UNIVERSITY OF SOUTHAMPTON

FACULTY OF SOCIAL, HUMAN AND MATHEMATICAL SCIENCES

Social Statistics and Demography

Dispersal, deprivation and data: Asylum seekers and refugees since 1999

by

Sarah Louise Nurse

Thesis for the degree of Doctor of Philosophy

January 2019

UNIVERSITY OF SOUTHAMPTON

ABSTRACT

FACULTY OF SOCIAL, HUMAN AND MATHEMATICAL SCIENCES

Social Statistics and Demography

Thesis for the degree of Doctor of Philosophy

DISPERSAL, DEPRIVATION AND DATA: ASYLUM SEEKERS AND REFUGEES SINCE 1999

Sarah Louise Nurse

In 2019, the existing contracts for housing dispersed asylum seekers will come to an end, therefore a new system of asylum accommodation and support is currently being developed. This research investigates the policy of dispersal, which has been implemented in the UK since 2000, by applying rigorous demographic methods and principles to the available data in order to contribute to a better understanding of the asylum settlement process. In particular, it explores the relationships between dispersal, deprivation and individual outcomes in the context of limited data.

Firstly, patterns of dispersal and deprivation are mapped to show the geographic spread of asylum seekers by support status compared to Local Authority deprivation levels, using Home Office Asylum Statistics and the English Indices of Multiple Deprivation. Findings confirm that settlement locations of asylum seekers housed by the government are different from those on subsistence only support, and reflect the policy aim to move settlement away from London. A more formal assessment of these relationships through cluster analysis highlights a distinct group of Local Authorities with high levels of dispersal and high deprivation.

Analysis of the Survey of New Refugees identifies statistically significant differences between refugees who were and were not dispersed, but the context of high attrition and increasing time since collection (baseline surveys from 2005-07) limits its use moving forward. A systematic review of the feasibility of combining data on the refugee and asylum seeking population suggests that augmenting existing datasets, by adding an indicator of dispersal, has the potential to greatly increase the number of variables, and therefore the topics, available for analysis. Examples of this are illustrated using the Survey of New Refugees and Annual Population Survey data on reason for migration.

Table of Contents

Table of Tables

Table of Figures

Academic Thesis: Declaration Of Authorship

I, Sarah Louise Nurse, declare that this thesis and the work presented in it are my own and has been generated by me as the result of my own original research.

Dispersal, deprivation and data: Asylum seekers and refugees since 1999.

I confirm that:

1. This work was done wholly or mainly while in candidature for a research degree at this University;
2. Where any part of this thesis has previously been submitted for a degree or any other qualification at this University or any other institution, this has been clearly stated;
3. Where I have consulted the published work of others, this is always clearly attributed;
4. Where I have quoted from the work of others, the source is always given. With the exception of such quotations, this thesis is entirely my own work;
5. I have acknowledged all main sources of help;
6. Where the thesis is based on work done by myself jointly with others, I have made clear exactly what was done by others and what I have contributed myself;
7. None of this work has been published before submission.

Signed:

Date: 17th January 2019

Acknowledgements

I am grateful to the Economic and Social Research Council for recognising the importance of this topic and funding my Demography Masters and this PhD research, without which this would not have been possible. I must also acknowledge my use of data provided by the Home Office, the Office for National Statistics and the UK Data Service Secure Lab.

I would not have been able to complete this PhD without the invaluable supervision and guidance of Professors Jakub Bijak and Nyovani Madise. In particular, I would like to thank Jakub for his initial enthusiasm about my research proposal which continued throughout the project. His expert advice, generosity with his time and his commitment to helping me overcome unforeseen challenges has been greatly appreciated. I am also grateful to Nyovani for sharing her wisdom and experience in the form of valuable advice for navigating the PhD process as well as more specific comments and feedback in the later stages.

I would like to thank Li-Chun Zhang and Beata Nowok for their insightful comments and helpful suggestions on an earlier draft which have undoubtedly strengthened this work.

My family have always had faith in my ability to accomplish what I've put my mind to and have cheered me along the way. Thank you for helping to inspire my interest in people, for encouraging my desire to do something that I find meaningful and entering into the small achievements along the way.

Finally, to my husband Jon, thank you for your unwavering support and encouragement through the 'ups and downs' of the PhD journey, for always believing I would get there in the end, and for your patient and meticulous proofreading to help me over the finish line.

Abbreviations

ADRN	Administrative Data Research Network
APS	Annual Population Survey
AS	Asylum Seeker
ASYS	Asylum Support System Database
BIS	Department for Business, Innovation and Skills
CI	Confidence Interval
CID	Case Information Database
COMPASS	Commercial and Operating Managers Procuring Asylum Support
DCLG	Department for Communities and Local Government
DL	Discretionary Leave
ELR	Exceptional Leave to Remain
EMN	European Migration Network
EU	European Union
HC	House of Commons
HO	Home Office
HP	Humanitarian Protection
IMD	Index of Multiple Deprivation
LA	Local Authority
LFS	Labour Force Survey
LSOA	Lower Super Output Area
NAO	National Audit Office
NASS	National Asylum Support Service
OECD	Organisation for Economic Co-operation and Development

Abbreviations

ONS	Office for National Statistics
SNR	Survey of New Refugees
SO	Subsistence Only
SUNRISE	Strategic Upgrade of National Refugee Integration Services
UKBA	UK Border Agency
UKDS	UK Data Service
UNHCR	United Nations High Commissioner for Refugees
VPRP	Syrian Vulnerable Persons Resettlement Programme

Chapter 1 Introduction

1.1 Background and aims

Dispersal can be described as the spatial distribution of a population or group, usually controlled as the result of government policy. It is intended to actively counter the 'concentration' of minority communities (Robinson et al, 2003, p. 3), and since the Asylum and Immigration Act of 1999 has been applied to asylum seekers in the UK during the time of their application for international protection under the 1951 Refugee Convention (UNHCR, 2010). Government dispersal contracts (Commercial and Operating Managers Procuring Asylum Support, COMPASS) end in 2019, and as policymakers consider the design of the dispersal process, they are being urged to adapt and improve it for the benefit of future applicants, as well as receiving communities (HC Home Affairs Committee, 2017).

