

# Administrative Data Quality: Investigating Record-Level Address Accuracy in the Northern Ireland Health Register

*Brian Foley<sup>1,2</sup>, Ian Shuttleworth<sup>1</sup>, and David Martin<sup>3</sup>*

Many national statistical institutes (NSIs) are seeking to supplement or replace their traditional population census with a methodology underpinned by administrative sources. Health service register data are key in this regard owing to their high population coverage; it is therefore important to improve understanding of data quality in this administrative source. This study investigated the factors associated with record-level address data mismatch between the Northern Ireland (NI) Health Card Registration System (HCRS) and the 2011 Census, using the NI Longitudinal Study (NILS). Address information in the form of anonymised Unique Property Reference Number (XUPRN) was available for circa 334,000 NILS members with census returns in 2001 and 2011, which provided a benchmark to assess XUPRN accuracy in their linked HCRS record for comparable time points. Multinomial logistic regression revealed a significantly greater likelihood of address mismatch in the HCRS for: males; young adults; individuals with no limiting long-term illness; migrants in the year prior to each census; and residents of communal establishments. Identification of population groups affected by poor quality address information in administrative sources can assist NSIs with the development and implementation of methodological improvements to ensure that official population statistics generated from these sources are fit for purpose.

*Key words:* Address data quality; census; population statistics; longitudinal data.

## 1. Introduction

The use of administrative data sources in official statistical systems is well established (Eurostat 2003; Karr 2012; Wallgren and Wallgren 2014; Agafitei et al. 2015). Many national statistical institutes (NSIs) draw upon these data, routinely collected by government departments, state agencies and other organisations via the operation of a service, transaction or registration to inform statistics on the economy and society. Challenges facing NSIs including financial constraints and the high cost of data collection via sample surveys (United Nations Economic Commission for Europe (UNECE) 2011)

<sup>1</sup> School of Natural and Built Environment, Queens University Belfast, Belfast, BT7 1NN, UK. Emails: [brian.foley@nirsra.gov.uk](mailto:brian.foley@nirsra.gov.uk) and [i.shuttleworth@qub.ac.uk](mailto:i.shuttleworth@qub.ac.uk)

<sup>2</sup> Northern Ireland Statistics and Research Agency, Belfast, BT9 5BF, UK.

<sup>3</sup> Geography and Environment, University of Southampton, Southampton, SO17 1BJ, UK Email: [d.j.martin@soton.ac.uk](mailto:d.j.martin@soton.ac.uk)

**Acknowledgments:** Author B.F. is supported by funding from the Economic and Social Research Council (ESRC) for the Administrative Data Research Centre in Northern Ireland. The help provided by the staff of the Northern Ireland Longitudinal Study (NILS) and the NILS Research Support Unit is acknowledged. The NILS is funded by the Health and Social Care Research and Development Division of the Public Health Agency (HSC R&D Division) and NISRA. The NILS-RSU is funded by the ESRC and the Northern Ireland Government. The authors alone are responsible for the interpretation of the data and any views or opinions presented are solely those of the author and do not necessarily represent those of NISRA/NILS.

are resulting in the increasing integration of administrative sources in statistical production processes.

Among the range of administrative sources in existence, one of the most informative in terms of providing basic demographic and address information is a health service register, referred to hereafter as a health register. These data are based on individual registrations to access primary care services, mainly via general practitioners. Such services are free at the point of use in many countries, for example the United Kingdom (UK), Italy and Denmark (Roland et al. 2012; Lo Scalzo et al. 2009; Pedersen et al. 2012), which is a strong incentive to register, hence the high population coverage generally provided by these data. Accordingly, health registers have utility from a statistical perspective, with many NSIs using these data in their methodology for producing annual subnational population estimates. Furthermore, a number of NSIs that conduct a traditional population census are investigating the potential of administrative sources to supplement or replace this approach; health registers are likely to be key in this regard given their value in a statistical context.

It is therefore important to improve understanding of quality issues in this administrative source given its use in producing population statistics; of specific interest for this study is record-level address data. Among various applications, these data inform the estimation of internal migration flows by some NSIs and they are one of the main matching variables for record linkage between administrative sources integrated in population statistics production processes. Our focus is the health register in Northern Ireland (NI), which is termed the Health Card Registration System (HCRS). Previous studies have investigated the quality of address information in this administrative source (Shuttleworth and Barr 2011; Barr and Shuttleworth 2012; Shuttleworth and Martin 2016) and the England and Wales equivalent, the Patient Register (Smallwood and Lynch 2010), using 2001 Census data as the reference. This study builds upon the existing research by investigating address data quality from both a cross-sectional and longitudinal perspective using data from the 2001 and 2011 censuses of NI. The first aim was to investigate the extent of address mismatch between the HCRS and census in 2011 and to identify some of the associated characteristics using a univariate approach. Secondly, combining results from the 2001- and 2011-based analyses, the aim was to identify the individual-, household-, and area-level factors associated with address mismatch from a longitudinal perspective, employing a multivariate methodology. These aims were met by the use of the NI Longitudinal Study (NILS), one of three UK census-based longitudinal studies (along with the Office for National Statistics (ONS) and Scottish longitudinal studies), but unique in its high sample fraction.

This article is presented in five sections. In Section 2, the use of administrative data to inform the production of population statistics is outlined in the context of the different methodologies employed by NSIs internationally; this is followed by specific reference to health register data and associated quality issues. A detailed description of the NILS is provided in section three, highlighting its value in facilitating a detailed assessment of address data quality in the NI HCRS. We also outline the methods underlying the analysis. The results in Section 4 reveal the factors associated with address mismatch in the HCRS in a cross-sectional and longitudinal context. In Section 5, the findings are interpreted from various perspectives including the wider implications for the statistical application of

health register data before, finally, the conclusions from the study are presented in Section 6.

## 2. Administrative Data and Population Statistics: National and International Perspectives

A census is fundamental to enumerating a country's population and producing statistics on its main demographic and socioeconomic characteristics from national to small area level. Most countries employ a traditional approach when conducting a census and while the principle of collecting information from all individuals and households remains, the methodology has evolved. This includes the introduction of online data collection in recent census rounds (Moore et al. 2008). The internet is becoming the primary mode of data collection in many of the traditional census-taking countries; the 2016 Canadian Census had an internet collection response rate of 68% (UNECE 2016), the target for the 2021 UK Census is a 65% online response rate (UNECE 2015), while the United States Census Bureau (USCB) are planning for the internet to be the main response mode in their 2020 Census (USCB 2015a). One of the main alternative census methods is a register-based system underpinned by administrative sources, which is well-established in many of the Nordic states (Martin 2006; UNECE 2007; Lange 2014). Owing to the advantages of using administrative data such as reduced respondent burden and the capacity for more frequent statistical outputs, other countries are turning increasingly to these sources. For example, Sweden and Austria made the transition to a completely register-based census in 2011 (Andersson et al. 2013; Kukutai et al. 2015), while Switzerland and Germany introduced integrated census systems combining administrative registers and sample surveys in 2010 and 2011, respectively (UNECE 2012a, 2012b). Although the UK conducts a traditional census, various administrative sources were used to quality assure population estimates from their 2011 event (National Records of Scotland (NRS) 2012; Northern Ireland Statistics and Research Agency (NISRA) 2012; ONS 2012a). Looking ahead, the ONS has committed to greater use of administrative data to enhance the quality of statistics from the 2021 Census of England and Wales (ONS 2014a); this approach has also been adopted by the other UK NSIs in Scotland (NRS 2014) and NI (NISRA 2014). It is clear therefore that administrative data are becoming increasingly relevant internationally in the context of the population census.

While the traditional census provides accurate data for a point in time, there is a requirement for high quality population statistics throughout the intercensal period. Many NSIs in the countries concerned use administrative sources to produce annual population statistics. Internal migration, the movement of individuals within a country, greatly influences the size and composition of populations at subnational level and its estimation is underpinned by administrative data in the UK (ONS 2016), Canada (Statistics Canada (SC) 2015), United States (USCB 2016), and Australia (ABS 2014); health register data are one of the main sources drawn upon in this context. Furthermore, these data are used in the methodology for estimating the subnational distribution of immigrants throughout England and Wales (ONS 2011). Looking forward, the NSIs of New Zealand and England and Wales are assessing the feasibility of replacing their traditional census with a model based largely on linked administrative sources that is less costly and can produce more

frequent population statistics (Bycroft 2015; ONS 2015a). In this regard, the ONS are releasing experimental population estimates at national and subnational level on an annual basis, which are based on linked administrative data sources including the Patient Register (ONS 2015b).

Like many administrative sources, health register data are affected by particular quality issues. Overcoverage is an acknowledged feature for reasons such as emigrants not de-registering prior to moving abroad and multiple area registrations; ‘list inflation’ of between four per cent and five per cent above the population estimate has been reported for the Patient Register in England and Wales (ONS 2012b) and the HCRS in NI (O’Reilly et al. 2012). Another issue concerns the lag in updating of residential moves, which is especially prevalent among males and in urban and deprived areas (Shuttleworth and Barr 2011; Statistics New Zealand 2013). From a temporal perspective, Barr and Shuttleworth (2012) found that 44% of an internal migrant study cohort in the NILS lagged in reporting a change of address to the HCRS by more than one year, while in a London-based study by Millett et al. (2005), seven per cent of participants took longer than three years to re-register with a general practitioner after a change of address.

There is a need to build upon the existing evidence base on address data quality in administrative sources such as a national health register. To provide context for this study, Figure 1 outlines, conceptually, major contributors to inaccurate address information in administrative data, drawing on the illustrative approach used by Raymer et al. (2015). The delayed or failed reporting of a residential move to the relevant administrative system is considered one of the main statistical challenges associated with the use of administrative data to produce population statistics (ONS 2012c).

