POPULATION, SPACE AND PLACE
Popul. Space Place (2010)
Published online in Wiley InterScience
(www.interscience.wiley.com) **DOI**: 10.1002/psp.614

Modelling Residential Mobility: Factors Associated with the Movement of Children in Greater Johannesburg, South Africa

Carren Ginsburg,[1,*] Fiona Steele,[2] Linda M. Richter[3] and Shane A. Norris[4]

[1]*Birth to Twenty Research Programme, Department of Paediatrics and Child Health, University of the Witwatersrand, Johannesburg, South Africa*
[2]*Centre for Multilevel Modelling, University of Bristol, United Kingdom*
[3]*Child, Youth and Family Development, Human Sciences Research Council, Durban; School of Psychology, University of KwaZulu-Natal, Durban; and Birth to Twenty Research Programme, Department of Paediatrics, University of the Witwatersrand, Johannesburg, South Africa*
[4]*MRC Mineral Metabolism Research Unit and Birth to Twenty Research Programme, Department of Paediatrics, University of the Witwatersrand, Johannesburg, South Africa*

ABSTRACT

Conceptualisation of child residential mobility has been influenced by developed country contexts; however, trends and models for movement are likely to differ in transitional societies. This paper uses event-history analysis to model the residential movements of South African urban children in the Birth to Twenty cohort over their first 14 years of life. Associations with mobility of children are tested over a set of domains relating to the child, the child's primary caregiver, and the child's household. A methodological approach is proposed for analysing repeated moves using multi-level models, which are adapted to maximise information from children who dropped out of the study or who had long gaps in their residential histories. The results indicate mobility is associated with economic disadvantage with children whose primary caregivers had no formal education and who lived in households with fewer assets and less access to services being more likely to change residence. The study suggests potential risks for mobile children in urban environments who may be more likely to be exposed to disruption or compromised living conditions. Copyright © 2010 John Wiley & Sons, Ltd.

Received 26 June 2009; revised 29 January 2010; accepted 13 February 2010

Keywords: residential mobility; internal migration; urban children; South Africa; event-history models

*Correspondence to: Carren Ginsburg, Birth to Twenty Research Programme, University of the Witwatersrand Medical School, 7 York Road, Parktown 2193, South Africa.
E-mail: carren.ginsburg@wits.ac.za

INTRODUCTION

The study of residential mobility among children has been recognised as important and has been researched in resource-rich countries, revealing dissimilar levels of mobility in different regions and environments (Long, 1992b). These empirical studies have been underpinned by a range of theoretical models for migration behaviour and processes. Early migration theorists, such as Lee (1966), posit that the drivers of migration are selective, with positive selection occurring where movement is motivated by opportunity and negative selection taking place where movement is prompted by necessity or disadvantage. Elaborating on this, more detailed frameworks for movement, such as that developed by De Jong (2000), propose a range of factors at the level of the individual, household, and community that impact on intentions to migrate, which, in turn, influence migration decision making. In the developed country context, where much of the conceptualisation of migration has occurred, movement of children has been assumed to be connected to a parent, and models describing mobility have therefore centred on adults. In low- and middle-income countries, residential mobility in children has been very under-researched; however, studies of children's living arrangements indicate that a proportion of children live independently of parents, with extended family members, for periods of varying duration (Hosegood *et al.*, 2007). In this setting, frameworks for describing mobility among children would need to be expanded to consider individual child characteristics, as well as the characteristics of parents, alternative caregivers, or family members who may be involved in movement decisions concerning children. Knowledge of these dynamics and attributes and how they might be associated with children's movement behaviour would yield significant insights into patterns of movement amongst children and, more importantly, lay the foundation for investigations into the impact of mobility on child health and development outcomes.

South Africa provides an interesting context in which to study internal population movements because of the shift within the country from politically controlled migration to movement based on choice, very often economically motivated. Further, the focus on children has become particularly relevant in the sub-Saharan region as a whole, in light of HIV prevalence and the potential effects of this on children's living arrangements and movements (Hosegood *et al.*, 2007). This paper presents the first longitudinal study of child residential mobility within urban South Africa and, in so doing, lays out methodological challenges to analysing children's movements.

South Africa's apartheid legacy significantly influenced internal population movement patterns through the implementation of policies, such as influx control and the Group Areas Act, which restricted permanent settlement of black[1] people within urban areas (Giliomee and Schlemmer, 1985). These regulations gave rise to a system of oscillatory labour migration between rural homes and urban places of employment where workers were accommodated temporarily (Wilson, 1972). High rates of urbanisation of black South Africans have been observed, immediately preceding and following South Africa's democratic transition in 1994. However, patterns of rural–urban circulation remain prominent, with evidence of increasing mobility within urban areas (South African Cities Network, 2004; Posel, 2006).

Movement to and within urban environments has the potential to render improved circumstances and conditions through better access to education, employment, health care, and social services. However, several studies have cautioned about the potential negative effects of urban living in large cities in low- and middle-income countries (Brockerhoff, 1995; UNICEF Innocenti Research Centre, 2002). In South Africa, for example, urban environments of large cities are associated with overcrowding, crime, unemployment, poverty, and susceptibility to disease (including increased mortality resulting from the AIDS pandemic).

Within these settings, children may be particularly at risk (UNICEF Innocenti Research Centre, 2002). These adverse circumstances may be exacerbated by unstable living arrangements and high residential mobility. For some children in South Africa, family life is characterised by residential insecurity, with frequent changes in household membership and child care arrangements (Murray, 1981; Spiegel *et al.*, 1996). South African children have been reported to move residence independently and/or in conjunction with a connected adult (Jones, 1992; Van der Waal, 1996). The factors prompting movement may be linked to circumstances surrounding the child's primary caregiver or family or to circumstances attached to the child directly. For example, children may move independently in response to the death of a caregiver or to access education by taking up residence with extended family (Ford and Hosegood, 2005; Kok and Collinson, 2006).

DOI: 10.1002/psp

Children may also accompany a caregiver in a move prompted by changes in a parental relationship or employment status (Kok and Collinson, 2006; Wentzel *et al.*, 2006).