In an area of research which persists as a topic of paramount policy importance and public debate, as in the case of asylum, the need for rigorous, robust and reliable analysis is crucial. This will not only facilitate informed thinking but help to encourage measured discussion and considered decision-making around a subject which is often dominated by emotive discourse and reactionary strategies, aimed at tackling real or perceived problems and threats as they come to people's attention. More specifically, policy evaluation is necessary in order to:

- investigate whether policy aims are being achieved;
- ensure that resources are used effectively;
- suggest how policy could be better implemented;
- inform further decision-making and policy creation.

The academic community, chiefly in the areas of demography and social statistics, can play a key role in providing a strong evidence base for policymakers by applying formal statistical methods with a systematic approach, focusing on current gaps in knowledge and presenting findings in a clear and coherent way.

Asylum seekers and refugees experience different factors driving their move compared to other migration routes (such as study, work or family reunion) and are subject to a unique legislative framework and a controlled settlement process. Therefore, analyses which are tailored to the specific context of this group are required. That said, further disaggregation within the population of asylum seekers and refugees is essential (and often lacking in existing research), in order to differentiate between different entitlements and experiences of individuals assigned to different statuses (i.e. support status or legal status) at various stages of the process. Geographic detail and the analysis of spatial patterns are also central to the study of this population, as controlling settlement location is intrinsic to the dispersal policy.

When studying a population where available data is limited, it is important to thoroughly explore what is currently possible, utilising innovative methodologies, as well as making recommendations of what could be achieved if further data were made available or collected. The overarching aim of this research is therefore to investigate the outcomes of the policy of dispersal, which has been implemented in the UK since 2000, by applying rigorous demographic and statistical methods and principles to the available data in order to contribute to a better understanding of the asylum settlement process. In particular, the main objective of this work to better understand the relationships between dispersal, deprivation and individual outcomes, in the context of limited data.

1.2 Chapters and content

The opening Chapter 2 sets out the context for the analyses undertaken here by first presenting the policy and legislative background, including the UK asylum process and the dispersal scheme, followed by a brief introduction to data availability and description of historical trends. A review of the existing literature in Chapter 3 then provides an overview of previous studies on the dispersal policy, deprivation and integration of asylum seekers and refugees; gaps in the current state of knowledge are highlighted and inform the subsequent chapters.

The three substantive pieces of analysis are presented in Chapters 4, 5 and 6, which respectively focus on patterns of dispersal and deprivation, selected socio-economic outcomes amongst refugees in the UK and the potential for combining data on dispersal. In particular, the first piece of analysis (Chapter 4) presents the national picture of dispersal and examines how the geographic patterns compare to those of asylum seekers who are not dispersed, as well as the spatial patterns of deprivation. This is done through measures of inequality analysis, mapping techniques and cluster analysis to identify similar characteristics across Local Authorities (LAs) in England.

The second analysis (Chapter 5) explores the only available national-level quantitative refugee dataset, the Survey of New Refugees (SNR). This includes descriptive analyses as well as cross-sectional and longitudinal modelling, carried out in order to identify differences in background characteristics and outcome variables between refugees who were dispersed during the asylum process and those who were not. An important part of this analysis is to assess the nature of the attrition that is observed and the extent to which robust conclusions can be made in this context.

Finally, in recognition of the enduring knowledge gaps which exist in this field as a result of a scarcity of available data, the feasibility of methods for combining existing sources is assessed. The analysis carried out in Chapter 6 highlights a selection of methods which could considerably increase the potential of existing datasets, whilst also considering the trade-off between the errors which are introduced and the information gained. Examples of how data from the Annual Population Survey (APS) may be 'augmented' by borrowing information Survey of New Refugees are presented. A discussion directly addressing the implications of this research for current and future policy, both in terms of development and implementation, as well as recommendations for improving data availability, concludes this work.

Chapter 2 Background

The 1951 United Nations Refugee Convention was the first formal framework to set out the right of individuals outside their national borders as a result of a 'well-founded fear of persecution', to the protection of another nation state (UNHCR, 2010). The Convention was drafted in response to the mass displacement of people in Europe following the Second World War and the UK was one of the first signatories in September 1954. Initially only displacement as a result of events occurring within Europe before 1951 were recognised, but the 1967 Protocol later removed these temporal and geographic limitations and extended protection to those displaced across the world.

A refugee is defined by the Convention as any person who:

'owing to well-founded fear of being persecuted for reasons of race, religion, nationality, membership of a particular social group or political opinion, is outside the country of his nationality and is unable or, owing to such fear, is unwilling to avail himself of the protection of that country' (UNHCR, 2010 p. 14).

