readOGR(
  dsn = "../../Data/Maps/gadm28_levels/", layer = 'ssa')

# Sites
temp2 <- SpatialPointsDataFrame(
  cbind(temp1$lon0, temp1$lat0),
  data = temp1, proj4string = CRS(proj4string(zim2)))
# Province of Zim
temp2$province <- sp::over(temp2, zim1)$NAME_1
# District of Zim
temp2$district <- sp::over(temp2, zim2)$NAME_2

# Visits
temp3 <- subset(temp1, !is.na(temp1$lat1) & !is.na(temp1$lon1)) %>%
SpatialPointsDataFrame(
  cbind(.$lon1, .$lat1),
  data = ., proj4string = CRS(proj4string(zim2)))
# Country
temp3$v.country <- sp::over(temp3, ssa)$NAME_ENGLI
# Province of Zim
temp3$v.province <- sp::over(temp3, zim1)$NAME_1
# District of Zim
temp3$v.district <- sp::over(temp3, zim2)$NAME_2

# Combine the sites and visits
```

```r
temp1 <- merge(temp2@data,
               temp3@data[,c('uid', 'round', 'visit.num',
                             'v.district', 'v.province', 'v.country')],
           by=c('uid', 'round', 'visit.num'), all.x=T)

# Change the locations of in other countries that were non-specific
# to the closest point from the site interviewed
# List of the countries named:
countries <-
as.data.frame(cbind(c('Botswana', 'Democratic Republic of the Congo', 'Malawi',
                      'Mozambique', 'Namibia','South Africa',
                      'Swaziland', 'Tanzania', 'Zambia'),
                  c(round(-20.541716,r), round(-12.1534304,r), round(-13.2543,r),
                    round(-16.9644355, r),round(-22.287214,r), round(-22.3813,r),
                    round(-26.5108178,r), round(-9.323102,r), round(-13.3136,r)),
                  c(round(27.725915,r), round(27.8197753,r), round(34.3015, r),
                    round(32.8672797, r), round(19.967102,r), round(30.0319,r),
                    round(30.34142,r), round(32.767857, r),round(28.086757, r))))
countries$V1 <- as.character(countries$V1)
countries$V2 <- as.numeric(as.character(countries$V2))
countries$V3 <- as.numeric(as.character(countries$V3))

# Loop over the countries
for (i in 1:nrow(countries)){
c <- ssa[ssa@data$NAME_ENGLI == countries[i,1],]
x <- temp1 %>%
   dplyr::filter(lat1 == countries[i,2] & lon1 == countries[i,3])  %>%
   SpatialPointsDataFrame(
     cbind(.$lon0, .$lat0),
     data = ., proj4string = CRS(proj4string(ssa)))

y <- gNearestPoints(x[1,], c)
  if (nrow(x@data)>1){
  for (j in 2:nrow(x@data)){
    y <- rbind(y, gNearestPoints(x[j,], c))
  }
  }
# Select only the points around the border
yy <- merge(as.data.frame(y),
            x@data, by.x = c('x', 'y'),
            by.y = c('lon0', 'lat0'), all = T) %>%
```

```r
    filter(is.na(id)) %>%
    dplyr::select(x, y) %>%
    cbind(., x@data$uid, x@data$visit.num)
colnames(yy) <- c('x', 'y', 'uid', 'visit.num')
# Change the lat lon coordinates for selected women
temp1 <- temp1 %>%
       merge(. , yy, by=c('uid', 'visit.num'), all.x = T) %>%
       mutate(lon1.2 = ifelse(lat1 == countries[i,2]
                              & lon1 == countries[i,3], x, lon1),
              lat1.2 = ifelse(lat1 == countries[i,2]
                              & lon1 == countries[i,3], y, lat1)) %>%
       mutate(lon1 = lon1.2, lat1 = lat1.2)

temp1$lat1.2 <- NULL
temp1$lon1.2 <- NULL
temp1$x      <- NULL
temp1$y      <- NULL
}

# Collect distances from Google maps, create a journey reference guide
temp1$start <- paste0(as.character(round(temp1$lat0, digits=5)),", ",
                      as.character(round(temp1$lon0, digits=5)))
temp1$end   <- paste0(as.character(round(temp1$lat1, digits=5)),", ",
                      as.character(round(temp1$lon1, digits=5)))
journeys    <- temp1[!is.na(temp1$start), c('uid', 'start', 'end')]