In this new phase of South Africa's socio-political development, patterns of urban mobility are of particular interest; however, little research has focused specifically on the analysis of internal migration and movement trends (Kok *et al.*, 2003; Collinson *et al.*, 2006a). South African researchers have highlighted the need for focused, localised survey research that addresses questions concerning residential mobility, life course migration, the profile of mobile groups, and reasons for movement (Kok *et al.*, 2003; Kok and Collinson, 2006). The reason for the dearth of research in this area is primarily the lack of available cross-sectional and longitudinal data concerning internal population movements. While national household and labour surveys and more recent population censuses have incorporated questions investigating internal migration, limitations have been identified with the applicability of these data in analyses of movement patterns (Posel, 2002; Kok *et al.*, 2003; Kok and Collinson, 2006). Furthermore, very few studies have investigated the movement patterns of children, particularly amongst those born and living in urban environments (Ford and Hosegood, 2005). A significant reason for this research gap is the difficulty in measuring child mobility over time and the need for analytical techniques that take account of the complexity of the data. Data sets pose difficulties because of missing data due to permanent or temporary attrition. Dropout is of particular concern in studies of migration because attrition is closely related to the outcome of interest – mobility – leading to bias if children with missing data are excluded.

In response to the research needs highlighted, an analysis of residential mobility amongst urban children was undertaken using data collected from the Birth to Twenty cohort (BT20) study located in Johannesburg–Soweto, in the Gauteng province. Gauteng is South Africa's most densely populated urban centre, containing approximately 8.8 million residents (Statistics South Africa, 2006). The province is regarded as the economic hub of the country and is the largest receiver of migrants from other provinces (Kok *et al.*, 2003; Statistics South Africa, 2006). The Johannesburg–Soweto metropolis consists of an inner city, surrounded by informal settlements and suburban areas comprising formal housing. During the apartheid era, segregation legislation separated regions on the basis of race, resulting in socio-economic inequalities between areas. These restrictions were lifted following democratisation; nevertheless, disparities persist in infrastructure development and service provision in many areas. In 1990, on the eve of South Africa's transition to democracy, a group of Johannesburg–Soweto born children were recruited into a longitudinal birth cohort, BT20. The aim of the study was to track children's physical and social development in the context of rapid urbanisation and social change (Richter *et al.*, 2007). Regular data collection conducted among the cohort generated longitudinal data for a range of child health and development areas, including children's places of residence.

A preliminary analysis of the frequency of residential mobility within the BT20 cohort revealed that the majority of children (64%) had moved home at least once during their first 14 years, with the largest proportion of moves occurring within the Greater Johannesburg urban area (Ginsburg *et al.*, 2009). The principal aim of this study is to model the occurrence of initial and repeated residential mobility of children in the cohort so as to identify factors associated with movement, relating to the child, the child's primary caregiver, and the child's household. Multi-level event-history analysis is used to allow for repeated moves and to explore the effects of time-varying characteristics, such as household socio-economic status and attributes relating to children's current primary caregivers. The paper further proposes a novel approach to deal with permanent and temporary attrition which avoids the exclusion of dropouts, thereby maximising the analysis sample and reducing the potential for selection bias.

DATA

The BT20 Study Sample and Data Collection

The BT20 study was conceptualised and initiated by researchers from the University of the Witwatersrand and the South African Medical Research Council. The study sample was designed to include all singleton children born within a seven-week period between April and

DOI: 10.1002/psp

June 1990 at mainly public clinics and hospitals in the Greater Johannesburg metropolitan area situated in the Gauteng province. Of the total births that took place over the defined period, a sample of 3273 children identified as permanently resident in the area was recruited into the longitudinal birth cohort (Richter *et al.*, 2004). At enrollment, the cohort was demographically representative of the study area and comprised roughly equal numbers of male (48.6%) and female (51.4%) participants. The majority of participants were black (78.5%), with white, coloured, and Asian children comprising 6.3%, 11.7%, and 3.5% of the cohort, respectively. At the birth of their child, the majority of biological mothers were aged between 19 and 34 years (79.3%). Mothers were primarily single (56.5%), and most had not completed secondary school (58.4%).

Data collection activities among the cohort have taken place over a series of waves beginning with questionnaires administered antenatally to pregnant women and continuing at intervals of either one or two years. The study has focused on a set of core themes that include children's household environments, health and nutrition, growth and development, and risk behaviours (Richter *et al.*, 2007). Data collection has taken the form of physical and biological measures and questionnaires, administered to cohort children and their primary caregivers at health care centres and through home visits. Over the course of the study, contact has been maintained with approximately 70% of the original cohort, with an average of 14% of the sample lost to follow-up in any data collection wave (Norris *et al.*, 2007). During the study's 15th wave of data collection, a survey of children's residential movements was conducted. The questionnaire included a section in which all historical address records were verified as correctly reflecting the children's primary places of residence at the time. Missing or incomplete address data and additional data concerning reasons for movement were also collected. This Residential Move Questionnaire (RMQ) was completed by 2158 members of the original residential cohort (66%), with the balance of 1115 cohort members identified as cases of study attrition. A more detailed account of the BT20 data collection processes and the development and implementation of the specific study of residential movement within the cohort can be found in Richter *et al.* (2007) and Ginsburg *et al.* (2009).

Construction of Residential Histories

The analyses conducted in this paper are based on a longitudinal data set of children's residential addresses. These address data were used to construct a residential history for each child from which movements could be identified. Baseline address data reflecting the biological mother's place of residence immediately preceding the birth of the child was collected. Thereafter, residential address data were available for a series of nine intervals when children were aged between 0 and 1 year, 1 and 2 years, 2 and 4 years, 4 and 6 years, 6 and 8 years, 8 and 10 years, 10 and 12 years, 12 and 13 years, and 13 and 14 years. These intervals correspond to the BT20 study's data collection waves, where each cohort member was seen either annually or within a two-year period. The addresses reflect the BT20 child's primary place of residence during the interval. Residential addresses were updated in a database on each occasion that a cohort member was seen or contact attempted. The address data were later verified through the RMQ during the study's 15th year. Based on the address information, it was possible to derive a binary indicator of whether there had been a change in the child's main place of residence between age intervals t and $t - 1$. This variable is taken as the outcome in the analysis of residential mobility. Movement was defined in terms of the child and therefore refers to both independent moves or moves in combination with a primary caregiver or household.