Migration for international protection differs from other forms of migration on a number of levels: not only are the reasons for individual decisions to migrate different from economic migrants, there are practical and legal differences which affect this group's experience of arrival and settlement in the UK. For example, application for permission to stay is usually made on or after arrival based on information gathered during interviews, rather than through visa applications in advance (as is the case for economic migrants or family reunification). While waiting for a decision applicants are defined as 'asylum seekers' and those who are granted leave to remain then become 'refugees'[1].

This chapter begins by outlining the broad picture of asylum policy and legislative changes in the UK since the 1990s, followed by a description of the dispersal policy introduced through the 1999 Immigration and Asylum Act, as well as subsequent developments in policy, its implementation and practice. Secondly, it assesses the current availability of data on asylum seekers and refugees in the UK, highlighting the gaps and limitations that still exist. Finally, a brief illustration of historic asylum trends is presented, along with an initial description of dispersal numbers, to give context for the subsequent analyses.

---

[1] Note that another category of permission to stay may also be granted for humanitarian or other reasons if an applicant does not qualify for refugee status (see: https://www.gov.uk/claim-asylum/decision as of 01/02/16).

2.1 Policy and legislation

Migration to the UK has been controlled since the Aliens Act of 1905, however, the 1993 Asylum and Immigration Appeals Act was the first piece of domestic legislation to incorporate the 1951 Refugee Convention[2]. At the same time, it also set out a standard of responsibility for housing authorities to provide adequate accommodation for asylum seekers, which was lower relative to the standards set out in homelessness legislation. Allowing Housing Associations to differentiate between those seeking asylum and the rest of the population signalled an era of increasingly restrictionist policies towards asylum support (Somerville, 2007).

Since the mid-1990s a new piece of legislation on immigration and asylum issues has been introduced every few years, reflecting a continuing concern about immigrant numbers and the political desire to be seen to be addressing this issue. Within these debates, asylum has been more or less prominent at different times. For example, in the two years before the 2005 general election, national newspapers repeatedly ran 'hostile' and 'alarmist' stories about asylum seekers, fuelling and promoting fear and xenophobia among their considerable combined readerships (Greenslade, 2005, p. 21). Greenslade (2005, p. 30) suggests that this was highly influential in putting asylum at the centre of the 2005 election campaign: 'a classic example of the press setting the political agenda'.

It is certainly the case that most policy development and legislative changes have focussed on restriction and deterrence, as well as some amendment of the application process and re-organisation of enforcement agencies. For example, the 1999 Immigration and Asylum Act introduced a system of dispersal, changing the focus of settlement away from London and removing the element of choice of location from asylum seekers; in addition, the right to work was removed through the Nationality, Immigration and Asylum Act of 2002[3]; and in 2004, the Asylum and Immigration Act was introduced to streamline the appeals process for those refused asylum. Friedman and Klein (2008, p. 58) argue that 'since 1996 a succession of laws has been introduced to deter asylum seekers from coming to Britain and, if they do arrive here, to support them at a level that British citizens would not find acceptable for themselves.'

[2] See Hynes (2011, pp. 10-11) for annotated chronology of British legislation relating to immigration and asylum.
[3] These rules were later amended so that an applicant who has not received an initial decision within 12 months may apply to work in the UK.

2.1.1 European policy and legislation

The UK asylum system should also be seen in the context of European Union (EU) law. The Dublin system - introduced initially at the 1990 Dublin Convention and replaced by later Regulations I, II and III[4] - prevented multiple claims from being made in different countries and allowed for the transfer of asylum seekers between member states. Applicants would be transferred when they were judged to have passed through a 'safe country'. Alongside Dublin was a process of 'harmonisation' which was being implemented through the creation of a Common European Asylum System (CEAS); the aim was to provide a common minimum standard in legal frameworks, to ensure 'fairness, efficiency, and transparency,' (Commission, 2007, p. 2) which would mean equal treatment for asylum seekers wherever they lodged their claim.

2.1.2 The 1996 and 1999 Acts and dispersal in the UK

The Asylum and Immigration Act of 1996 changed the support arrangements for those seeking asylum, removing any right to benefits for those who applied 'in-country'. The consequence of this was that many were left potentially destitute; LAs therefore became responsible for supporting those within their boundaries (Robinson et al. 2003). This resulted in some LAs - particularly in and around London - feeling that they had a disproportionate burden on their finances and services. This set the scene for changes introduced through the 1999 Immigration and Asylum Act, which included the creation of the National Asylum Support Service (NASS), a new division of the Home Office, to administer welfare and housing support to asylum seekers, removing their right to the mainstream benefits system.

More broadly, it was also hoped that the policy would act as a deterrent for those intending to seek asylum in Britain (Schuster, 2003). The two main challenges which were identified in the new Labour Government's White Paper, 'Fairer, Faster and Firmer', published in July 1998, were firstly the 'shambles' of the asylum support system, and secondly the 'abuse' of the system by economic migrants that had caused 'huge backlogs' (Home Office, 1998, Preface). Therefore the stated aims were to 'ensure that genuine asylum seekers are not left destitute, but … [to] minimise the

---

[4] Please see: *Convention determining the State responsible for examining applications for asylum lodged in one of the Member States of the European Communities – Dublin Convention* of 15 June 1990, OJ C 254, 19.8.1997, pp. 1–12, available at http://eur-lex.europa.eu/legal-content/EN/ALL/?uri=CELEX:41997A0819%2801%29 and for the most recent Regulation (Dublin III), *Regulation (EU) No 604/2013 of the European Parliament and of the Council of 26 June 2013 establishing the criteria and mechanisms for determining the Member State responsible for examining an application for international protection lodged in one of the Member States by a third-country national or a stateless person*, OJ L 180, 29.6.2013, pp. 31–59, at http://eur-lex.europa.eu/legal-content/EN/ALL/?uri=CELEX:32013R0604.

attractions of the UK to economic migrants' (Home Office, 1998 Preface). The Government was attempting to design a policy which would balance these two – potentially conflicting – objectives. The inherent dilemma faced here was the underlying desire to provide generously only to asylum applicants who would prove to be 'genuine' when their case was decided upon, despite the impossibility of determining this before that decision is made.