# For women who visited more than one place, calculate the distance between the places
temp2 <- reshape(data = temp1,
               direction  = 'wide',
               v.names    = c("lat1", "lon1"),
               timevar    = "visit.num",
               idvar      = "uid")
temp2 <- temp2 %>%
        dplyr::select(uid,
                      lat1.1.1, lon1.1.1,
                      lat1.1.2, lon1.1.2,
                      lat1.1.3, lon1.1.3,
                      lat1.1.4, lon1.1.4,
                      lat1.1.5, lon1.1.5,
                      lat1.2.1, lon1.2.1,
                      lat1.2.2, lon1.2.2,
```

```
                      lat1.2.3, lon1.2.3,
                      lat1.2.4, lon1.2.4,
                      lat1.2.5, lon1.2.5)

m <- matrix(c(1,2,3,4,2,3,4,5), ncol=2)

# Journeys for work
for (i in 1:nrow(m)){
lat.dest1 <- paste0('lat1.1.', m[i,1])
lon.dest1 <- paste0('lon1.1.', m[i,1])
lat.dest2 <- paste0('lat1.1.', m[i,2])
lon.dest2 <- paste0('lon1.1.', m[i,2])

temp3          <- temp2[c('uid', lat.dest1, lon.dest1, lat.dest2, lon.dest2)]
temp3          <- cbind(temp3, apply(temp3[,2:3], 1, conc.coords),
                        apply(temp3[,4:5], 1, conc.coords))
colnames(temp3) <- c('uid', lat.dest1, lon.dest1, lat.dest2, lon.dest2, 'start', 'end')
temp3          <- temp3[temp3$start!='NA,NA' & temp3$end!='NA,NA', c('uid', 'start', 'end')]
journeys       <- rbind(journeys, temp3)
}

# Journeys not for work
for (i in 1:nrow(m)){
lat.dest1 <- paste0('lat1.2.', m[i,1])
lon.dest1 <- paste0('lon1.2.', m[i,1])
lat.dest2 <- paste0('lat1.2.', m[i,2])
lon.dest2 <- paste0('lon1.2.', m[i,2])

temp3          <- temp2[c('uid', lat.dest1, lon.dest1, lat.dest2, lon.dest2)]
temp3          <- cbind(temp3, apply(temp3[,2:3], 1, conc.coords),
                        apply(temp3[,4:5], 1, conc.coords))
colnames(temp3) <- c('uid', lat.dest1, lon.dest1, lat.dest2, lon.dest2, 'start', 'end')
temp3          <- temp3[temp3$start!='NA,NA' & temp3$end!='NA,NA', c('uid', 'start', 'end')]
journeys       <- rbind(journeys, temp3)
}

# Remove duplicates
journeys <- journeys[journeys$start!='NA, NA' & journeys$end!='NA, NA' &
                     !duplicated(journeys[c('start', 'end')]), ]

# Loop through Google searches with system delay to avoid error on Google server
```

```r
journey.ref <- c()

for (i in 1:nrow(journeys)){
  x <- mapdist(journeys[i,"start"], journeys[i, "end"], mode = "driving")
  journey.ref <- rbind.fill(journey.ref, x)
  Sys.sleep(.1)
}

# Merge the journey distances with the main dataset
temp1 <- merge(temp1,
               journey.ref, by.x = c('start', 'end'),
               by.y = c('from', 'to'), all.x = T)
names(temp1)[names(temp1) == 'km'] <- 'dist.s2v'
names(temp1)[names(temp1) == 'hours'] <- 'time.s2v'
temp1[c('m', 'miles', 'seconds', 'minutes', 'hours')] <- NULL

# Merge the journey distances between places visited, in order
temp2  <- temp2[c('uid',
                'lat1.1.1', 'lon1.1.1',
                'lat1.1.2', 'lon1.1.2',
                'lat1.1.3', 'lon1.1.3',
                'lat1.1.4', 'lon1.1.4',
                'lat1.1.5', 'lon1.1.5',
                'lat1.2.1', 'lon1.2.1',
                'lat1.2.2', 'lon1.2.2',
                'lat1.2.3', 'lon1.2.3',
                'lat1.2.4', 'lon1.2.4',
                'lat1.2.5', 'lon1.2.5')]
# Create lat and lon columns in the journeys reference list
journey.ref$lat.start <- as.numeric(sub("\\, .*", "", journey.ref$from))
journey.ref$lon.start <- as.numeric(sub("^.*? ", "", journey.ref$from))
journey.ref$lat.end   <- as.numeric(sub("\\, .*", "", journey.ref$to))
journey.ref$lon.end   <- as.numeric(sub("^.*? ", "", journey.ref$to))