The NILS facilitated a novel approach for this study, providing address information for a large sample from consecutive censuses and for regular intervals from the HCRS over a thirteen-year period to conduct a detailed longitudinal analysis. Although NI is the smallest country in the UK, its varied settlement pattern with a large urban centre (Belfast) and rural countryside interspersed with towns is similar to that observed in many other

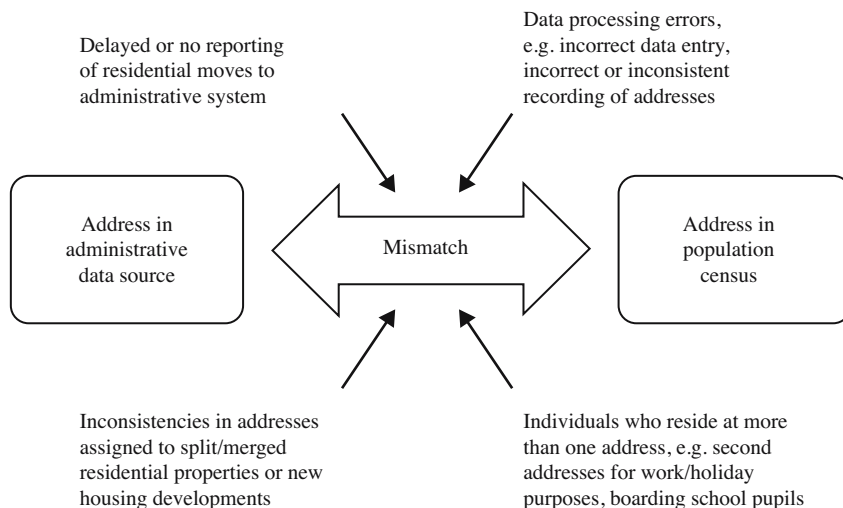


Fig. 1. Overview of contributors to address inaccuracy in an administrative data source.

countries, which broadens the applicability of the findings. Our study may be of interest to organisations and NSIs that use health register data for statistical or research purposes as it provides an insight on record-level inaccuracies in geographical referencing and the associated factors.

### 3. Data and Methods

#### 3.1. The NILS

The analysis was based on the NILS, which is a representative circa 28% sample of the 1.84 million population of NI. The spine of the NILS is records from the NI HCRS, which is administered by the Health and Social Care Business Services Organisation (BSO) and includes almost 100% of the NI population ([O'Reilly et al. 2012](#)). The HCRS data are obtained from the National Health Authority Information Registration System, which is linked to all general practitioner practices in NI. Although each person registered to access primary care services is assigned a unique Health and Care Number, duplicate registrations are identified as one of the factors contributing to overcoverage in the HCRS; however, quality assurance within NISRA sees the removal of duplicate records prior to the use of the data for statistical and research purposes ([NISRA 2016a](#)). Once an individual is registered on the HCRS, they can voluntarily update their address information via their general practitioner practice or the BSO website ([ONS 2017](#)).

Membership of the NILS is based on having one of 104 predesignated birth dates, with records linked to decennial census returns from 1981 to 2011. There is no complete household structure in the NILS as it is a sample, unlike census data, with some NILS members being a single representative from a household and sometimes multiple members from the same household. The linkage of HCRS and census records in the NILS employs a sequential match-key approach; this methodology enables records for the same individual to be linked where their address differs across both sources by assigning highest matching weights to name, gender and date of birth and using part of the address information in one of the match-keys. Manual verification of linked records and thorough quality checks ensure a high match rate. Further information on the NILS record linkage processing is available in a data matching methodology working paper ([Northern Ireland Longitudinal Study 2015](#)). Census data provide a rich source of information on demographic and socioeconomic characteristics of NILS members at individual- and household-level. In addition, residential moves of NILS members from 2001 onwards can be determined from routine HCRS updates provided to the NILS at six-month intervals. Note that the HCRS and census are separate data sources with neither used to update the other; the aforementioned data linkage is for the purpose of the NILS. A more detailed description of the NILS is available in [O'Reilly et al. \(2012\)](#).

Census data in the NILS is based on the enumerated population and therefore excludes wholly imputed individuals and households from the standard coverage adjustment process. Some NILS members do not have census information; this arises where individuals emigrated or died yet remained on the HCRS or where they simply did not provide a census return. This study was based on NILS members with a census record, which facilitated a comparison of address information with that in the HCRS. The lowest

level of geography available for NLS research is Super Output Area (SOA), which was introduced by NISRA for the 2001 Census; there are 890 SOAs in NI, with an average population size of 2,000 (NI Neighbourhood Information Service 2013).

Pointer, maintained by Land and Property Services, is the main address database for NI, in which each property has a Unique Property Reference Number (UPRN). In the HCRS and NI census data, address information includes the UPRN. Record-level address data in the NLS are provided in the form of anonymised UPRN, termed XUPRN. For this study, the availability of XUPRN in the HCRS and census data permitted an indirect comparison of the record-level address recorded in both sources. Owing to issues such as address formats not recognised by Pointer or incomplete address information, it is not possible to assign a UPRN in all cases. Accordingly, in the NLS, XUPRN can be missing in the census or HCRS data, missing in both or present in both; in the case of the latter, XUPRN can be a match or mismatch. This categorisation was used in presenting the cross-sectional results for 2011. The incidence of unassigned XUPRN was particularly high among addresses in Fermanagh, a predominantly rural area in the south-west of NI. This was mainly due to the use of a non-standard addressing system based on geographical units called townlands, which was not entirely incorporated in Pointer.

### 3.2. Methods

#### 3.2.1. Research Approach

The main element of this study was a longitudinal assessment of address data accuracy in the HCRS in terms of identifying the individual-, household-, and area-level factors associated with XUPRN mismatch. The census in 2001 and 2011 provided benchmark address information for NLS members as this was the most accurate record of where individuals resided at the time of enumeration. In addition, the routine updates of HCRS data gave a detailed address history for NLS members. The reference for one of the twice-yearly updates of HCRS data in the NLS was April, making it possible to obtain XUPRN data for a time point close to the 2001 and 2011 censuses conducted on April 29th and March 27th, respectively. A cross-sectional comparison of record-level XUPRN between the census and HCRS was undertaken for 2001 and 2011. This facilitated classification of NLS members with a census record and valid XUPRN recorded in both sources over time into four categories, namely (i) matching XUPRN between the census and HCRS in 2001 and 2011, (ii) matching XUPRN in 2001 only, (iii) matching XUPRN in 2011 only and (iv) mismatching XUPRN in 2001 and 2011. The primary focus of this study is the longitudinal assessment of address data accuracy, with descriptive results from the 2011 cross-sectional analysis also provided. In addition, the Supplemental data include further regression model output from the longitudinal analysis and 2001 and 2011 cross-sectional analyses. For clarification, *address* and *XUPRN* are used interchangeably, with *mismatch* and *match* referring to NLS members having an address recorded in the HCRS that was different or the same, respectively, to that from which their census questionnaire was returned.

#### 3.2.2. Explanatory Variables

Most of the person- and household-level explanatory variables used in the analysis were obtained from the 2001 and 2011 censuses of NI. Their selection was informed by similar

studies undertaken by [Shuttleworth and Barr \(2011\)](#), [Barr and Shuttleworth \(2012\)](#), and [Shuttleworth and Martin \(2016\)](#), which point to similarities between the hard-to-enumerate population from a census perspective and those with a greater likelihood of having inaccurate address information recorded in administrative systems. Accordingly, age and gender were included alongside marital status, level of education, socioeconomic status, and health status. Country of birth was chosen to investigate if XUPRN mismatch was prevalent among the immigrant population in NI. Migration, based on the ‘One year ago, what was your usual address’ census question was considered very relevant as it reflected residential moves, while housing tenure and household accommodation type and composition were informative in the context of residential mobility. Transition variables were derived for the longitudinal analysis based on the characteristics of NILES members in both the 2001 and 2011 censuses, for example, single in 2001 and married in 2011. The variables in question were marital status, limiting long-term illness (LLTI), National Statistics Socio-economic Classification (NSSEC), migration, and housing tenure. In addition, a derived variable capturing the frequency of record level address changes reported in the HCRS at six-month intervals between 2001 and 2011 was included. For the variables describing educational qualifications and religion, 2001 Census data were used on account of their antecedent status.

To investigate the geography of XUPRN mismatch, a number of area-level variables were included in the analysis. The 2005 NI Multiple Deprivation Measure (MDM) and component Proximity to Services domain score at SOA level ([NISRA 2005a](#)) provided a measure of spatial deprivation and the extent to which people had poor geographical access to key services, respectively, close to the mid-point of the 2001 to 2011 period. Based on the official statistical classification of settlements in NI ([NISRA 2005b](#)), the SOA of the census-recorded address was assigned urban or rural status. A variable based on whether the census-recorded address was in Fermanagh or elsewhere in NI was created to assess the effect of the addressing problems associated with the former. Explanatory variables that did not follow an approximate normal distribution were log transformed and expressed as quartiles where necessary.

### 3.2.3. Analysis

An initial descriptive analysis provided the percentage distributions of selected variables across the various XUPRN status categories for the 2011 cross-sectional analysis initially and, subsequently, the longitudinal assessment of address data quality in the HCRS. To further explore the multivariate relationships in the latter, multinomial logistic regression analysis was conducted using NILES records in the 16 to 74 age group; this restricted age category was relevant in the context of the education and socioeconomic status explanatory variables. For the four-category dependent variable, matching XUPRN in 2001 and 2011 was selected as the reference and the coefficients were expressed as relative risk ratios. Regarding the individual explanatory variables, the reference category was chosen on the basis of having the most records with a status of XUPRN match in both years. Data analysis was undertaken at the secure research environment of the NILES Research Support Unit, primarily using STATA version 14 ([StataCorp 2016](#)).