Under the resulting system of dispersal implemented since the Act, housing has been provided in LA areas with no element of choice from the person seeking asylum. As Chapter 3 will introduce, it has been argued that moving the focus of settlement from the South East to the North, Midlands and Scotland indicates a broad shift from areas of relative affluence to high deprivation levels.

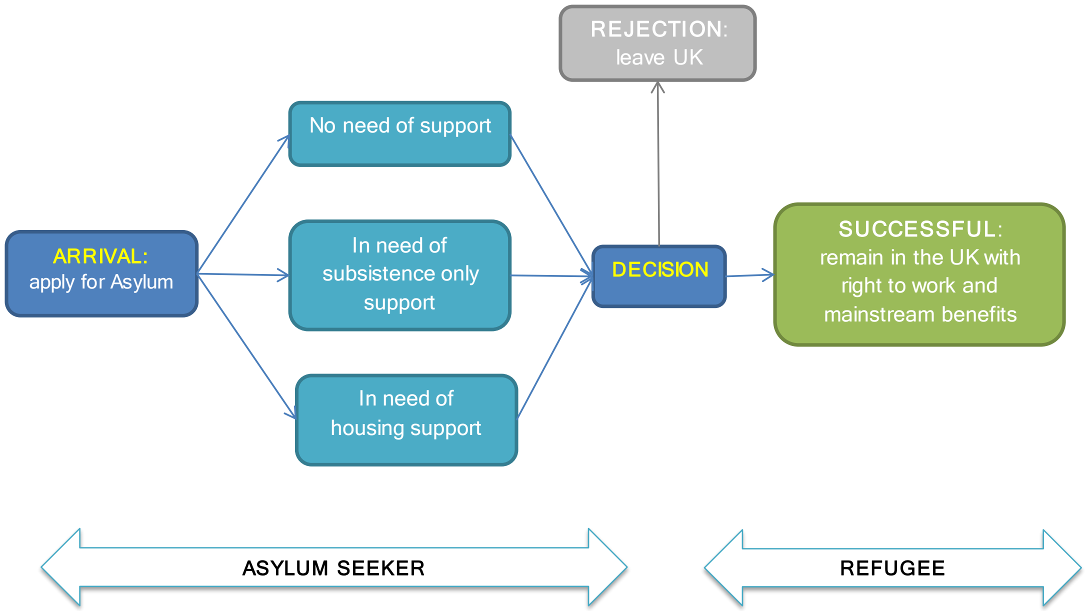
2.1.3 Definitions and the UK system

The process of application usually follows the broad outline represented in Figure 2.1. This diagram is a simple representation of the asylum process and support system that has been in place since the 1999 Act. A person is defined as an asylum seeker from the time that an application for asylum is made (often at or a short time after arrival) until the final decision on whether the application is successful. In theory, those who receive a 'rejection' will then leave the country; in practice, many will go through a lengthy process of appeals, or may remain illegally. Under the current rules[5], applicants that are not rejected outright may be given permission to stay under one of three categories: as a refugee, for humanitarian reasons, or for other reasons.

The government aims to make decisions on asylum cases within six months of the submission of an application. In practice this waiting time varies dramatically between individuals and at times of greater applicant numbers. Official statistics show that since 2010 (when these data were first published), the percentage of applicants with initial decisions pending for more than six months has varied from 20 to 64 percent (Home Office, 2018). The total number of decisions pending each quarter has ranged from 5,947 to 24,213. The length of time spent waiting for initial decisions and subsequent decisions on any appeals will impact on the duration of stay of those receiving housing support in dispersal accommodation, although data on this is not published.

[5] See https://www.gov.uk/claim-asylum/decision.

Figure 2.1: The asylum process



Source: Author's own creation.

For the purposes of this research, an individual who has received a decision allowing them to remain in the UK under any of these three categories (is 'successful') will be treated as a 'refugee' unless otherwise stated. Under permission to stay as a refugee or for humanitarian reasons, 'leave to remain' or 'leave to enter' is given for five years, after which an application may be made to settle in the UK indefinitely. Length of permission to remain granted for other reasons varies according to individual circumstances.

2.1.4 Subsequent developments, implementation and practice

Since 1999, the body responsible for administering asylum support has undergone multiple reorganisations: the NASS ceased to exist as a directorate due to Home Office restructuring in 2006 and all asylum support issues were then dealt with and processed by the Home Office's Border and Immigration Agency (BIA), which in turn became the UK Border Agency in 2008 until this was also abolished in 2013; UK Visas and Immigration then took over, working back within the Home Office. There have also been developments in who is responsible for providing housing, with contracts increasingly moving to the private sector. Burnett (2011) argues that since the recession there has been a 'gradual shift in dispersal policies: one which has seen LAs abdicating, or being absolved of, their responsibilities to house and provide shelter for asylum seekers'.