# Merge with the journeys between places in the order that they were named
for (i in 1:4){

# Places to work
temp2 <- merge(temp2,
               journey.ref,
               by.x = c(paste0('lat1.1.',i),
```

```
                        paste0('lon1.1.',i),
                        paste0('lat1.1.', i+1),
                        paste0('lon1.1.',i+1)),
               by.y = c('lat.start',
                        'lon.start',
                        'lat.end',
                        'lon.end'), all.x=T)
names(temp2)[names(temp2) == 'km']    <- paste0('dist.v2v1.', i)
names(temp2)[names(temp2) == 'hours'] <- paste0('time.v2v1.', i)
temp2[c('m', 'miles', 'seconds', 'minutes', 'hours', 'from', 'to')] <- NULL

# Places to stay
temp2 <- merge(temp2, journey.ref, by.x = c(paste0('lat1.2.',i),
                                            paste0('lon1.2.',i), paste0('lat1.2.', i+1),
                                            paste0('lon1.2.',i+1)),
               by.y = c('lat.start', 'lon.start', 'lat.end', 'lon.end'), all.x=T)
names(temp2)[names(temp2) == 'km'] <- paste0('dist.v2v2.', i)
names(temp2)[names(temp2) == 'hours'] <- paste0('time.v2v2.', i)
temp2[c('m', 'miles', 'seconds', 'minutes', 'hours', 'from', 'to')] <- NULL

}

temp2 <- temp2[c('uid', 'dist.v2v1.1', 'dist.v2v1.2','dist.v2v1.3','dist.v2v1.4',
                      'dist.v2v2.1', 'dist.v2v2.2','dist.v2v2.3','dist.v2v2.4',
                      'time.v2v1.1', 'time.v2v1.2','time.v2v1.3','time.v2v1.4',
                      'time.v2v2.1', 'time.v2v2.2','time.v2v2.3','time.v2v2.4')]
temp3 <- reshape(data = temp2,
               direction = 'long',
               varying   = c(list(c('dist.v2v1.1', 'dist.v2v1.2','dist.v2v1.3','dist.v2v1.4',
                             'dist.v2v2.1', 'dist.v2v2.2','dist.v2v2.3','dist.v2v2.4'),
                             c('time.v2v1.1', 'time.v2v1.2','time.v2v1.3','time.v2v1.4',
                             'time.v2v2.1', 'time.v2v2.2','time.v2v2.3','time.v2v2.4'))),
               v.names   = c('dist.v2v', 'time.v2v'),
               timevar   = "visit.num",
               idvar     = "uid")

# Add one to the visit number
temp3$v.num[temp3$visit.num<=4] <- paste0('1.', temp3$visit.num[temp3$visit.num<=4])
temp3$v.num[temp3$visit.num>4]  <- paste0('2.', temp3$visit.num[temp3$visit.num>4]-4)
temp3$visit.num <- as.numeric(temp3$v.num)
```

```r
# Merge with the primary dataset
temp1 <- merge(temp1, temp3, by = c('uid', 'visit.num'), all.x = T)


# Add worked as SW in interview location
temp1$sw.int <- 1


# Add worked as SW at the visit location - all 'yes' since that is the question
temp1$sw.v <- 1


# Add types of place dummies
temp1$towncity[!is.na(temp1$category)] <- 0
temp1$towncity[temp1$category == 'border town' |temp1$category == 'city'
           |temp1$category == 'town'|temp1$category == 'suburb'] <- 1
temp1$growthpoint[!is.na(temp1$category)] <- 0
temp1$growthpoint[temp1$category == 'growth point' |temp1$category == 'business centre'] <- 1
temp1$minefarm[!is.na(temp1$category)] <- 0
temp1$minefarm[temp1$category == 'mine' |temp1$category == 'farm'] <- 1
temp1$other[!is.na(temp1$category)] <- 0
temp1$other[!is.na(temp1$category)
            & temp1$towncity!=1
            & temp1$minefarm!=1
            & temp1$growthpoint!=1] <- 1
```