## 4. Results

### 4.1. Cross-Sectional Analysis (2011): Descriptive Results

The following descriptive results apply to 485,185 NILS members with a 2011 Census return. Regarding age, the XUPRN match rate between the census and HCRS fluctuated slightly around 85% for the 16 and under and 50 plus age groups (Figure 2). The trough in the XUPRN match rate for the 20 to 40 age group was reflected in the XUPRN mismatch rate, which peaked at 26% for those in their late twenties. The proportions of those with an unassigned XUPRN in the census, HCRS or both remained below ten per cent across all ages; some data were unavailable for the very old on account of aggregation for disclosure control purposes.

There was a reasonably large range in the XUPRN match and mismatch rates across other selected variables (Table 1). Individuals with the following characteristics were associated with a noticeably higher rate of XUPRN mismatch: male; marital status of single; without an LLTI; migrated within NI during the previous twelve months; private renter; living in a flat, apartment or communal establishment; living in a one person (under 65 age group), cohabiting couple or student household. In terms of area effect, the XUPRN match rate of 76% for rural SOAs was around seven per cent lower compared to SOAs classified as urban. Another finding of note was the relatively high proportion in the 'XUPRN not assigned' category of those who moved to NI from outside the country

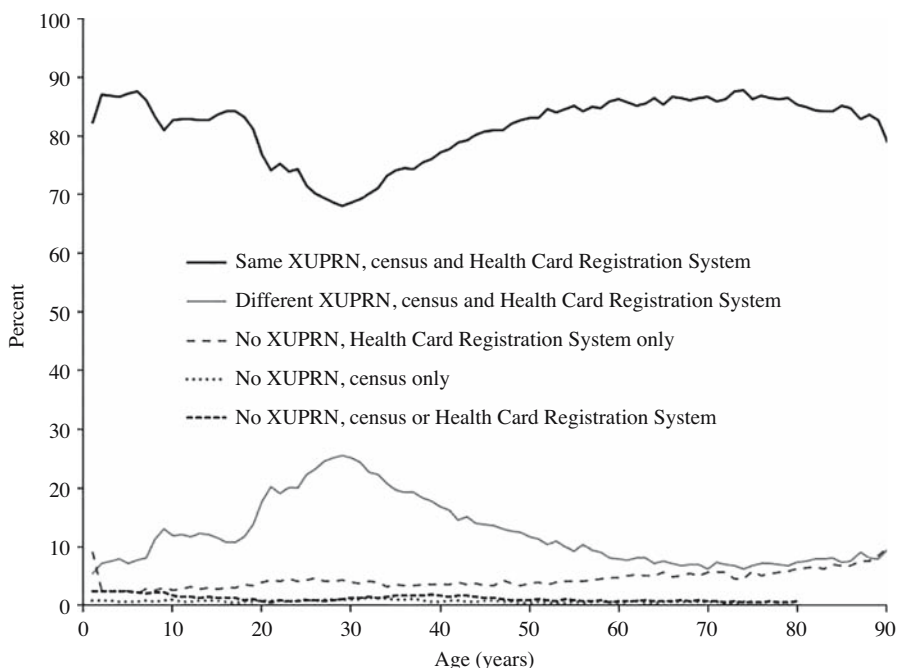


Fig. 2. Percentage distribution of Northern Ireland Longitudinal Study members by age according to XUPRN status category, based on a comparison of XUPRN in the census and Health Card Registration System in 2011; XUPRN is the anonymised Unique Property Reference Number.



Table 1. Distribution of individual-, household- and area-level characteristics based on a comparison of XUPRN in the census and Health Card Registration System (HCRS) in 2011 (figures are percentages based on the count numerator in the final column); XUPRN is the anonymised Unique Property Reference Number.

Variable	No XUPRN: census or HCRS	No XUPRN: census only	No XUPRN: HCRS only	Same XUPRN	Different XUPRN	Total
Sex						
Male	1.22	0.76	4.45	78.46	15.12	234,910
Female	1.21	0.66	3.80	83.01	11.32	250,275
Marital status						
Single	1.28	0.73	3.89	79.07	15.03	235,626
Married	1.34	0.77	4.22	82.68	10.99	186,524
Limiting long-term illness						
Yes	0.79	0.53	4.83	83.69	10.16	102,362
No	1.33	0.76	3.91	80.06	13.95	379,468
Address one year ago						
Same	1.19	0.63	3.93	82.66	11.60	450,766
Moved from within NI	1.59	2.14	5.10	51.58	39.59	25,538
Moved from outside NI	1.45	1.13	15.68	64.88	16.86	3,368
Housing tenure						
Owned	1.38	0.72	3.97	83.14	10.80	353,354
Social rented	0.47	0.45	2.57	82.97	13.54	57,643
Private rented	0.94	0.84	4.78	68.91	24.52	60,968
Accommodation type						
Detached	2.26	1.10	5.64	79.58	11.42	210,150
Semi-detached/terraced	0.28	0.23	1.96	84.71	12.82	245,546
Flat/apartment	1.27	1.72	8.98	61.19	26.83	23,835
Communal establishment	1.84	1.80	21.80	38.39	36.15	5,045

Table 1. Continued.

Variable	No XUPRN: census or HCRS	No XUPRN: census only	No XUPRN: HCRS only	Same XUPRN	Different XUPRN	Total
Household composition						
Couple, dependent children	1.75	0.78	3.78	83.62	10.07	217,390
Lone parent family	0.60	0.41	2.60	82.69	13.69	71,411
One person (over 65 years)	0.49	0.45	5.27	86.36	7.44	21,893
One person (other)	0.66	0.82	4.43	71.75	22.35	32,607
Cohabiting couple	0.80	0.84	3.45	70.48	24.43	27,094
Couple, no dependent children	1.12	0.75	4.25	81.51	12.37	40,649
Students/Communal establishment	1.53	1.79	19.34	34.93	42.41	6,204
Other	0.85	0.66	5.01	80.28	13.20	67,874
Urban/rural classification						
Urban	0.44	0.38	2.55	83.51	13.11	311,508
Rural	2.61	1.31	6.91	75.94	13.24	173,677

within the previous twelve months (15.7%) and residents of communal establishments (21.8%). The minor inconsistencies in the variable totals were due to the ‘No code required’ category in many of the census-based variables.

## 4.2. Longitudinal Analysis (2001 and 2011)

### 4.2.1. Descriptive Results

Descriptive results in [Table 2](#) from the 2001- and 2011-based longitudinal analysis were based on 334,670 NILS members with returns from both censuses. The distribution across the four XUPRN status categories was: 246,506 individuals (73.7%) with a matching XUPRN in the census and HCRS in 2001 and 2011; 28,661 (8.6%) with a matching XUPRN in 2001 only; 41,338 (12.4%) with a matching XUPRN in 2011 only; and 18,165 (5.4%) with a mismatching XUPRN in 2001 and 2011. A matching XUPRN in both years was more common among females and those born in the UK or Republic of Ireland (ROI). Consistency over time in the following transition variables was associated with the highest proportions in the category of XUPRN match in both years: being married; having an LLTI; not having migrated in the twelve months prior to the census; and residing in housing that was owned outright. For the category of XUPRN mismatch at both time points, there were noticeably high proportions among individuals who migrated in the twelve months before one or both of the censuses and those whose housing tenure changed from rented in 2001 to owner-occupied in 2011. A large proportion of those who migrated in 2001 only were in the category of XUPRN mismatch in 2001 only, as was the case with reference to 2011. As with [Table 1](#), the minor differences in the variable totals were due to the ‘No code required’ category.

### 4.2.2. Multivariate Results

Results from the multinomial logistic regression model in [Table 3](#) detail the multivariate relationships with longitudinal XUPRN status for 243,088 NILS members in the 16 to 74 age group. Statistical significance of variables was widespread on account of the large sample size so most attention is given to the size of effects. Address mismatch in the HCRS was more prevalent among males, who were 2.2 times more likely to have a mismatched XUPRN in 2001 and 2011 relative to matching XUPRN in both years. There was a greater likelihood of address mismatch in both years for those aged 25 to 34 in 2001 compared to the 35 to 44 age group, with those aged 45 to 74 in 2001 less likely to have a mismatched XUPRN between the HCRS and census. While the level of educational qualifications and country of birth were not strongly associated with longitudinal XUPRN status, Catholics had a greater likelihood of address mismatch in one or both years relative to Protestants and other Christians. The variable capturing the frequency of address changes in the HCRS between 2001 and 2011 was influential; compared to non-movers, those who reported at least one address change were less likely to have a mismatched XUPRN in 2011 only or both years relative to individuals with a matching XUPRN over time. Furthermore, the category of XUPRN mismatch in 2001 only had extremely high relative risk ratios, which is discussed further below.

Table 2. Distribution of individual-, household- and area-level characteristics by longitudinal XUPRN status, 2001 to 2011 (figures are percentages based on the count numerator in the final column); XUPRN is the anonymised Unique Property Reference Number.