Full movement histories (for each of the nine age intervals) were available for 99% of children whose residential details were confirmed in the RMQ. However, address data corresponding to a particular age interval or set of intervals may have been missing for children who were out of contact with the study at age 15 when the RMQ was administered. At each wave of data collection, a proportion of the cohort was identified as lost to follow-up for reasons such as caregiver or child mortality, study fatigue, or movement (see Norris *et al.*, 2007). In some instances, these losses to follow-up were classified as permanent (such as migrating out of the study area, emigrating to

DOI: 10.1002/psp

Europe, or child death), which meant that the child did not rejoin the study at a later stage. In other cases, non-participation was periodic, with children rejoining the study after a period of absence – for example, after spending some time living with a relative in a rural area. In instances where a child did not return to the study, the child's movement history was treated as censored from the interval corresponding to the first missing address and after which there was no further knowledge of the child's residential locations. In cases where a child left and then returned to the study, missing address data for a particular age interval may have been followed by one or more age intervals for which the residential address was recorded. A move was inferred in these cases through a comparison with the last recorded address, with the assumption that only one move had occurred if the address had changed. This approach allowed for the inclusion of all age intervals in which residential address information was available, with movement histories treated as censored at the last point at which information about children's residences was known. However, cases where address information was missing for seven or more consecutive years were excluded from the analyses on the basis that information was too sparse, and children may have experienced multiple moves within this time.[2]

The completeness of the residential histories of children included in the sample is presented in Table 1.

From the original cohort of 3273 children, 66 permanent dropouts were excluded from the analysis sample as these children were not present at any follow-up after the baseline and therefore no information was available on their residential moves. A further 230 children with address information missing for more than six years were excluded, resulting in a sample of 2977 children observed for 21,830 age intervals.

Explanatory Variables

The selection of explanatory variables considered in the event-history analysis was governed by theory and prior empirical findings. Potential predictors were conceptualised across three domains: attributes of the child, characteristics of the child's biological mother or current primary caregiver, and variables representing features of the child's current household (see Table 2).

The explanatory variables are a mixture of time-invariant and time-varying characteristics. While residential addresses and corresponding movement status was known at each data collection wave, other information was collected periodically. It is therefore necessary to assume that

Table 1. Completeness of residential histories in analysis sample.

Completeness of residential information	No. of children	Percentage
Complete history[1]	2155	65.84
Dropout after baseline	66	2.02
Permanent attrition[2]	306	9.35
Temporary dropout[3]	746	22.79
Total	3273	100
Max consecutive years missing for temporary dropouts		
1	–	–
2	85	11.39
3	87	11.66
4	268	35.92
5	7	0.94
6	69	9.25
>6	230	30.83
Total	746	100

[1] Children present at all nine waves with any gaps filled in from the RQM at age 15.
[2] Children with some mobility information who were permanently lost at any age.
[3] Children who were absent at one or more waves not compensated for by information collected in the RMQ.

DOI: 10.1002/psp

Table 2. Explanatory variables.

Variable	Description
Time invariant	
Child sex	Male, female
Child ethnicity	White, black, coloured, Asian
Residential area at birth	Soweto/Diepmeadow, former coloured/Asian, inner city, suburban
Hospital of birth	Public, private
Maternal age at delivery	<= 18, 19–34, 35+ years
Biological mother born in the Gauteng province	Yes, no
Time varying	
Age of child (t)	Intervals in years: [0,1], (1,2], (2,4], (4, 6], (6, 8], (8, 10], (10, 12], (12, 13], (13, 14]
Child moved previously[1]	Coded 0 prior to the first move and 1 after the first move
Duration since child's last move[1]	Length of time between moves in years
Caregiver schooling	No formal schooling, primary school, secondary school
Caregiver marital status	Married/living with partner, single/widowed/divorced/separated
Household socio-economic index	Constructed from the following binary items: Home ownership at birth (owned, other), house type (house, other), water indoors, flush toilet, electricity, TV, car, fridge, washing machine

[1]Only included in the multi-level models where repeated moves are analysed.

the time-varying covariates remained constant between those waves at which information was updated. Starting values of the time-varying variables were derived from questionnaire data collected antenatally or when children were aged 1 year. These values were held constant over the age intervals of [0,1], (1,2], and (2,4] years. Questionnaire data collected in the children's seventh year were then used to update the values of these variables over the three age ranges of (4, 6], (6, 8], and (8, 10] years. Values for the remaining intervals, (10,12], (12, 13], and (13,14] years, were based on information collected when the children were 12 or 13 years of age.

A socio-economic index was constructed from a set of 10 time-varying binary variables measuring access to a range of services and household assets: home ownership, house type, indoor water supply, indoor toilet, electricity supply, and telephone and household ownership of a television, motor vehicle, refrigerator, and washing machine. A probit factor model (see, for example, Bartholomew *et al.*, 2008: Chap. 8) was fitted using Markov chain Monte Carlo (MCMC) methods to incorporate children with missing data under a missing at random assumption (Browne, 2009); estimated factor loadings for each item were then used to compute a 'wealth index'. The socio-economic factor values ranged from −2.22 to 1.71, with a mean of 0.03 and a standard deviation of 0.82. A higher positive score on this index indicates greater relative wealth or household assets.

The multi-level analyses of repeated moves included a binary dummy variable indicating, for each age interval, whether a move had occurred in any previous age interval. Also included in the model was an interaction between the previous move indicator and the duration since the last move. This interaction variable was coded zero up to the first move, so its coefficient is interpreted as a duration effect among movers.

STATISTICAL METHODS

Multi-level discrete-time event-history analysis (e.g. Steele *et al.*, 1996) was used to model the timing of children's residential moves, allowing for the possibility that a child may be exposed to the risk of a move more than once over the observation period of 14 years. Residential histories can be viewed as a type of two-level hierarchical structure with episodes of exposure (periods between moves) at level 1 nested within individuals at level 2.

Denote by y_{it}, a binary response coded 1 if child i moves during age interval t and 0 otherwise. We assume that y_{it} follows a binomial distribution with probability π_{it} and denominator n_{it}, where, in the present application, n_{it} equals the length of interval t for child i. A multi-level logit model for the probability of a move, π_{it}, can be written as

$$\log\left(\frac{\pi_{it}}{1-\pi_{it}}\right) = \alpha_t + \boldsymbol{\beta}^T \boldsymbol{x}_{it} + u_i, \quad (1)$$

where α_t is the coefficient of a dummy variable for age interval t, $\boldsymbol{x}_{it}$ is a vector of time-varying and time-invariant characteristics of the child, caregiver, or household with coefficients $\boldsymbol{\beta}$, and u_i is a child-specific random effect assumed to be normally distributed with a mean of zero and variance of σ_u^2. The random effect represents unmeasured time-invariant child characteristics affecting the probability of a move throughout the study period. The child's residential history up to interval t is captured by an indicator of a previous move and the duration since the last move, both included as time-varying covariates in $\boldsymbol{x}_{it}$.