From 2000, the 1999 Immigration and Asylum Act was implemented with NASS entering into contracts with LAs across the UK to house dispersed asylum seekers. Most of this accommodation came from a combination of hard-to-let social housing and private landlords sub-contracted by

LAs (Darling, 2013). However, contracts with LAs from 2000 were renewed from 2006 with increasing use of private providers. Further re-negotiation of contracts in 2011 resulted in the provision of dispersal accommodation being procured solely from three private providers: G4S, Serco and Clearsprings Group. The six new regional contracts (collectively known as Commercial and Operational Managers Procuring Asylum Support Services or COMPASS) have been found to result in 'patchy' and 'poor quality' provision of housing, with considerable delays and problems with over-expenditure (House of Commons Committee of Public Accounts, 2014).

Within this system, LA approval is required for providers to acquire dispersal accommodation within their boundaries. In practice, a LA has 72 hours to raise any specific objections to a proposed property (or properties) proposed by a provider, and even then the Home Office can decide to override them, if the LA has otherwise broadly agreed to housing asylum seekers (HC Home Affairs Committee, 2017). According to John Whitwam, Managing Director of G4S Immigration and Borders, wide-scale refusal of LAs to allow providers to acquire properties to house asylum seekers, or even to respond to requests, has hindered attempts to broaden the geographic spread beyond existing dispersal locations (HC Home Affairs Committee, 2017 para. 36). The Government has reportedly written to LA leaders across the country to encourage them to participate in the dispersal scheme, but in spite of the 1999 Immigration and Asylum Act giving the power for this to be enforced, this is yet to be exercised (Politowski and McGuinness, 2016).

The intention for dispersal was to create 'clusters' where asylum seekers sharing the same languages and countries of origin would be housed together (NAO, 2014, p.11). It was hoped that this would provide community support networks and access to economic infrastructure as well as enabling efficient and effective provision of services and information (Zetter et al, 2002). However, in practice this did not happen, largely due to the volume of demand for temporary housing (Friedman and Klein, 2008) and service provision, such as health information, has reportedly been difficult as a result (Johnson, 2003). The Yorkshire and Humberside Consortium for asylum seekers and Refugees Integration Strategy (2003, p. 32) highlights that 'concentrations of asylum seekers have placed some local health services under strain', but also notes that 'it has been estimated that approximately 10% of refugees have had some form of medical training,' and suggests that these skills may be utilised to help ease the pressures.

Changes to the contracting process and the resulting organisational structures based around COMPASS have had an impact on how dispersal is implemented across the country. For example, in Manchester, the fact that a private company (Serco) have been the sole provider of dispersal accommodation since 2013 has resulted in the identification of some key risks, including that there may be a 'disproportionate placement of asylum seekers in the most deprived areas', according to a report for the communities scrutiny committee (Manchester City Council Communities Scrutiny Committee, March 2014). The suggestion that reception areas have typically been deprived urban areas where there are large supplies of vacant housing has also been identified in wider literature (e.g. Phillimore and Goodson, 2006).

2.1.5 European context

Policies to disperse asylum seekers and refugees have been implemented in a number of ways in different countries where concern about spreading the 'burden', in particular the financial and social 'costs', is the main motivation for introducing dispersal (European Migration Network (EMN), 2014, p. 3). For some countries this extends to encouraging long-term settlement of those with right to remain in particular regions or localities.

The development of policy as a response to evolving pressures and to address problems identified during the settlement process has been seen repeatedly since the 1990s. For example dispersal policy in the Netherlands emerged in order to centralise control of initial housing of asylum seekers, but it has shared the UK experience of friction between national policy and local agreement for implementation, with controversy around choice of locations and the subsequent impact on communities and services (Robinson et al, 2003). Similar issues have been observed in Sweden with the introduction of dispersal aiming to shift the focus of refugee settlement away from large urban areas; similar concerns around integration and onward migration following enforced dispersal have prompted ongoing debate and further policy changes over time.

2.2 Data

The recording of absolute numbers of asylum seekers is inherent in the system, as an application must have been submitted to the authorities in order for an individual to be defined as such. These data are published in the regular quarterly Home Office Asylum Statistics, along with other characteristics (such as age, sex, nationality), but data on timing and transitions between statuses are lacking. The absence of information on length of time between stages of the asylum process limits the analytical potential considerably, as does the lack of detail available for earlier years (for example nationality breakdown is only available from 2001). Data are also published on numbers of asylum seekers in receipt of housing or 'subsistence only' support, either by LA or nationality; understanding geographic variation is essential to allocate resources and effectively implement policy in an efficient, fair and transparent way, but data are not published below LA level.

The only administrative data available on asylum seekers after they have received their final decision is a record of the numbers deported or who have left the country through notified/assisted voluntary return. Data is not collected on how many remain in the country or leave without informing the authorities, whether their application was accepted or not. Data on the refugee population more generally is notably scarce (Stewart, 2004). In an attempt to address the data limitations recognised in assessing asylum and refugee policy, the Survey of New Refugees (SNR) was commissioned by the Home Office, primarily to provide information on the integration of new refugees between 2005 and 2007, and to assess the SUNRISE[6] intervention (Home Office, 2010a-c). The SNR is described and analysed in detail in Chapter 5, particularly in relation to what the data can tell us about differences in background characteristics and outcomes that are observed between dispersed and non-dispersed refugees. Furthermore, since 2010 the Labour Force Survey (LFS) has included a question on 'main reason for migration' (Home Office, 2014). These datasets can provide some information but the considerable issues often apparent in survey data mean that potential for analysis is limited.