	XUPRN match, 2001 and 2011	XUPRN match 2001, mismatch 2011	XUPRN mismatch 2001, match 2011	XUPRN mismatch, 2001 and 2011	Total
Sex					
Female	76.51	7.61	11.66	4.22	176,883
Male	70.46	9.63	13.12	6.78	157,787
Country of birth					
UK/Republic of Ireland	73.92	8.45	12.23	5.39	327,988
Outside UK/Republic of Ireland	69.65	8.91	15.81	5.63	3,782
Marital status (2001–2011)					
Single – Single	70.20	11.95	11.93	5.93	133,688
Married – Married	81.04	3.76	10.36	4.85	120,964
Other	73.94	6.36	14.48	5.22	35,866
Single – Married	54.52	16.58	22.37	6.54	24,739
Married – Separated/divorced/widowed	75.36	9.08	10.99	4.56	19,413
Limiting long-term illness (2001–2011)					
No – No	71.72	9.50	13.02	5.77	232,767
Yes – Yes	81.19	5.34	9.24	4.23	40,682
No – Yes	78.18	6.11	11.32	4.39	40,231
Yes – No	75.41	7.08	12.08	5.44	15,342
Address one year ago (2001–2011)					
Same – Same	78.17	7.23	9.94	4.66	289,754
Different – Same	40.66	4.72	43.28	11.34	22,184
Same – Different	45.53	40.15	7.74	8.58	13,340
Different – Different	26.75	14.43	32.33	26.49	1,525
Housing tenure (2001–2011)					
Owned – Owned	78.08	6.35	10.64	4.93	233,520
Rented – Rented	65.35	13.06	15.17	6.41	78,585
Owned – Rented	65.44	17.42	12.94	4.20	7,978
Rented – Owned	47.49	8.01	34.53	9.97	7,966
Communal establishment – Communal establishment	43.34	39.41	8.45	8.80	2,898

Table 3. Relative risk ratios and associated 95% confidence intervals for individual-, household- and area-level variables from a multinomial logistic regression model of longitudinal XUPRN status (2001 and 2011) for the 16 to 74 age group; XUPRN is the anonymised Unique Property Reference Number.

Variable	XUPRN match 2001, mismatch 2011		XUPRN mismatch 2001, match 2011		XUPRN mismatch 2001 and 2011	
	Relative risk ratio	95% confidence interval	Relative risk ratio	95% confidence interval	Relative risk ratio	95% confidence interval
Constant	0.04*	0.04–0.05	0.02*	0.01–0.02	0.04*	0.04–0.05
Sex (ref: Female)						
Male	1.64*	1.58–1.69	1.55*	1.51–1.59	2.18*	2.10–2.26
Age group, 2001 (ref: 35 to 44)						
16 to 24	2.52*	2.36–2.69	0.70*	0.66–0.74	1.06	0.98–1.14
25 to 34	1.78*	1.69–1.88	1.03	0.99–1.07	1.36*	1.29–1.43
45 to 54	0.67*	0.63–0.71	0.92*	0.88–0.97	0.71*	0.67–0.75
55 to 64	0.60*	0.55–0.64	0.77*	0.72–0.81	0.53*	0.50–0.57
65 to 74	0.62*	0.57–0.67	0.79*	0.73–0.85	0.52*	0.48–0.57
Educational qualifications, 2001 (ref: None)						
Level 1 to 3/apprenticeship	1.14*	1.09–1.19	0.98	0.94–1.02	1.06*	1.01–1.11
Level 4 and above	1.24*	1.17–1.32	1.05*	1.00–1.10	1.06	1.00–1.13
Religion belong to, 2001 (ref: Protestant/other Christian)						
Catholic	1.11*	1.07–1.15	1.27*	1.23–1.31	1.28*	1.23–1.33
Other religions	0.89	0.66–1.22	1.29*	1.02–1.64	1.01	0.71–1.43
No religion/not stated	1.10	0.98–1.24	1.17*	1.06–1.30	1.28*	1.12–1.47
Country of birth, 2001 (ref: UK/Ireland)						
Non-UK/Ireland	1.16*	1.01–1.34	0.94	0.83–1.06	0.97	0.82–1.14
Frequency of address changes in the HCRS, 2001–2011 (ref: None)						
1 to 2	0.71*	0.68–0.74	20.13*	19.34–20.96	0.88*	0.85–0.92
3 and above	0.45*	0.41–0.48	22.45*	21.19–23.79	0.74*	0.68–0.81

Table 3. Continued.

Variable	XUPRN match 2001, mismatch 2011		XUPRN mismatch 2001, match 2011		XUPRN mismatch 2001 and 2011	
	Relative risk ratio	95% confidence interval	Relative risk ratio	95% confidence interval	Relative risk ratio	95% confidence interval
Marital status transition, (ref: Married – Married)						
Single – Single	1.75*	1.66–1.84	1.21*	1.15–1.26	1.24*	1.17–1.31
Single – Married	2.45*	2.31–2.61	0.97	0.93–1.02	1.15*	1.08–1.24
Married – Separated/divorced/widowed	2.23*	2.09–2.37	0.86*	0.81–0.91	1.10*	1.01–1.19
Other	1.82*	1.71–1.93	1.19*	1.14–1.25	1.39*	1.31–1.47
Limiting long-term illness (LLTI) transition, (ref: No LLTI – No LLTI)						
LLTI – LLTI	0.84*	0.79–0.89	0.79*	0.75–0.83	0.77*	0.72–0.82
LLTI – No LLTI	0.84*	0.77–0.91	0.89*	0.83–0.95	0.88*	0.81–0.96
No LLTI – LLTI	0.82*	0.77–0.86	0.96	0.92–1.01	0.84*	0.79–0.89
Other	0.56	0.26–1.23	0.45	0.19–1.05	0.48	0.19–1.24
Socio-economic classification transition, (ref: Professional – Professional)						
Routine – Routine	0.85*	0.81–0.90	0.89*	0.85–0.93	0.80*	0.76–0.85
LT unemployed/never worked – LT unemployed/never worked	0.87*	0.77–0.98	0.97	0.88–1.08	0.89	0.79–1.01
Student – In employment	1.12*	1.05–1.19	0.71*	0.66–0.76	0.85*	0.78–0.94
Other	0.96	0.92–1.00	0.95*	0.92–0.99	0.86*	0.82–0.91
Address one year ago transition, (ref: Same address – Same address)						
Same address – Mover	4.19*	3.95–4.45	0.63*	0.58–0.69	2.53*	2.31–2.77
Mover – Same address	0.96	0.88–1.04	4.96*	4.75–5.18	4.22*	3.98–4.47
Mover – Mover	1.59*	1.27–2.00	3.98*	3.37–4.70	13.58*	11.38–16.22
Other	2.33*	1.85–2.94	2.48*	2.15–2.85	2.47*	2.05–2.99

Table 3. Continued.

Variable	XUPRN match 2001, mismatch 2011		XUPRN mismatch 2001, match 2011		XUPRN mismatch 2001 and 2011	
	Relative risk ratio	95% confidence interval	Relative risk ratio	95% confidence interval	Relative risk ratio	95% confidence interval
Housing tenure transition, (ref: Owned – Owned)						
Owned – Rented	3.24*	3.08–3.41	0.62*	0.59–0.66	1.42*	1.32–1.53
Rented – Owned	1.12*	1.05–1.19	1.32*	1.26–1.39	1.42*	1.33–1.52
Rented – Rented	1.31*	1.24–1.39	1.03	0.98–1.08	1.50*	1.41–1.60
Communal establishment – Communal establishment	9.23*	5.92–14.40	8.31*	5.59–12.37	25.14*	17.57–35.97
Other	3.41*	3.13–3.72	0.93	0.85–1.02	1.75*	1.56–1.96
Urban/rural classification of SOA, 2001 (ref: Urban)						
Rural	1.04	0.97–1.11	1.23*	1.17–1.30	1.29*	1.20–1.39
2005 Multiple Deprivation Measure score by SOA, log transformed	0.95	0.90–1.01	1.15*	1.10–1.21	1.05	0.99–1.13
2001-based SOA population density, log transformed	0.98	0.93–1.02	0.96	0.93–1.00	0.94*	0.89–0.98
Proximity to Services score by SOA, log transformed (ref: Quartile 1)						
Quartile 2	1.00	0.96–1.05	0.87*	0.83–0.90	0.84*	0.79–0.89
Quartile 3	0.99	0.93–1.05	0.95*	0.90–1.00	0.84*	0.78–0.90
Quartile 4	1.00	0.91–1.10	1.19*	1.10–1.30	1.04	0.94–1.16
Address location transition, (ref: Elsewhere in NI – Elsewhere in NI)						
Fermanagh – Fermanagh	1.23*	1.07–1.42	1.36*	1.20–1.55	1.53*	1.33–1.76
Fermanagh – Elsewhere in NI	1.21	0.77–1.91	0.99	0.69–1.42	2.18*	1.40–3.40
Elsewhere in NI – Fermanagh	3.69*	2.71–5.03	0.66*	0.47–0.92	1.51	0.97–2.35

\*Statistically significant.



Clear patterns were evident for the transition variables. Those with an unchanged marital status of single compared to individuals married at the two time points had a greater propensity for XUPRN mismatch in one or both years relative to matching XUPRN at both time points. Individuals with an LLTI in 2001 and/or 2011 and NILS members in routine employment in 2001 and 2011 relative to professionals were less likely to exhibit XUPRN mismatch in one or both years. There was a strong positive association between migration and address mismatch, with those who made an address change prior to each census at 14 times the risk of XUPRN mismatch in both 2001 and 2011. Housing tenure was influential, with renters and especially communal establishment residents at both time points having a greater likelihood of XUPRN mismatch in both years. Compared to the individual- and household-level factors, there was less variation in the relative risk ratios for the area-level variables, with many close to the threshold value of one. This indicates that from a longitudinal perspective, area characteristics were less influential in terms of the determinants of address mismatch in the HCRS. There were some noteworthy findings however, with rural relative to urban areas and having a census-recorded address in Fermanagh in both years as opposed to elsewhere in NI generally associated with a greater likelihood of XUPRN mismatch in 2001 and/or 2011.