Equation (1) defines a proportional odds model where the effects of the covariates $\boldsymbol{x}_{it}$ are assumed to be constant across age intervals. Non-proportional effects may be accommodated by adding interactions between elements of $\boldsymbol{x}_{it}$ and the age dummies, but in our application, the proportionality assumption was found to be reasonable for all covariates.

When all time intervals are of equal width, the denominator for the binary response, n_{it}, equals 1 for all t and i, and Equation (1) can be estimated as a standard multi-level logit model for binary data. In the present application, however, age intervals vary in width. Children who were present at every wave contribute nine age intervals, where the width of an interval is either one or two years. As described in the Data section, children who dropped out permanently contribute one- or two-year intervals up to the point of being lost to follow-up. If a child temporarily left the study, the interval for the missing wave is combined with the interval for the wave at which the child rejoined the study, and n_{it} is updated to equal the width of the new interval. An adjustment to the coding of the dummy variables for the age intervals being aggregated is also needed. For example, consider a child who is absent at the age 2 interview but present at age 4. Age intervals (1,2] and (2,4] are combined to give a three-year interval, and the dummy variables for these intervals are each coded 0.5. In general, if k intervals are combined, the dummy variables for these intervals will each be coded $1/k$ regardless of the relative widths of the interval (see the Appendix for further details and an example of the required data structure).

The multi-level event-history model in Equation (1) is estimated using procedures for multilevel binomial response data (Steele *et al.*, 2004). We use MCMC methods as implemented in the MLwiN software (Browne, 2009; Rasbash *et al.*, 2009).[3]

To aid interpretation of the fitted model, predicted probabilities may be calculated for a range of values of each covariate (or each value in the case of categorical covariates), holding constant the values of all other covariates in the model. To obtain mean probabilities, it is necessary to average across child-specific unobservables by integrating out the random effect or by simulating random effect values. The simulation approach involves generating a large number of random effect values from a normal distribution with variance $\hat{\sigma}_u^2$, calculating a predicted probability based on each of these values and the estimated coefficients, and taking the mean across the simulated values. This procedure is implemented in MLwiN v2.10 and described in Rasbash *et al.* (2009).

RESULTS

An event-history analysis was conducted to examine the occurrence of the children's first residential move, with cases censored after the first move or at the last time interval when information regarding their movements was available (15,844 age intervals of 3146 children).

The conditional probability of the first residential move in age interval t, given no move occurred before t, is shown in Figure 1. The probability of a first move is highest between ages 1 and 2 years ($\hat{\pi} = 0.147$). By age 4, the probability of a first residential move decreased, with the lowest predicted probability of a first move in age interval (13, 14] ($\hat{\pi} = 0.046$).

While most children had experienced at most one residential move by age 14, 15% moved more

DOI: 10.1002/psp

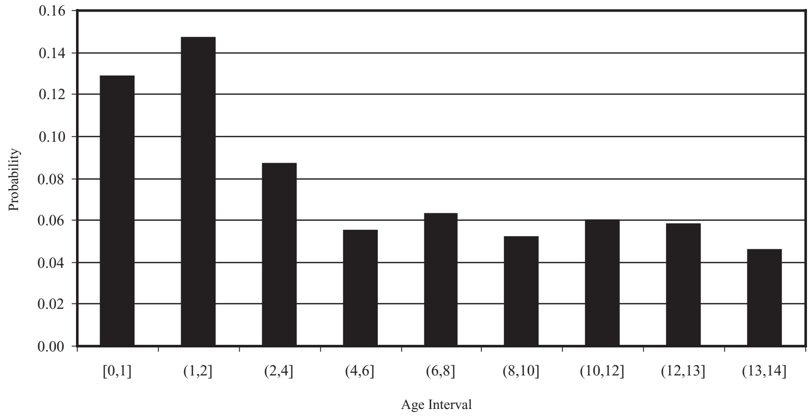

Figure 1. Probability of first move by age interval.

Table 3. Distribution of number of residential moves per child.

No. of moves	No. of children	Percentage
Never moved	1287	43.2
1	1245	41.8
2	368	12.4
3	61	2.0
4	14	0.5
5	2	0.1
Total	2977	100

than once (see Table 3). Multi-level event-history analysis was therefore used to consider repeated moves and to estimate the effect of previous mobility on the probability of a subsequent move.

The first model was based on the complete data set of 2977 children, observed for 21,830 age intervals. Due to a substantial number of missing values, the covariates *caregiver schooling* and *caregiver marital status* were initially excluded. A model was then fitted with these covariates included, where records were dropped from the data set after the first missing value. The reduced sample contains 2853 children who contributed 15,761 age intervals. Although the exclusion of these observations leads to the omission of some repeated moves (the percentage of children with more than one move decreases to 9.6%), the effects of the caregiver variables were of substantive interest because previous research has suggested that movements amongst children have been linked to caregiver characteristics and circumstances. Furthermore, a comparison between the models fitted to the full and reduced data sets revealed little difference in the magnitude and statistical significance of the regression coefficients of other covariates. The results presented in Table 4 are therefore based on the reduced data set with the inclusion of the two caregiver variables.

Other covariates were tested for significance using a combination of forward selection and backward elimination. The covariate *hospital of birth* was excluded from the analysis because it failed to achieve significance in any preliminary analyses. The variable *biological mother born in the Gauteng province* was also excluded from the models because of a substantial number of missing values together with non-significance at the 5% level. This variable had a negative effect on the probability of moving, indicating that

DOI: 10.1002/psp

Table 4. Parameter estimates (and standard errors) from the multi-level event-history model of residential mobility.

Variable	β	Standard error	Wald	p
Female child	0.060	0.047	1.658	0.198
Child ethnicity (ref: white)			3.051[1]	0.384
Black	−0.143	0.171	0.706	0.401
Coloured	−0.326	0.206	2.494	0.114
Asian	−0.146	0.227	0.415	0.519
Residential area at birth (ref: Soweto/Diepmeadow)			16.146[1]	0.001
Former coloured/Asian	−0.039	0.178	0.049	0.825
Inner city	0.658	0.202	10.625	0.001
Suburban	0.317	0.133	5.688	0.017
Maternal age at delivery (ref: <= 18)			41.347[1]	<0.001
19–34	0.197	0.078	6.31	0.012
35+	−0.379	0.119	10.101	0.001
Caregiver schooling (ref: no formal schooling)			8.853[1]	0.012
Primary school	−0.487	0.185	6.968	0.008
Secondary school	−0.534	0.180	8.783	0.003
Caregiver single/widowed/divorced/separated	−0.273	0.049	31.437	<0.001
Household socio-economic index	−0.162	0.031	26.738	<0.001
Child moved previously	0.590	0.083	50.599	<0.001
Child moved previously × duration since child's last move	−0.036	0.020	3.274	0.070
Age of child in years (t)				
Age [0, 1]	−1.375	0.281	23.961	<0.001
Age (1, 2]	−1.139	0.280	16.542	<0.001
Age (2, 4]	−1.781	0.280	40.457	<0.001
Age (4, 6]	−2.343	0.285	67.741	<0.001
Age (6, 8]	−2.257	0.284	63.179	<0.001
Age (8, 10]	−2.235	0.286	61.092	<0.001
Age (10, 12]	−2.246	0.293	58.842	<0.001
Age (12, 13]	−2.220	0.310	51.232	<0.001
Age (13, 14]	−2.311	0.312	54.813	<0.001
Child-level random effect variance	0.006	0.003		