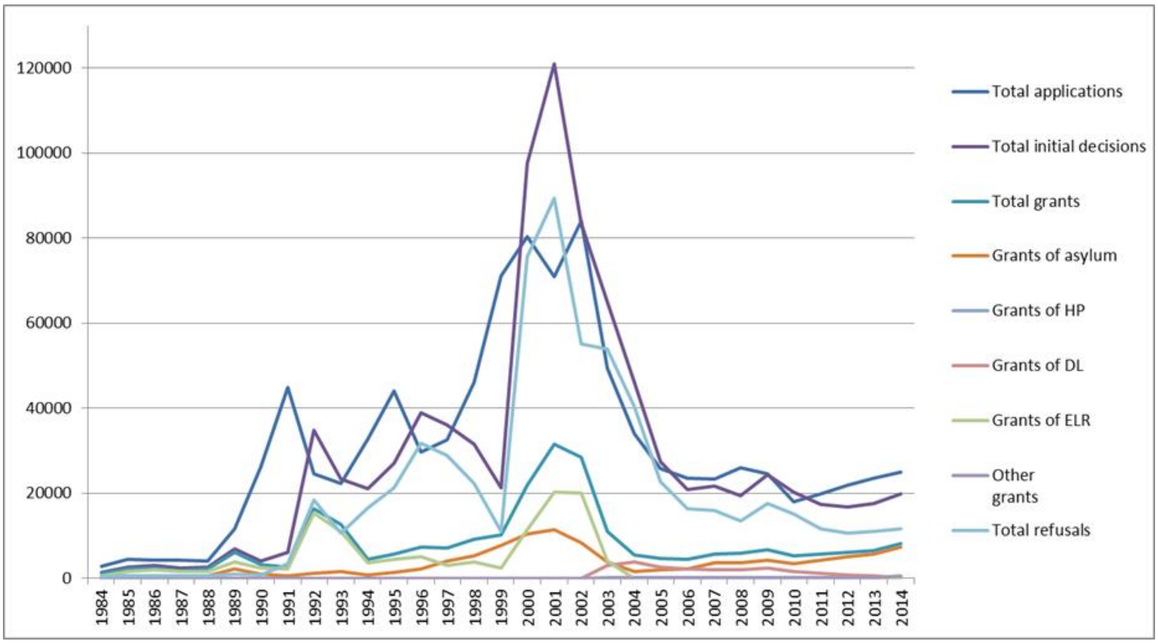
A number of qualitative studies have collected data on refugees' experiences of the UK asylum process, support system and life as they settle in a new country. These are discussed in more detail in the Literature Review chapter but have focussed on local level, small scale studies, little of which have been published in full for further analysis; this will therefore only be utilised to the extent that it adds to the body of research and adds context to the following quantitative analyses.

2.2.1 Historical trends

The number of displaced people throughout history has varied depending on factors such as the existence, as well as the nature and intensity, of conflict. The vast majority of refugees still live in neighbouring countries within the region of conflict, but as transport links have proliferated throughout the 20th and 21st Centuries, there has been an increase in those travelling considerably greater distances to seek protection. The number of refugees globally increased from 2.4 million in 1975 to 14.9 million in 1990 (Castles and Miller, 2009). Home Office Statistics presented in figure 2 show considerable fluctuations in the patterns of asylum applications and outcomes in the UK over the last three decades.

[6] The Strategic Upgrade of National Refugee Integration Services (SUNRISE) pilot scheme was intended to support refugees in the integration process; however not enough participants responded to allow meaningful analysis of its impact.

Figure 2.2: Annual asylum applications, decisions, grants and refusals in the UK, 1984 to 2014.



Source: Author’s creation from Home Office Asylum Statistics (Home Office, 2014b).

Note: Humanitarian Protection (HP) and Discretionary Leave (DL) replaced Exceptional Leave to Remain (ELR) in April 2003. Applications received in-country were first differentiated from applications received at ports in 2000.

A changing world context with different regions experiencing conflict and instability, along with increasing availability of transport as well as changes to asylum policy and legislation in the UK, are all factors which influence the variation observed in the official figures. For example, the number of applicants rose in the early 1990s with the breakup of the Soviet Union and conflict in the Former Yugoslavia and the peak in the total number of initial decisions in 2000-2001 reflects that clearing the backlog of asylum cases had reached the top of the New Labour policy agenda (Home Office, 1998). The subsequent reduction in applications, decisions and grants is largely a result of the introduction of progressively more stringent controls, limiting of asylum support and greater focus on removals (Somerville, 2007). During the period covered in the first two analysis chapters, 2005 to 2010, there were 141,086 asylum applications in total and 134,044 initial decisions (Home Office, 2018). While there is some evidence that increasing the focus on deterrence can reduce asylum flows, an unintended consequence may be a rise in the levels of undocumented arrivals (Czaika and Hobolth, 2016). In addition to restriction and the aim to differentiate those ‘genuinely in need of protection’ from ‘economic migrants’, in the UK during this decade, there was a desire to better understand and encourage refugee resettlement and community cohesion (Home Office, 2002).