The interaction of migration transition with age was investigated (see Supplemental data, Table 1 found online at: <http://dx.doi.org/10.1515/JOS-2018-0004>). Compared to older age groups, 16–24 and 25–34 year olds in 2001 who migrated in the year prior to each census had a greater likelihood of consistent address mismatch relative to non-movers. A further interaction of migration and housing tenure transitions indicated that NILS members in rented accommodation who made pre-census residential moves were at greater risk of XUPRN mismatch compared to those in owner-occupied housing.

The findings from the 2001 and 2011 cross-sectional logistic regression models of XUPRN mismatch (see Supplemental data, Table 2 found online at: <http://dx.doi.org/10.1515/JOS-2018-0004>) were similar to those from the longitudinal model of address mismatch (Table 3). The former models revealed common characteristics associated with address mismatch at each time point, namely males, young adults (aged 25–35) relative to older age groups, single people compared to those who were married, pre-census migrants, renters and communal establishment residents as opposed to those in owner-occupied housing, and in terms of geography, living in Belfast or Fermanagh relative to elsewhere in NI.

#### 4.2.3. Residential Movement

The longitudinal analysis revealed 18,165 NILS members with an address mismatch in the HCRS in 2001 and 2011. Of particular interest was whether this group made more residential moves compared to those with a matching XUPRN in both census years. The updates of address data from the HCRS every six months from 2001 to 2014 facilitated analysis of the proportion in each of the four XUPRN status categories with a reported address change at each time point. Figure 3 indicates little difference in the proportion of individuals reporting an address change between 2001 to 2011 for the categories of matching and mismatching XUPRN in both years, which suggests similar levels of residential movement over the decade. The main explanation seems to be that individuals who mismatched in 2001 and 2011 had happened to change address in the year before the

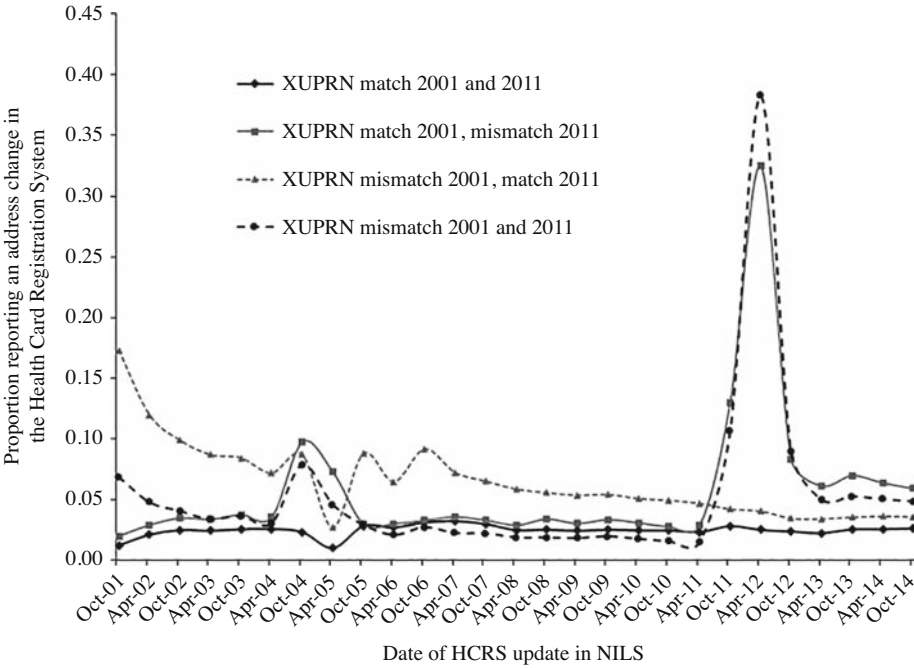


Fig. 3. Proportion of Northern Ireland Longitudinal Study (NILS) members reporting an address change by six-monthly extract of the Health Card Registration System (HCERS) from 2001 to 2014, according to longitudinal XUPRN status; XUPRN is the anonymised Unique Property Reference Number.

respective census but not that they were more spatially mobile in general. The sharp peak for those with a mismatched XUPRN in 2011 can be attributed in part to NILS members who made a residential move prior to the 2011 Census, which was not reported to the HCERS until after March 2011. However, this peak is somewhat inflated by a data cleaning exercise undertaken by BSO during 2011 and 2012, which resulted in some non-genuine address changes being recorded in the HCERS data provided to the NILS. Although the time series begins in October 2001, the remnants of a similar peak associated with lagged reporting of an address change to the HCERS is evident for those with a mismatched address in 2001.

### 5. Discussion

This study provides a useful insight on an element of data quality in a key administrative source and highlights the value of the NILS as a research resource. Of the individual-, household-, and area-level factors associated with address mismatch in the HCERS, many are intuitive and others less so. In general, the frequency of residential moves is highest among young adults for reasons such as leaving home, pursuing higher education and entering the labour market. In addition, this population group has a low level of engagement with primary care services, thus increasing the likelihood of address changes not being communicated to the associated administrative system. It is very likely that these combined factors contributed to the high rates of address mismatch in the HCERS for the 20 to 34 years age group observed in the cross-sectional and longitudinal analyses. Those

with an LLTI can be expected to engage with their general practitioner to a greater extent compared to individuals in good health, thereby increasing the likelihood of maintaining an accurate address in the HCRS (Shuttleworth and Barr 2011; Barr and Shuttleworth 2012). The greater prevalence of XUPRN mismatch among Catholics may be explained in part by their geographical distribution across NI; rural relative to urban SOAs had a lower likelihood of address match between the HCRS and census and both the former and SOAs with a majority Catholic resident population are prevalent in the west and south-west of the country. The often unforeseen movement of individuals into care-homes, hospitals or prisons along with the transience of many communal establishment populations is a probable explanation for the poor quality address information in the HCRS displayed by individuals residing in this accommodation type. Regarding area effects, the address referencing problem in Fermanagh appeared to exacerbate XUPRN mismatch, which was more prevalent in this part of NI. Although records with an address in this area amounted to just 1.5% and 1.9% of all NLS records with an assigned XUPRN in 2001 and 2011, respectively, the Fermanagh effect may represent a wider issue with address formats and referencing in rural parts of NI that warrants further investigation.

The 2011 cross-sectional analysis provided evidence of higher incidence of address mismatch in the HCRS for those in households that deviated from the traditional nuclear family structure. This is important as increasingly diverse households have become a feature of modern society. For example, in the UK in 2015, cohabiting couples were the fastest growing family type, accounting for 17% of all families, while 29% of households consisted of just one person (ONS 2015c). These household types are also becoming more common in Canada (SC 2012), the United States (USCB 2013), and several European countries (Sanchez Gassen and Perelli-Harris 2015). In addition, multi-family and multi-generational households are becoming more numerous (ONS 2015c; Fry and Passel 2014). For cohabiting couples without dependent children and single-person households in particular, there is likely to be greater scope for residential movement and, consequently, their address information to be out-dated in administrative systems. In the context of NSIs using administrative sources to inform population statistics, the frequency of engagement by particular population groups with the administrative system(s) in question is a key consideration regarding address information accuracy. Indeed, the increasing complexity of household structure presents a challenge to NSIs when conducting a census irrespective of whether they continue with a traditional approach or adopt a method more reliant on administrative sources. Households are a fundamental observation unit for population statistics; therefore, the potential of administrative sources to provide an insight on household structure should be explored so that NSIs can define households accurately and appropriately.

Student households were shown to be problematic in terms of their residents maintaining accurate address information in the HCRS; a contributory factor is likely to be the highly mobile nature of student populations (Duke-Williams 2009; Finney 2011). Areas around higher education institutions are subject to considerable population turnover, including inflows of new and continuing students in line with the academic year and outflows of graduates pursuing employment or further education opportunities elsewhere. This group present challenges for the production of high quality population statistics. For example, there were specific questions included in the 2011 UK Census to determine the appropriate address at which to classify a student as usually resident. Furthermore, the

methodology underlying the estimation of annual internal migration flows in UK constituent countries now incorporates higher education data to provide more accurate student address information (NISRA 2016b; ONS 2016). Regarding health register data, students are infrequent users of primary care services, which compounds the problem of address inaccuracy. In using secondary data such as health registers to produce population statistics, it is prudent for NSIs to incorporate supplementary sources that provide better quality address data on highly mobile groups such as students. Furthermore, these findings highlight the importance of addressing key areas of weakness in administrative sources when used by NSIs to augment or replace a traditional census; a robust statistical system drawing upon administrative sources should seek to improve the quality of the source data.

During the 2000s, international migration to NI experienced an unprecedented surge, with around 122,000 immigrants estimated to have arrived in the country over the decade. European and North-American studies have shown that the propensity to migrate internally among immigrants is greatest for recent arrivals to the host country (Reher and Silvestre 2009; King and Newbold 2011), which is relevant in the context of maintaining accurate address information in administrative systems that immigrants interact with. In general, this study did not find a greater prevalence of address mismatch in the HCRS for NILS members born outside the UK or ROI compared to the native population. A possible factor is that the vast majority of immigrants were from Poland and Lithuania (Krausova and Vargas-Silva 2014); there is evidence of the former developing strong ties with their place of residence after being most internally mobile in the initial stages of their immigration to the UK (Trevena et al. 2013). As immigration to NI increased steadily from 2000 before peaking in 2007, it is likely that a large proportion of this population group had settled in residential terms by 2011, thus providing greater scope for having accurate address information recorded in the HCRS.