[1] For categorical variables with more than two categories, the results of two types of Wald test are presented: (i) a joint test of the null hypothesis that the coefficients of the dummy variables for each category are simultaneously equal to zero and (ii) individual tests comparing each category with the reference.

children whose biological mothers were born in the province were less inclined to experience residential mobility. The interaction between child ethnicity and household socio-economic status was of interest because of the possible differences in the effect of socio-economic indicators on child movement between more and less advantaged ethnic groups; however, the term was found to be non-significant when tested, and was therefore excluded from the models.

Table 5 shows predicted probabilities of a move during age interval (1,2], the period when moves were most frequent. The probabilities were calculated by varying the values of one variable at a time, holding all other covariates at their sample mean values. In the case of a categorical variable, the dummy variable associated with a particular category takes on the value of the sample proportion in that category instead of the usual 0 or 1 value. The two variables associated with a previous move were fixed at a value of 0 so that probabilities refer to a first move (which is reasonable given the probabilities are calculated for ages 1–2 years). Although the probabilities will be different for other age intervals, their general pattern will be the same because the effects of covariates were found to be independent of age.

DOI: 10.1002/psp

Table 5. Predicted probabilities of a first move between 1 and 2 years of age.

Variable	Probability
Child sex	
Male	0.137
Female	0.144
Child ethnicity	
White	0.161
Black	0.143
Coloured	0.122
Asian	0.142
Residential area at birth	
Soweto/Diepmeadow	0.139
Former coloured/Asian	0.134
Inner city	0.238
Suburban	0.181
Maternal age at delivery (years)	
≤18	0.127
19–34	0.151
35+	0.091
Caregiver schooling	
No. formal schooling	0.216
Primary school	0.145
Secondary school	0.139
Caregiver marital status	
Married/living with partner	0.161
Single/widowed/divorced/separated	0.127
Household socio-economic index	
1 standard deviation above mean	0.125
Mean	0.141
1 standard deviation below mean	0.158

Of the child characteristics, sex and ethnic differences in the probability of a move were found not to be statistically significant. However, controlling for ethnic group, the effect of residential area at birth was found to be significant with a higher chance of moving among children born in the inner city and suburbs compared with those born in former Asian or coloured areas or in Soweto/Diepmeadow. The variables *child ethnicity* and *residential area at birth* are highly correlated due to the racial segregation of residential areas during the apartheid era; nevertheless, a significant effect of area that is independent of ethnic differences was found. Children who experienced a previous residential move were more likely to experience a repeated move as compared with non-movers, and there is some evidence (at the 10% level) that the probability of a move decreases with the duration since the last move.

Children born to older mothers (aged 35 or more) were less likely to move as compared with children born to younger mothers. Children whose biological mothers or primary caregivers were single, widowed, divorced, or separated rather than married or living with a partner were less likely to move. Similarly, a negative effect on the rate of residential movement was found amongst children whose primary caregivers had attained either primary or secondary level schooling as compared with caregivers with no formal education. The probability of a first residential move for children aged 1 and 2 years was highest for the group whose primary caregivers had no formal schooling.

The analysis revealed a significant negative relationship between household socio-economic status (as measured by the socio-economic factor values) and residential mobility. Holding

DOI: 10.1002/psp

household socio-economic status at its lowest level of −2.22, the probability of a first move for a child aged 1–2 years was 0.19, while the probability was 0.11 when household socio-economic status was held at its highest level of 1.71.

After controlling for child, caregiver, and household characteristics, there remains a small amount of unobserved heterogeneity between children ($\hat{\sigma}_u^2 = 0.006$, $SE = 0.003$). As expected, indicators of children's prior residential history – whether they had moved previously and the duration since the last move – explained a large proportion of the between-child variance; before accounting for these variables, the random effect variance was estimated as 0.198 ($SE = 0.054$).

DISCUSSION

This is the first South African study to explore longitudinal patterns of residential mobility amongst urban children. Using data from the BT20 cohort, children's residential movements over the first 14 years of their lives were analysed with the aim of identifying child, caregiver, and household factors associated with movement. The study looked both at the timing of children's first residential moves and at repeated residential mobility, with the conclusion that the more disadvantaged children in the cohort had a higher likelihood of experiencing residential change. Furthermore, standard event-history methods were adapted to handle permanent attrition and gaps in children's movement histories.

Knowledge of the patterns of child mobility in South Africa is scarce, and, consequently, it is important to develop a more detailed understanding of this area. Movement and its timing can have an important influence on future events and transitions in an individual's life course (Amoateng, 2007). In the current study, children's first residential moves occurred most frequently in early childhood (at age 2 years or younger). This finding is consistent with results from rural South African studies, in which movement was found to be highest amongst preschool children (Ford and Hosegood, 2005; Collinson *et al.*, 2006b). Similarly, findings from studies of residential mobility in developed countries have found relatively high levels of movement among one to four year olds, suggesting that the birth and early care of a child may prompt parents to move (Long, 1992a, 1992b). In the South African case, we hypothesise that the higher levels of first movement in very early childhood is reflective of changes in the life cycle of mothers who may be moving to access employment or to enter into cohabitating relationships. In addition, children may move more frequently in their preschool years, after which families attempt to stabilise children's status in the interest of minimising interruptions of schooling. Children who experience a first move early in life are more likely to experience repeated residential relocations during childhood. In the BT20 sample, 15% of the children had experienced repeated residential movement. Although we found that the statistical significance and effects of the covariates were the same regardless of whether children had experienced a single or multiple moves (results not shown), it is possible that the group of multiple movers may be more at risk of disrupted living conditions (by virtue of having shifted households more often).