Chapter 3 Literature Review

The multifaceted nature of this topic and the limited data available for analysis means that an accumulation of research from different academic disciplines as well as input from practitioners, charities, policy documents and official reports is crucial, not only to give important context for the current study, but also to contribute to the overall knowledge. This chapter reviews the available literature, firstly giving a brief theoretical background to the development of dispersal and understanding of deprivation, highlighting qualitative research that has been done into the relationship between the two and studies undertaken at a local level. The discourse on refugee integration is then presented, identifying important indicators and highlighting the limitations of existing quantitative analyses.

3.1 Dispersal and deprivation

Since the 1960s dispersal has been regarded by many (e.g. Cullingworth Committee, 1969, cited in Robinson et al. 2003) as desirable for race relations and the integration and improvement of prospects for ethnic minorities. Before 1999, it was used in the reception and settlement of specific groups of migrants arriving in the UK as a result of conflict or persecution. For example, Ugandan Asians arriving in 1972 were dispersed to LAs that volunteered vacant housing; Vietnamese refugees were also received in the late 1970s and early 1980s with the intention of housing them in ‘clusters’ of between four and ten families (Kushner and Knox, 1999, p. 318). However, Kushner and Knox (1999) suggest that the planned dispersal failed for these groups as the geography of settlement was dependent on which housing authorities offered accommodation.

Dispersal has developed as a response to the ‘problem’ of financial and social ‘costs’ resulting from refugee settlement, particularly when concentrated in specific localities (Robinson et al, 2003, p. 23). Policies have been based on what respective governments believe will be acceptable to the British public and justified by reference to the morality of ‘burden sharing’ in order to protect local populations from a disproportionate share of the costs and to protect their cultural identities (Robinson et al, 2003, p, 23). Bloch and Schuster (2005, p. 493) have argued that the lack of freedom asylum seekers have to choose where they settle in Britain under the dispersal policy means that they may be separated from crucial community organisations, kinship and other social networks, which can leave them marginalised and socially excluded.

Furthermore, it has been observed that many dispersal locations feature in the list of the 88 most multiply deprived districts in England (Hynes 2011, p. 71), and that the pattern of dispersal is related to the availability of unpopular or low-demand accommodation (Zetter et al, 2002; Robinson et al, 2003; Friedman and Klein, 2008). Townsend (1987, p. 126) describes deprivation as a relative condition, when the standard of living of an individual is below the 'socially accepted or institutionalised' level of the wider population. Multiple factors are identified as contributing to deprivation, including a lack of diet, housing, fuel, environment, education, access to working and social conditions, activities and facilities 'which are customary in the society to which they belong' (Townsend, 1987, p. 126). It is this definition of deprivation which underpins the indices of multiple deprivation which have been published by the Department for Communities and Local Government since 2000 (DCLG, 2008; DCLG, 2011). The indices incorporate seven main types of deprivation – income, employment, health, education, housing and services, living environment and crime – which are combined to produce a single, relative measure (DCLG, 2011).

The link between dispersal and deprivation is further highlighted in a Home Office report on factors affecting successful dispersal (Anie et al, 2005). It found some evidence of a positive correlation between dispersal locations with higher proportions of dispersed asylum seekers and higher levels of vacant housing, as well as a higher proportion of residents in 'social grade E' (i.e. on state benefit, unemployed or in the lowest-grade jobs). The main focus of the Home Office study was to assess what factors are associated with successful dispersal; that said, the level of 'success' was measured solely by the number of incidents of verbal and racial harassment and physical assault of asylum seekers reported to NASS, and was based on very low counts. Incidents of harassment clearly represent a negative experience for an individual but this is an extremely narrow remit for assessing the success of dispersal: an absence of incidents does not indicate successful dispersal; therefore Anie et al (2005) recommend further research.

Hynes (2011) carried out in-depth qualitative research investigating the presence of a wide range of experiences and factors that may impact on the settlement process. Dispersed asylum seekers and refugees were interviewed with the aim of understanding their experiences of dispersal. Interviewees were asked to describe the neighbourhoods where they had been located and other details of their lives. The findings reported suggest a link between dispersal and indicators of social exclusion, such as the presence of crime and a lack of access to services. It should be noted that it is not possible to identify from this research whether the evidence gathered indicates a lack of the *presence* of service provision, or an inability of asylum seekers to access those services. Nevertheless, Hynes (2011, p. 71) argues that it was a focus on 'bed-space' – mainly to be found in multiply deprived areas – which 'was, and remains, key to the reinforcement of formal and informal social exclusion of asylum seekers in the UK'. Unfortunately, a systematic description of methods (for example, how many people were interviewed at how many locations) is not reported but interviews appear to have been carried out largely in 2003, only three years after dispersal was first implemented. The picture of dispersal over a decade later is likely to be different; a decrease in asylum applicants, a change of government and policy, as well as major economic change, could all have an impact on how this policy is implemented and the resulting outcomes.

Case studies looking at the locations of asylum seekers and refugees at a local level have found evidence that there tend to be concentrations living in deprived wards. For example, Phillimore et al (2003) used a range of sources to map the locations of asylum seeker and refugee communities and compare these with the locations of education services in Coventry and Warwickshire. It provides an insight into the barriers to education and employment experienced by asylum seekers and refugees living in the region. Although this research can contribute to the understanding of where asylum seekers and refugees are living across the region, it is clear that this is not a homogenous population; therefore, it is not possible to draw conclusions about the national policy of dispersal or to identify where dispersed asylum seekers are living, or how this pattern compares to levels of deprivation.