Many of the determinants of address mismatch in the HCRS indicated by this study exhibit a strong relationship with residential mobility and, consequently, internal migration. This is unsurprising since all of these phenomena are intertwined. Males, young adults, individuals who are single, those in good health and professionals have been shown to be more residentially mobile and migrate internally to a greater extent (Owen and Green 1992; Bailey and Livingston 2007; Champion et al. 1998; Finney and Simpson 2008); these characteristics also exhibited a positive relationship with XUPRN mismatch in the HCRS. A number of the aforementioned factors are also relevant in the context of census under-enumeration. The issue of nonresponse when conducting a population census is most pronounced among the likes of young adult males and residents of single occupant and student households (Rahman and Goldring 2006; Martin 2010); likewise, these had a higher likelihood of exhibiting XUPRN mismatch. For NSIs planning to supplement their traditional census process with administrative data, it will be important to consider these individual and household characteristics not just in terms of the hard-to-enumerate population but also where address information in the administrative sources will be drawn upon for specific purposes such as informing an address register or data linkage.

Having migrated in the year prior to each census was strongly associated with address mismatch in the HCRS, based on the main effect and interactions with age and housing tenure. As migration is a continual demographic process, a cohort with questionable locational information persists, which needs to be accounted for when using address

information in health register data in the population statistics production process. Another interesting finding from the longitudinal analysis was the similar levels of residential mobility between individuals with a mismatched XUPRN in 2001 and 2011 and those with a matching XUPRN in both years. A reasonable initial hypothesis would have been more residential movement by the former group. Along with the post-census peak in address changes recorded in the HCRS for individuals with a mismatched XUPRN at census time, it reaffirms the existence of a cohort who persistently lagged in reporting or failed to report an address change to the health register, as observed by [Barr and Shuttleworth \(2012\)](#) and [Shuttleworth and Martin \(2016\)](#). For mobile population groups such as young adult males who are more prone to address inaccuracy in administrative systems, the longitudinal structure of the NILS would facilitate investigation of the extent to which address information at a point in time could be used to infer accurate address data at a previous juncture.

The comparable findings from the longitudinal and cross-sectional models suggests that it was the status at the time of census enumeration and not necessarily the 2001 to 2011 transitions that contributed to address inaccuracy in the HCRS. Therefore, the latter models were sufficient to gain an understanding of the factors associated with this aspect of data accuracy in the administrative source. Nonetheless, the longitudinal model was a novel extension of the approach adopted by [Shuttleworth and Martin \(2016\)](#) in their 2001-based study and was appropriate for the inclusion of the variable based on the frequency of address changing in the HCRS over the 2001–2011 period, which proved to be an influential factor. In this regard, the extremely high relative risk ratios for the category of XUPRN mismatch in 2001 only warranted further investigation. From the raw data, 74% and 15% of this cohort reported 1–2 and 3 or more address changes, respectively, to the HCRS over the decade; the corresponding distribution for the XUPRN mismatch in 2011 only group was 35% and 7%, and 29% and 7% for those mismatching in both years. This suggests a prevalence of address changing prior to March 2001 within the XUPRN mismatch in 2001 only group that was not reported to the HCRS until after the 2001 Census.

The extent of address mismatch in the HCRS has implications for the statistical application of this administrative source. National health registers elsewhere are likely to experience similar data quality issues, so while the potential impacts are described primarily in a UK context, they are relevant to other countries. The quality of internal migration estimates, based largely on address changes in health register data, is affected by address inaccuracy. Internal moves are missed or lagged where individuals fail to report a change of address or a period of time elapses before doing so. As this component greatly influences population size and composition at local level, inaccurate address information in the health register underlying the estimates affects the quality of annual subnational population estimates; this is important given the widespread use of these data, which includes informing the allocation of funding from central government to local authorities, councils and health bodies. For the next England and Wales census in 2021, there are a number of broad uses of administrative data proposed. These include using activity information to enhance the estimates of the size and location of the population ([ONS 2015d](#)); health registers are likely to be a key source in this regard based on the successful use of HCRS data for the 2011 Census under-enumeration project in NI ([NISRA 2015](#)). Furthermore, in countries where the traditional census is discontinued, administrative

sources such as the health register are likely to have a key role in the production of population statistics. Record linkage between data sources will be a fundamental aspect of this alternative approach, which typically uses address information along with name and date of birth as linking variables. Address inaccuracies in health register data could therefore adversely affect the matching quality achieved in record linkage processing.

An important aspect underlying this analysis is the increasing difficulty in referencing individuals to a single address. Features of modern society such as second addresses for holiday, work or study purposes and fragmented family structures causing shared custody of children result in individuals residing at more than one address. This was acknowledged in the 2011 Census of England and Wales, with ONS releasing statistics based on alternative population bases, namely the out-of-term student population and usually resident dependent children with a parental second address (ONS 2014b, 2014c). Therefore, genuine address inconsistencies between census and administrative data sources can exist. The extent to which this contributed to the XUPRN mismatch between census and HCRS data observed in this study is unknown. However, it is one of many challenges for NSIs that move away from a traditional census in favour of a methodology based on linked administrative sources.

## 6. Conclusion

This study took a novel longitudinal approach, facilitated by the NILS, to investigate the extent of address mismatch in a key administrative source and the associated factors. Our findings provide evidence of the commonalities with internal migration, residential mobility and hard-to-enumerate groups from a census perspective. Currently, health registers are used by numerous NSIs to inform the ongoing production of population statistics. In addition, given the apparent discontinuation of the traditional census in many countries, these and other administrative sources are likely to be fundamental to the generation of population and census statistics in the future. While national health registers are a valuable statistical resource owing to their high population coverage, this and other studies have shown the quality of their record-level address information to be impaired for certain population groups. It is important to further improve understanding of quality issues in this and other administrative sources that are likely to underpin official statistical systems in the future. Many of the limitations of administrative data such as undercoverage and lags in updating are acknowledged in the context of their application for statistical purposes. Indeed, some NSIs incorporate additional administrative sources to address specific shortcomings. However, it is prudent to build upon existing research on administrative data quality; effective methodological improvements can then be developed and implemented to ensure that official statistics generated from these sources are fit for purpose and sufficiently accurate.

## 7. References

Agafitei, M., F. Gras, W. Kloek, F. Reis, and S. Vâju. 2015. "Measuring Output Quality for Multisource Statistics in Official Statistics: Some Directions." *Statistical Journal of the IAOS* 31: 203–211. Doi: <http://dx.doi.org/10.3233/sji-150902>.



- Andersson, C., A. Holmberg, I. Jansson, K. Lindgren, and P. Werner. 2013. "Methodological Experiences from a Register-Based Census." In Proceedings of the Section on Survey Research Methods: American Statistical Association, August 6, 2009. 3289–3296. Alexandria, US. Available at: [https://ww2.amstat.org/sections/srms/Proceedings/y2013/Files/309549\\_82867.pdf](https://ww2.amstat.org/sections/srms/Proceedings/y2013/Files/309549_82867.pdf) (accessed August 2017).
- Australian Bureau of Statistics (ABS). 2014. *Information Paper: Review of Interstate Migration Method*. Available at: <http://www.abs.gov.au/AUSSTATS/abs@.nsf/Lookup/3412.0.55.003Main+Features1Mar%202014?OpenDocument> (accessed August 2017).
- Bailey, N. and M. Livingston. 2007. *Population Turnover and Area Deprivation: Final Report*. Joseph Rowntree Foundation. Available at: <https://www.jrf.org.uk/sites/default/files/jrf/migrated/files/2004-population-census-deprivation.pdf> (accessed August 2017).
- Barr, P. and I. Shuttleworth. 2012. "Reporting Address Changes by Migrants: The Accuracy and Timeliness of Reports via Health Card Registers." *Health & Place* 18: 595–604. Doi: <http://dx.doi.org/10.1016/j.healthplace.2012.01.005>.
- Bycroft, C. 2015. "Census Transformation in New Zealand: Using Administrative Data without a Population Register." *Statistical Journal of the IAOS* 3(13): 401–411. Doi: <http://dx.doi.org/10.3233/SJI-150916>.
- Champion, T., S. Fotheringham, P. Rees, P. Boyle, and J. Stillwell. 1998. *The Determinants of Migration Flows in England: A Review of Existing Data and Evidence*. Report for the Department of the Environment, Transport and the Regions, UK. Available at: <http://www.geog.leeds.ac.uk/publications/DeterminantsOfMigration/report.pdf> (accessed August 2017).
- Duke-Williams, O. 2009. "The Geographies of Student Migration in the UK." *Environment and Planning A* 41(8): 1826–1848. Doi: <http://dx.doi.org/10.1068/a4198>.
- Eurostat. 2003. *Quality Assessment of Administrative Data for Statistical Purposes*. Eurostat, Luxembourg. Available at: <https://unstats.un.org/unsd/EconStatKB/KnowledgebaseArticle10264.aspx> (accessed August 2017).
- Finney, N. and L. Simpson. 2008. "Internal Migration and Ethnic Groups: Evidence for Britain from the 2001 Census." *Population, Space and Place* 14: 63–83. Doi: <http://dx.doi.org/10.1002/psp.481>.
- Finney, N. 2011. "Understanding Ethnic Differences in the Migration of Young Adults within Britain from a Lifecourse Perspective." *Transactions of the Institute of British Geographers* 36(3): 455–470. Doi: <http://dx.doi.org/10.1111/j.1475-5661.2011.00426.x>.
- Fry, R. and J.S. Passel. 2014. *In Post-Recession Era, Young Adults Drive Continuing Rise in Multi-generational Living*. Pew Research Center's Social and Demographic Trends project Washington DC, US. Available at: <http://www.pewsocialtrends.org/files/2014/07/ST-2014-07-17-multigen-households-report.pdf> (accessed August 2017).
- Karr, A.F. 2012. "Discussion on Statistical Use of Administrative Data: Old and New Challenges." *Statistica Neerlandica* 66(1): 80–84. Doi: <http://dx.doi.org/10.1111/j.1467-9574.2011.00510.x>.
- King, K.M. and K.B. Newbold. 2011. "Internal Migration of Canadian Immigrants, 1993–2004: Evidence from the Survey of Labour and Income Dynamics." *Canadian Studies in Population* 38: 1–18.