The multi-level analysis of repeated moves revealed no significant gender differences, and ethnicity was not significantly associated with residential mobility. Previous empirical studies investigating inter-provincial migrations in South Africa have shown strong ethnic differences in the profile of migrants, with higher levels of movement amongst white and black South Africans and lower levels of movement amongst coloureds (Kok *et al.*, 2003). In the current study, a significant neighbourhood effect was present, with children born in Johannesburg's inner city or suburban regions more likely to change residence as compared with children in the areas of the city formerly designated as black or coloured/Asian. The finding is suggestive of a more integrated social geography in these regions following the dismantling of apartheid policy where residential areas were strictly segregated according to ethnic group membership. The higher levels of stability amongst those in the coloured and black township areas is potentially explained by the tendency amongst these communities towards extended family household structures (Amoateng *et al.*, 2007), suggesting that these families may be less inclined to move home.

At any particular time, children whose biological mothers or current primary caregivers were married or living with a partner were more likely to change residence as compared with

DOI: 10.1002/psp

children whose caregivers were single. Research has shown that partnership formation and breakdown are likely to result in residential mobility as part of changes in family cycles (Speare and Goldscheider, 1987; Long, 1992a). In addition, children living with parents or with a caregiver and her partner may be more geographically mobile because employment or accommodation options are increased by the presence of two adults as compared with one. Levels of movement may be lower amongst children being cared for by single women who could have fewer residential choices available to them and are potentially more likely to be living in extended family accommodation.

Mobility among children was found to be associated with lower levels of educational attainment of mothers or primary caregivers. South African studies of inter-provincial migration have found a correlation between labour migration (which may be prompted by instability or vagaries of circumstance) and lower levels of education, while higher levels of education have been associated with relocation linked to economic opportunity and options (Kok *et al.*, 2003; Wentzel *et al.*, 2006). The evidence of a link between intra-urban mobility and lower levels of education is suggestive of a group of children whose movements may be necessitated by limited employment or accommodation options for their mothers or caregivers. Similarly, the association between lower household socio-economic status and higher levels of mobility for children in the cohort further connects residential mobility to economic disadvantage. The negative relationship between household socio-economic status and movement suggests that negative selection, described by Lee (1966), is likely to be occurring within the cohort.

A strength of the current study is the focus on the movement of children. The data suggest an expansion of De Jong's (2000) model of migration decision making in that child individual characteristics, as well as the characteristics of a current primary caregiver and household, need to be taken into account in explaining movements involving children. This must allow for the fact that a child may not necessarily reside continuously with the same primary caregiver and in the same household. South African census data indicate that only 36.4% of black children aged between 5 and 13 years lived in a household together with both parents, 31.5% lived with a mother only, and 25.7% lived in households with neither parents. These rates are higher amongst coloured, white, and Asian children, where 58.7%, 80.0%, and 83.8%, respectively, lived in households with both parents (Statistics South Africa, 2001). In addition to evaluating the relative importance of child, caregiver, and household factors on movement decisions concerning children; the extent of children's broader care networks as well as context specific drivers and constraints would need to be incorporated into a framework explaining mobility in children.

The study contributes to the development of a broader understanding of the principle of migration selection in relation to child mobility by revealing a number of associations with child movement, which could be used to define a set of *a priori* hypotheses for future investigation within different sub-populations. For example, changes in the lifecycle of children's primary caregivers (such as partnership and employment status) may be associated with a higher probability of mobility for connected children. The study also lays the foundation for future research into impact studies. In order to begin to assess the consequences of relocation on child well-being, attributes of movement destinations at the neighbourhood or community level would be significant and may be analysed through the application of more complex multi-level modelling techniques.

Missing data and sample attrition encountered in a cohort study present a complex challenge, but an approach to handle children who drop out of the study and return at a later wave (non-monotone attrition) is proposed. This method of analysing all available data allows for higher levels of data retention than would have been the case if movement histories were censored at the first instance of lost contact. It thus includes into the analysis children who, often due to mobility, may not have been traceable over all data collection time points.

A limitation of the study relates to the potential underestimate in the total number of residential moves per child reported. It is likely that permanent dropouts and children with long gaps in their residential histories are more mobile and may have experienced multiple moves during their time out of the study, which would not have been known. A further limitation relates to lack

DOI: 10.1002/psp

of data concerning shifts in children's caregiving structures and its relationship to child and caregiver movements. Research is currently underway in BT20 to explore caregiving patterns longitudinally, and these will then be mapped onto children's movement trajectories.

In conclusion, the results of the analysis reveal a set of characteristics associated with residential mobility amongst a group of urban South African children that is suggestive of socio-economic disadvantage. For example, children experiencing residential mobility were more likely to have mothers or current primary caregivers with no formal education and reside in households with less access to assets and services. Moves were most likely to occur before the age of 2, and approximately 15% of all children studied experienced repeated moves during childhood. This group of mobile children may therefore have experienced economic adversities and lack of stability in living arrangements.

The study addresses a critical gap in children's developmental research in South Africa and suggests the need for comparative research on child mobility, both in rural regions of South Africa and in other low- and middle-income countries. Insight into the drivers and processes around child mobility within different contexts would contribute to current frameworks describing movement among adults, and thus fill an important research gap. Given that movement may be one response to disadvantage, understanding the consequences of mobility for children is a key priority. Knowledge of the impact of movement on children's adjustment, physical health and education would significantly inform local policy initiatives centred on vulnerable children.

ACKNOWLEDGEMENTS

Studies of children's migration in Birth to Twenty are funded by the Wellcome Trust (UK), the Human Sciences Research Council (South Africa), The Medical Research Council (South Africa), and a grant from the Mellon Foundation to the University of the Witwatersrand (South Africa). Carren Ginsburg has also received funding from the South African Department of Science and Technology (DST) and the National Research Foundation (NRF). The assistance of Saul Margolis with the database is gratefully acknowledged.

NOTES

(1) The categories black, coloured (mixed ancestral origin), white, and Asian are carried over from South Africa's apartheid past. While they no longer have legislative force, they have so influenced South African society, and in many ways continue to do so, that there is consensus on the importance of retaining these categories for social analyses. In this paper, we have used these racial categories in our analyses as opposed to ethnic categories (for example, Zulu, Xhosa, Sotho, and Afrikaans).

(2) The analysis was repeated with a different exclusion rule for children with gaps in their residential histories. The results were found to be robust to whether the cutoff was more than four, five, or six consecutive years.

(3) MCMC methods are used to estimate statistical models in a Bayesian framework. In the Bayesian approach, each unknown parameter in the model is viewed as a random variable with an associated probability distribution that incorporates any prior beliefs about the value of that parameter. MCMC methods are simulation-based procedures in which a chain of random draws is taken from the current conditional probability distribution for each parameter. A point estimate of a parameter may be obtained by taking the mean, median, or mode of the parameter values across the chains, while the standard deviation of parameter values corresponds to a frequentist standard error. See Browne (2009) for an introduction to MCMC methods for multi-level analysis. The estimates presented in this paper are from 50,000 chains using approximate quasi-likelihood estimates (Goldstein, 2003: 112–113) as starting values for the sampling.

REFERENCES

Amoateng AY. 2007. Towards a conceptual framework for families and households. In *Families and Households in Post-Apartheid South Africa: Socio-Demographic Perspectives*, Amoateng AY, Heaton TB (eds). HSRC Press: Cape Town; 27–42.

Amoateng AY, Heaton TB, Kalule-Sabiti I. 2007. Living arrangements in South Africa. In *Families and Households in Post-Apartheid South Africa: Socio-Demographic Perspectives*, Amoateng AY, Heaton TB (eds). HSRC Press: Cape Town; 43–59.

Bartholomew DJ, Steele F, Moustaki I, Galbraith JI. 2008. *Analysis of Multivariate Social Science Data*. 2nd ed. Chapman & Hall/CRC: Boca Raton, FL.

Brockerhoff M. 1995. Child survival in big cities: the disadvantages of migrants. *Social Science and Medicine* **40**: 1371–1383.

DOI: 10.1002/psp

Browne WJ. 2009. *MCMC Estimation in MLwiN v2.10*. Centre for Multilevel Modelling, University of Bristol: Bristol.

Collinson M, Kok P, Garenne M. 2006a. *Migration and Changing Settlement Patterns: Multilevel Data for Policy*. Report 03-04-01, Statistics South Africa: Pretoria.

Collinson M, Tollman S, Kahn K, Clark S, Garenne M. 2006b. Highly prevalent circular migration: households, mobility and economic status in rural South Africa. In *Africa on the Move: African Migration and Urbanisation in Comparative Perspective*, Tienda M, Findley S, Tollman S, Preston-Whyte E (eds). Wits University Press; Johannesburg; 194–216.

De Jong GF. 2000. Expectations, gender, and norms in migration decision-making. *Population Studies* **54**: 307–319. DOI: 10.1080/713779089.

Ford K, Hosegood V. 2005. AIDS mortality and the mobility of children in Kwazulu Natal, South Africa. *Demography* **42**(4): 757–768.

Giliomee H, Schlemmer L. 1985. *Up Against the Fences: Poverty, Passes and Privilege in South Africa*. David Philip: Cape Town.

Ginsburg C, Norris SA, Richter LM, Coplan DB. 2009. Patterns of residential mobility amongst children in Greater Johannesburg-Soweto, South Africa: observations from the Birth to Twenty cohort. *Urban Forum* **20**(4): 397–413. DOI: 10.1007/s12132-009-9069-6.

Goldstein H. 2003. *Multilevel Statistical Models*. 3rd ed. Edward Arnold: London.

Hosegood V, Floyd S, Marston M, Hill C, McGrath N, Isingo R, Crampin A, Zaba B. 2007. The effects of high HIV prevalence on orphanhood and living arrangements of children in Malawi, Tanzania, and South Africa. *Population Studies* **61**(3): 327–336. DOI: 10.1080/00324720701524292.

Jones S. 1992. Children on the move: parenting, mobility and birth-status among migrants. In *Questionable Issue: Illegitimacy in South Africa*, Burman S, Preston-Whyte EM (eds). Oxford University Press: Cape Town; 247–281.

Kok P, O'Donovan M, Bouare O, van Zyl J. 2003. *Post-Apartheid Patterns of Internal Migration in South Africa*. HSRC Press: Cape Town.

Kok P, Collinson M. 2006. *Migration and Urbanisation in South Africa*. Report 03-04-02, Statistics South Africa: Pretoria.

Lee ES. 1966. A theory of migration. *Demography* **3**(1); 47–57.

Long L. 1992a. Changing residence: comparative perspectives on its relationship to age, sex, and marital status. *Population Studies* **46**: 141–158.

Long L. 1992b. International perspectives on the residential mobility of America's children. *Journal of Marriage and the Family* **54**(4): 861–869.

Murray C. 1981. *Families Divided: The Impact of Migrant Labour in Lesotho*. Ravan: Johannesburg.

Norris SA, Richter LM, Fleetwood SA. 2007. Panel studies in developing countries: case analysis of sample attrition over the past 16 years within the Birth to Twenty cohort in Johannesburg, South Africa. *Journal of International Development* **19**: 1143–1150.

Posel D. 2002. *Moving out of the Household and out of the Household Questionnaire? The Coverage of Labour Migration in National Surveys in South Africa (1993–2001)*. Presented at the Development Policy Research Unit's Second Annual Conference on Labour Markets and Poverty in South Africa. September 2002 (available at http://www.commerce.uct.ac.za/ResearchUnits/DPRU/Conf2003PDF/Conf2002pdf/DorritPosel.pdf).

Posel D. 2006. Moving on: patterns of labour migration in post-Apartheid South Africa. In *Africa on the Move: African Migration and Urbanisation in Comparative Perspective*, Tienda M, Findley S, Tollman S, Preston-Whyte E (eds). Wits University Press: Johannesburg; 217–231.

Rasbash J, Charlton C, Browne WJ, Healy M, Cameron B. 2009. *MLwiN Version 2.10*. Centre for Multilevel Modelling, University of Bristol: Bristol.

Richter LM, Norris SA, De Wet T. 2004. Transition from Birth to Ten to Birth to Twenty: the South African cohort reaches 13 years of age. *Paediatric and Perinatal Epidemiology* **18**(4): 290–301.

Richter LM, Norris SA, Pettifor J, Yach D, Cameron N. 2007. Cohort profile, Mandela's Children: the 1990 Birth to Twenty study in South Africa. *International Journal of Epidemiology* **36**: 504–511.

Spiegel A, Watson V, Wilkinson P. 1996. Domestic diversity and fluidity among some African households in Greater Cape Town. *Social Dynamics* **22**(1): 7–30.

Speare A, Goldscheider FK. 1987. Effects of marital status change on residential mobility. *Journal of Marriage and the Family* **49**(2): 455–464.