- Krausova, A. and C. Vargas-Silva. 2014. *Northern Ireland: Census Profile*. Migration Observatory briefing, COMPAS, University of Oxford, UK. Available at: <http://www.migrationobservatory.ox.ac.uk/resources/briefings/northern-ireland-census-profile/> (accessed August 2017).
- Kukutai, T., V. Thompson, and R. McMillan. 2015. "Whither the Census? Continuity and Change in Census Methodologies Worldwide, 1985–2014." *Journal of Population Research* 32(1): 3–22. Doi: <http://dx.doi.org/10.1007/s12546-014-9139-z>.
- Lange, A. 2014. "The Population and Housing Census in a Register Based Statistical System." *Statistical Journal of the IAOS* 30(1): 41–45. Doi: <http://dx.doi.org/10.3233/SJI-140798>.
- Lo Scalzo, A., A. Donatini, L. Orzella, A. Cicchetti, S. Profili, and A. Maresso. 2009. *Italy: Health system review*. World Health Organisation, Copenhagen, Denmark. Available at: [http://www.euro.who.int/\\_\\_data/assets/pdf\\_file/0006/87225/E93666.pdf](http://www.euro.who.int/__data/assets/pdf_file/0006/87225/E93666.pdf) (accessed August 2017).
- Martin, D. 2006. "Last of the Censuses? The Future of Small Area Population Data." *Transactions of the Institute of British Geographers, New Series* 31: 6–18. Doi: <http://dx.doi.org/10.1111/j.1475-5661.2006.00189.x>.
- Martin, D. 2010. "Understanding the Social Geography of Census Undercount." *Environment and Planning A* 42(11): 2753–2770. Doi: <http://dx.doi.org/10.1068/a43123>.
- Millett, C., C. Zelenyanszki, K. Binysh, J. Lancaster, and A. Majeed. 2005. "Population Mobility: Characteristics of People Registering with General Practices." *Public Health* 119(7): 632–638. Doi: <http://dx.doi.org/10.1016/j.puhe.2004.09.004>.
- Moore, T., L. Bailie, and G. Gilmour. 2008. *Building a Business Case for Census Internet Data Collection*. In Proceedings of Statistics Canada Symposium. Statistics Canada, Ontario, Canada. Available at: <http://www.statcan.gc.ca/pub/11-522-x/2008000/article/10978-eng.pdf> (accessed August 2017).
- National Records of Scotland (NRS). 2012. *Overview of Administrative Comparator Data Used in 2011 Census Quality Assurance*. Edinburgh, UK: NRS. Available at: <http://www.scotlandscensus.gov.uk/documents/methodology/census-overview-comparative-sources-uksa.pdf> (accessed August 2017).
- NRS. 2014. *Background information on the Beyond 2011 Programme*. Edinburgh, UK: NRS. Available at: <http://www.nrscotland.gov.uk/files/census/2021-census/Background/2021-census-background.pdf> (accessed August 2017).
- Northern Ireland Longitudinal Study (NILS). 2015. *NILS Data Matching Methodology, NILS Working Paper 3.1*. Belfast, UK: NISRA. Available at: <http://www.qub.ac.uk/research-centres/NILSResearchSupportUnit/FileStore/Fileupload,425661,en.DOCX> (accessed August 2017).
- Northern Ireland Neighbourhood Information Service. 2013. *NISRA Geography Fact Sheet*. Belfast, UK: NISRA. Available at: <http://www.ninis2.nisra.gov.uk/public/documents/NISRA%20Geography%20Fact%20Sheet.pdf> (accessed August 2017).
- Northern Ireland Statistics and Research Agency (NISRA). 2005a. *Northern Ireland Multiple Deprivation Measure 2005*. Belfast, UK: NISRA. Available at: <https://www.nisra.gov.uk/sites/nisra.gov.uk/files/publications/NIMDM2005FullReport.pdf> (accessed August 2017).

- NISRA. 2005b. *Report of the Inter-Departmental Urban-Rural Definition Group: Statistical Classification and Delineation of Settlements*. Belfast, UK: NISRA. Available at: [http://www.ninis2.nisra.gov.uk/public/documents/ur\\_report.pdf](http://www.ninis2.nisra.gov.uk/public/documents/ur_report.pdf) (accessed August 2017).
- NISRA. 2012. *Quality Assurance of the 2011 Census in Northern Ireland*. Belfast, UK: NISRA. Available at: <https://www.nisra.gov.uk/sites/nisra.gov.uk/files/publications/2011-census-quality-assurance-strategy.pdf> (accessed August 2017).
- NISRA. 2014. *The Future Provision of Census of Population Information for Northern Ireland*. Belfast, UK: NISRA. Available at: <https://www.nisra.gov.uk/sites/nisra.gov.uk/files/publications/the-future-provision-of-census-of-population-information-for-northern-ireland.pdf> (accessed August 2017).
- NISRA. 2015. *Northern Ireland Census 2011 General Report*. Belfast, UK: NISRA. Available at: <https://www.nisra.gov.uk/sites/nisra.gov.uk/files/publications/2011-census-general-report.pdf> (accessed August 2017).
- NISRA. 2016a. *Population Estimates and Projections Data Quality Document*. Belfast, UK: NISRA. Available at: [https://www.nisra.gov.uk/sites/nisra.gov.uk/files/publications/Population-DataQuality\\_1.pdf](https://www.nisra.gov.uk/sites/nisra.gov.uk/files/publications/Population-DataQuality_1.pdf) (accessed August 2017).
- NISRA. 2016b. *Methodology Paper – Mid-Year Population Estimates for Northern Ireland*. Belfast, UK: NISRA. Available at: <https://www.nisra.gov.uk/sites/nisra.gov.uk/files/publications/Methodology-2015.PDF> (accessed August 2017).
- Office for National Statistics (ONS). 2011. *Improved Immigration Estimates to Local Authorities in England and Wales: Overview of Methodology*. Hampshire, UK: ONS. Available at: <http://www.ons.gov.uk/ons/guide-method/method-quality/imps/improvements-to-local-authority-immigration-estimates/overview-of-improved-methodology.doc> (accessed August 2017).
- ONS. 2012a. *Overview of Administrative Comparator Data Used in 2011 Census Quality Assurance*. Hampshire, UK: ONS. Available at: <http://www.ons.gov.uk/ons/guide-method/census/2011/the-2011-census/processing-the-information/data-quality-assurance/overview-of-administrative-comparator-data-used-in-2011-census-quality-assurance.pdf> (accessed August 2017).
- ONS. 2012b. *Beyond 2011: Administrative Data Sources Report: NHS Patient Register*. Hampshire, UK: ONS. Available at: <http://www.ons.gov.uk/ons/about-ons/who-ons-are/programmes-and-projects/beyond-2011/reports-and-publications/sources-reports/beyond-2011-administrative-data-sources-report-nhs-patient-register-s1-.pdf> (accessed December 2015).
- ONS. 2012c. *Beyond 2011: Exploring the Challenges of Using Administrative Data*. Hampshire, UK: ONS. Available at: <http://www.ons.gov.uk/ons/about-ons/who-ons-are/programmes-and-projects/beyond-2011/reports-and-publications/methods-and-policies-reports/beyond-2011-exploring-the-challenges-of-using-administrative-data.pdf> (accessed August 2017).
- ONS. 2014a. *The census and future provision of population statistics in England and Wales: Recommendation from the National Statistician and Chief Executive of the UK Statistics Authority*. Hampshire, UK: ONS. Available at: <http://www.ons.gov.uk/ons/about-ons/who-ons-are/programmes-and-projects/beyond-2011/beyond-2011-report-on-autumn-2013-consultation-and-recommendations/national-statisticians-recommendation.pdf> (accessed August 2017).