South African Cities Network. 2004. *State of the Cities Report*. South African Cities Network: Johannesburg.

Statistics South Africa. 2001. *Census 2001: Stages in the Life Cycle of South Africans*. Report 03-02-46, Statistics South Africa: Pretoria.

Statistics South Africa. 2006. *Provincial profile 2004: Gauteng*. Report 00-91-07, Statistics South Africa: Pretoria.

Steele F, Diamond I, Amin S. 1996. Immunization uptake in rural Bangladesh: a multilevel analysis. *Journal of the Royal Statistical Society. Series A-Statistics in Society* **159**(2): 289–299.

Steele F, Goldstein H, Browne WJ. 2004. A general multilevel multistate competing risks model for event history data, with an application to a study of contraceptive use dynamics. *Statistical Modelling* **4**(2): 145–159.

UNICEF Innocenti Research Centre. 2002. *Poverty and Exclusion Among Urban Children*. Innocenti Digest No 10, UNICEF Innocenti Research Centre: Florence.

Van der Waal CS. 1996. Rural children and residential instability in the Northern Province of South Africa. *Social Dynamics* **22**(1): 31–54.

Wentzel M, Viljoen J, Kok P. 2006. Contemporary South African migration patterns and intentions. In *Migration in South and Southern Africa: Dynamics and Determinants*, Kok P, Gelderblom D, Oucho JO, van Zyl J (eds). HSRC Press: Cape Town; 171–204.

Wilson F. 1972. *Labour in the South African Gold Mines 1911–1969*. Cambridge University Press: Cambridge.

APPENDIX: TREATMENT OF RESIDENTIAL HISTORIES FOR TEMPORARY DROPOUTS

Table A1 shows the data structure for two children, where $D_1, D_2, \ldots, D_9$ are the dummy variables for age interval t with coefficients given by α_t in Equation (1). Complete information is available for Child 1, who contributes records for the full set of nine age intervals. Child 2 is a combination of a permanent and temporary dropout, being absent at the age 2 interview, then present at age 4, before being lost to follow-up after the age 6 interview. For this child, age intervals (1,2] and (2,4] are combined to give a three-year interval. The values of the dummy variables for these intervals are also changed from the usual (0,1) coding to reflect the fact that the second interval is now an aggregate of intervals $t = 2$ and $t = 3$. Specifically, the dummies for intervals (1,2] and (2,4] are each coded 0.5.

This coding of the dummies for age is based on the following approximation. Consider a simplified specification of the model in Equation (1) with only age effects and no child-specific random effects. Omitting child subscripts, the model can be written as

$$\text{logit}(\pi_t) = \alpha_t. \tag{A1}$$

Combining age intervals t and $t + 1$, the probability of a move in the joint interval is

$$\pi = \pi_t + \pi_{t+1}. \tag{A2}$$

When the probability of a move is small within each interval t, the logit transformation is well approximated by the log transformation, so that

$$\log(\pi_t) \approx \alpha_t. \tag{A3}$$

Exponentiating Equation (A3) and substituting in Equation (A2) gives

$$\pi \approx \exp(\alpha_t) + \exp(\alpha_{t+1}). \tag{A4}$$

We next carry out a Taylor series expansion of $\exp(\alpha_{t+1})$ around α_t:

$$\exp(\alpha_{t+1}) = \exp(\alpha_t) + (\alpha_{t+1} - \alpha_t)\exp(\alpha_t) + O(\alpha^2), \tag{A5}$$

Table A1. Example of data structure for complete and partial residential histories.

Child i	Interval t	n_{it}	D_1	D_2	D_3	D_4	D_5	D_6	D_7	D_8	D_9
1	[0,1]	1	1	0	0	0	0	0	0	0	0
1	(1,2]	1	0	1	0	0	0	0	0	0	0
1	(2,4]	2	0	0	1	0	0	0	0	0	0
1	(4,6]	2	0	0	0	1	0	0	0	0	0
1	(6,8]	2	0	0	0	0	1	0	0	0	0
1	(8,10]	2	0	0	0	0	0	1	0	0	0
1	(10,12]	2	0	0	0	0	0	0	1	0	0
1	(12,13]	1	0	0	0	0	0	0	0	1	0
1	(13,14]	1	0	0	0	0	0	0	0	0	1
2	[0,1]	1	1	0	0	0	0	0	0	0	0
2	(1,4][1]	3	0	0.5	0.5	0	0	0	0	0	0
2	(4,6]	2	0	0	0	1	0	0	0	0	0

[1] Combined interval.

DOI: 10.1002/psp

where $\alpha = \alpha_{t+1} - \alpha_t$. Substituting Equation (A5) in Equation (A4) leads to

$$\begin{aligned}\pi &\approx \exp(\alpha_t) + \{\exp(\alpha_t) + (\alpha_{t+1} - \alpha_t)\exp(\alpha_t) + O(\alpha^2)\}\\ &= \exp(\alpha_t)\{2 + (\alpha_{t+1} - \alpha_t) + O(\alpha^2)\}\\ &= 2\exp(\alpha_t)\{1 + (\alpha_{t+1} - \alpha_t)/2 + O(\alpha^2)\}\end{aligned} \tag{A6}$$

Using the first-order McLaurin series expansion $\exp(z) = 1 + z + O(z^2)$, we can write

$$\exp\{(\alpha_{t+1} - \alpha_t)/2\} = 1 + (\alpha_{t+1} - \alpha_t)/2 + O(\alpha^2). \tag{A7}$$

Finally, substituting Equation (A7) in the last line of Equation (A6) and assuming that the difference in the log probability between intervals t and $t + 1$ is small (so that $O(\alpha^2) \to 0$), we have the following first-order approximation for the probability of an event in the joint interval:

$$\begin{aligned}\pi &\approx 2\exp(\alpha_t)\exp\{(\alpha_{t+1} - \alpha_t)/2\}\\ &= 2\exp(0.5\alpha_t + 0.5\alpha_{t+1}),\end{aligned} \tag{A8}$$

which can be written in log-linear form as

$$\log(\pi) \approx \log(2) + 0.5\alpha_t + 0.5\alpha_{t+1}. \tag{A9}$$

The log probability implied by Equation (A9) is fitted by including log(2) as an offset term and coding the dummy variables for intervals t and $t + 1$ as 0.5 and the dummies for all other intervals as 0. Reverting to the original logit scale, the width of the joint interval ($n_t = 2$) is included as a denominator for the binary response. Note that the approximation in Equation (A9) holds for combining any two intervals regardless of their width.