- ONS. 2014b. *Out of Term Population of England and Wales: An Alternative 2011 Census Population Base*. Hampshire, UK: ONS. Available at: [http://www.ons.gov.uk/ons/dcp171776\\_377904.pdf](http://www.ons.gov.uk/ons/dcp171776_377904.pdf) (accessed August 2017).
- ONS. 2014c. *Dependent Children Usually Resident in England and Wales with a Parental Second Address, 2011*. Hampshire, UK: ONS. Available at: [http://www.ons.gov.uk/ons/dcp171776\\_372236.pdf](http://www.ons.gov.uk/ons/dcp171776_372236.pdf) (accessed August 2017).
- ONS. 2015a. *Beyond 2011 research strategy and plan – 2015 to 2017*. Hampshire, UK: ONS. Available at: <http://www.ons.gov.uk/ons/about-ons/who-ons-are/programmes-and-projects/beyond-2011/reports-and-publications/research-strategy-and-plan—2015-2017.pdf> (accessed August 2017).
- ONS. 2015b. *2021 Administrative Data Research Report 2015: Methodology and Analysis of Estimates Produced from a Statistical Population Dataset*. Hampshire, UK: ONS. Available at: <http://www.ons.gov.uk/ons/guide-method/census/2021-census/progress-and-development/research-projects/beyond-2011-research-and-design/research-outputs/administrative-data-research-outputs—2015.pdf> (accessed August 2017).
- ONS. 2015c. *Families and Households: 2015*. Hampshire, UK: ONS. Available at: <http://www.ons.gov.uk/peoplepopulationandcommunity/birthsdeathsandmarriages/families/bulletins/familiesandhouseholds/2015-11-05/pdf> (accessed March 2015).
- ONS. 2015d. *2021 Census Design Document*. Hampshire, UK: ONS. Available at: <http://www.ons.gov.uk/ons/guide-method/census/2021-census/progress-and-development/research-projects/2021-census-design-document.pdf> (accessed December 2015).
- ONS. 2016. *Internal migration by local authorities in England and Wales*. Hampshire, UK: ONS. Available at: <https://www.ons.gov.uk/file?uri=/peoplepopulationandcommunity/populationandmigration/migrationwithintheuk/qmis/internalmigrationestimatesqmi/qmiimelajune16finalforpub.pdf> (accessed August 2017).
- ONS. 2017. *Northern Ireland internal migration and Northern Ireland Medical Card information: quality assurance of administrative data used in population statistics, Feb 2017*: ONS. Available at: <https://www.ons.gov.uk/peoplepopulationandcommunity/populationandmigration/populationestimates/methodologies/northernirelandinternalmigrationandnorthernirelandmedicalcardinformationqualityassuranceofadministratedatausedinpopulationstatisticsfeb2017/pdf> (accessed August 2017).
- O'Reilly, D., M. Rosato, G. Catney, F. Johnston, and M. Brolly. 2012. "Cohort Description: The Northern Ireland Longitudinal Study (NILS)." *International Journal of Epidemiology* 41(3): 634–641. Doi: <http://dx.doi.org/10.1093/ije/dyq271>.
- Owen, D.W. and A.E. Green. 1992. "Migration Patterns and Trends." In *Migration Processes and Patterns Volume 1: Research Progress and Prospects*, edited by A.G. Champion and A. Fielding, 17–38. London: Belhaven.
- Pedersen, K.M., J.S. Andersen, and J. Søndergaard. 2012. "General Practice and Primary Health Care in Denmark." *The Journal of the American Board of Family Medicine* 25(Suppl 1): S34–S38. Doi: <http://dx.doi.org/10.3122/jabfm.2012.02.110216>.
- Rahman, N. and S. Goldring. 2006. "Factors Associated with Household Non-Response in the 2001 Census." Newport, UK: ONS. *Survey Methodology Bulletin* 59: 11–24.

- Raymer, J., P. Rees, and A. Blake. 2015. "Frameworks for Guiding the Development and Improvement of Population Statistics in the United Kingdom." *Journal of Official Statistics* 31(4): 699–722. Doi: <https://doi.org/10.1515/jos-2015-0041>.
- Reher, D.S. and J. Silvestre. 2009. "Internal Migration Patterns of Foreign-Born Immigrants in a Country of Recent Mass Immigration: Evidence from New Micro Data for Spain." *International Migration Review* 43: 815–849. Doi: <http://dx.doi.org/10.1111/j.1747-7379.2009.00785.x>.
- Roland, M., B. Guthrie, and D.C. Thomé. 2012. "Primary Medical Care in the United Kingdom." *The Journal of the American Board of Family Medicine* 25(Suppl 1): S6–S11. Doi: <http://dx.doi.org/10.3122/jabfm.2012.02.110200>.
- Sanchez Gassen, N. and B. Perelli-Harris. 2015. "The Increase in Cohabitation and the role of Union Status in Family Policies: A Comparison of 12 European Countries." *Journal of European Social Policy* 25(4): 431–449. Doi: <http://dx.doi.org/10.1177/0958928715594561>.
- Shuttleworth, I. and P. Barr. 2011. "Who Reports Address Changes Through the Health Service System? The Characteristics of Laggards and Non-Reporters Using the Northern Ireland Longitudinal Study." Newport, UK: ONS. *Population Trends* 144: 48–54.
- Shuttleworth, I. and D. Martin. 2016. "People and Places: Understanding Geographical Accuracy in Administrative Data from the Census and Health Service Systems." *Environment and Planning A* 48: 594–610. Doi: <http://dx.doi.org/10.1177/0308518X15618205>.
- Smallwood, S. and K. Lynch. 2010. "An Analysis of Patient Register Data in the Longitudinal Study – What does it Tell Us about the Quality of the Data?" Newport, UK: ONS. *Population Trends* 141: 151–169.
- StataCorp. 2016. Stata Statistical Software: Release14. StataCorp LP, College Station, Texas, US.
- Statistics Canada (SC). 2012. *Canadian households in 2011: Type and growth*. Ontario, Canada: SC. Available at: [https://www12.statcan.gc.ca/census-recensement/2011/as-sa/98-312-x/98-312-x2011003\\_2-eng.cfm](https://www12.statcan.gc.ca/census-recensement/2011/as-sa/98-312-x/98-312-x2011003_2-eng.cfm) (accessed August 2017).
- SC. 2015. *Population and Family Estimation Methods at Statistics Canada*. Ontario, Canada: SC. Available at: <http://www.statcan.gc.ca/pub/91-528-x/2015001/ch/ch7-eng.htm> (accessed August 2017).
- Statistics New Zealand. 2013. *Evaluation of Administrative Data Sources for Subnational Population Estimates*. Wellington, NZ: SNZ. Available at: <http://www.stats.govt.nz/~media/Statistics/browse-categories/population/estimates-projections/eval-admin-subnat-pop-est/eval-admin-data-subnat-pop-est.pdf> (accessed August 2017).
- Trevena, P., D. McGhee, and S. Heath. 2013. "Location, Location? A Critical Examination of Patterns and Determinants of Internal Mobility Among Post-accession Polish Migrants in the UK." *Population, Space and Place* 19: 671–687. Doi: <http://dx.doi.org/10.1002/psp.1788>.
- United Nations Economic Commission for Europe (UNECE). 2007. *Register-Based Statistics in the Nordic Countries: Review of Best Practices with Focus on Population and Social Statistics*. Geneva, Switzerland: UNECE. Available at: <http://unstats.un.org/unsd/dnss/docViewer.aspx?docID=2764> (accessed August 2017).

- UNECE. 2011. *Using Administrative and Secondary Sources for Official Statistics: A Handbook of Principles and Practices*. Geneva, Switzerland: UNECE. Available at: [https://www.unece.org/fileadmin/DAM/stats/publications/Using\\_Administrative\\_Sources\\_Final\\_for\\_web.pdf](https://www.unece.org/fileadmin/DAM/stats/publications/Using_Administrative_Sources_Final_for_web.pdf) (accessed August 2017).
- UNECE. 2012a. *Conference of European Statisticians, The Swiss Census System: a Comprehensive System of Household and Person Statistics*. Geneva, Switzerland: UNECE. Available at: [http://www.unece.org/fileadmin/DAM/stats/documents/ece/ces/2012/55-SP\\_Switzerland\\_-\\_the\\_swiss\\_census\\_system.pdf](http://www.unece.org/fileadmin/DAM/stats/documents/ece/ces/2012/55-SP_Switzerland_-_the_swiss_census_system.pdf) (accessed August 2017).
- UNECE. 2012b. *Conference of European Statisticians. Lessons Learnt from a Mixed-Mode Census for the Future of Social Statistics*. Geneva, Switzerland: UNECE. Available at: [http://www.unece.org/fileadmin/DAM/stats/documents/ece/ces/2012/37-SP\\_Germany.pdf](http://www.unece.org/fileadmin/DAM/stats/documents/ece/ces/2012/37-SP_Germany.pdf) (accessed August 2017).
- UNECE. 2015. *Conference of European Statisticians. From the 2010 to the 2020 Census Round in the UNECE Region – Plans by Countries on Census Methodology and Technology*. Geneva, Switzerland: UNECE. Available at: [https://www.unece.org/fileadmin/DAM/stats/documents/ece/ces/ge.41/2015/mtg1/UNECE\\_paper\\_Paolo\\_draft\\_0925\\_rev2.pdf](https://www.unece.org/fileadmin/DAM/stats/documents/ece/ces/ge.41/2015/mtg1/UNECE_paper_Paolo_draft_0925_rev2.pdf) (accessed August 2017).
- UNECE. 2016. *Conference of European Statisticians. Innovative Approaches Used in the 2016 Canadian Census*. Geneva, Switzerland: UNECE. Available at: [http://www.unece.org/fileadmin/DAM/stats/documents/ece/ces/ge.41/2016/mtg1/CES\\_GE.41\\_2016\\_16E.pdf](http://www.unece.org/fileadmin/DAM/stats/documents/ece/ces/ge.41/2016/mtg1/CES_GE.41_2016_16E.pdf) (accessed August 2017).
- United States Census Bureau (USCB). 2013. *America's Families and Living Arrangements: 2012*. Washington DC, US: USCB. Available at: <https://www.census.gov/prod/2013pubs/p20-570.pdf> (accessed August 2017).
- USCB. 2015a. *2020 Census Operational Plan*. Washington DC, US: USCB. Available at: <http://www2.census.gov/programs-surveys/decennial/2020/program-management/planning-docs/2020-oper-plan.pdf> (accessed August 2017).
- USCB. 2016. *Methodology for the United States Population Estimates*. Washington DC, US: USCB. Available at: <https://www2.census.gov/programs-surveys/popest/technical-documentation/methodology/2010-2016/2016-natstcopr-meth.pdf> (accessed August 2017).
- Wallgren, A. and B. Wallgren. 2014. *Register-Based Statistics: Statistical Methods for Administrative Data, 2nd Edition*. New York: John Wiley & Sons.

Received April 2016

Revised September 2017

Accepted October 2017