Further research was carried out into the patterns of asylum seeker and refugee settlement using mapping techniques in two other study areas: Birmingham and Solihull (Phillimore et al, 2004) and the Black Country[7] (Goodson et al, 2005). These, along with Coventry and Warwickshire, are discussed further in relation to deprivation more broadly by Phillimore and Goodson (2006): the findings suggest that in the West Midlands clusters of asylum seekers and refugees are in the most deprived wards, and that these are some of the most deprived parts of Britain (Phillimore and Goodson, 2006, p. 1,722). Again, this analysis is based on data for one region, but in some general comments on the context of the dispersal policy, it is argued that the availability of housing was the most important criteria used to identify dispersal locations, and it was this which drove the practice of housing asylum seekers in highly deprived areas (Phillimore and Goodson, 2006). Furthermore, a recommendation is also made that the economic potential of refugees to contribute to existing deprived areas should be recognised, highlighting the multi-directional relationships involved in the process of refugee settlement.

Detailed local analysis and qualitative studies can make an important contribution to the understanding of the relationship between dispersal and deprivation, but findings are not generalisable to the national picture resulting from this policy. This research can clearly contribute

[7] Including Wolverhampton, Walsall and Sandwell.

important context and evidence towards understanding the mechanisms involved and experience of individuals, but a higher level approach to assess the national picture is also essential.

3.2 Integration

The analysis of the social and economic outcomes of immigrants are often framed within the discourse on integration, however, this concept lacks a widely agreed definition. One approach taken by Ager and Strang (2004) presents a framework for measuring the success of integration policies and projects by identifying factors which contribute to the process of integration. Measurable indicators are framed within ten key domains and include outcomes and means as well as indicators which provide a foundation for or facilitate integration. An attempt to capture refugees' experiences of the process is also included through measures of social connections and relationships. Achieving equity of outcomes with the host population is a theme in Ager and Strang's (2004) paper and is also picked up by Phillimore and Goodson (2008), who suggest that this is lacking in the Home Office (2005) definition, which focuses more on assimilation, rather than the 'bi-directional features' of the settlement process (Phillimore and Goodson, 2008, p. 311). Phillimore and Goodson (2008 p. 312) do recognise, however, that the indicators pursued by the Home Office 'represent the Government's functional policy vision on integration'.

The Home Office (2005, p. 81) national integration strategy focuses on 'what can feasibly be achieved' and also relies on analysis of a set of indicators, aiming for demonstrable 'improvements' in the outcomes of refugees so that they 'more closely match those of the people in the communities in which they are living,' rather than achieving equity with the broader UK population.

The findings (Home Office, 2005, p. 81) suggest that indicators for achieving full potential are:

- the employment rates of refugees; and
- levels of English language attainment over time.

For contributing to the community, the indicators are:

- the number of refugees involved in voluntary work;
- the numbers of refugees, and their children, in touch with community organisations (including local groups and wider community life);
- the proportion of refugees taking up British citizenship once they are qualified to do so; and
- the proportion of refugees reporting racial, cultural or religious harassment.

Indicators for accessing services are:

- the rates of access to housing services by refugees; and
- the proportion of refugee parents indicating their satisfaction with the education received by their children.

The intention is to measure progress in these indicators in order to monitor the 'effectiveness of local projects and policy interventions to remove barriers to integration' (Home Office, 2005, p. 81).

Phillimore (2008) also carried out qualitative interviews in order to assess which indicators were considered most important by refugees themselves. It was found that employment and housing were both considered crucial in order to enable advancement in other areas such as language development, cultural understanding, accessing services and a feeling of security. These two areas were also found to be inadequate for many refugees who experienced extremely high rates of unemployment, a lack of secure housing and high levels of homelessness. However, it should be noted that the population sampled may be skewed towards those who have more difficulties integrating and are therefore more likely to be visible and continuing to identify as a 'refugee'. Finally, the research points to the need for individual level analysis in order to better understand how indicators vary in importance for individuals with different characteristics, such as age, sex, and ethnicity.

There is clearly considerable literature on integration which aims to identify indicators that will measure the extent to which integration has occurred; however, the quantitative data on characteristics and outcomes relating specifically to asylum seekers and refugees that is necessary to assess the progress of these indicators has been long recognised to be inadequate (Stewart, 2004). Therefore rigorous quantitative analysis of integration for this group was not possible until the publication of the Survey of New Refugees in 2010 (Home Office, 2010a-c). This survey, commissioned by the Home Office, was an attempt to address this gap in knowledge on a topic of intense interest within academic and political spheres, as well as in the media and public debate.

Cheung and Phillimore (2013) carried out the first substantial exploration of integration issues using the survey data. Their aim was to utilise the unprecedented breadth of information that this dataset provides on the settlement experience, background characteristics and outcome variables of new refugees in order 'to increase understanding about the role of social networks and social capital in refugee integration' (Cheung and Phillimore, 2013, p. 3). The complexity of the relationships between indicators meant that the available variables needed to be organised into a useable framework. The analysis shows that housing was consistently found to be important in terms of integration outcomes observed in the SNR. In particular, it was noted that living in NASS accommodation at the baseline wave of the survey was associated with worse outcomes, such as an increased likelihood to not be employed, to have less contact with family and friends and poorer health in following waves. In spite of the ambitious goals of this paper – to disentangle the relationships between social capital, social networks and integration – the analysis does not venture much beyond descriptive presentation, considering the relationships between pairs of variables with only limited consideration of what can be deduced from the patterns observed across the waves.

The Home Office (2010a-c) reports published alongside the SNR data summarise the findings of their analysis and describe the importance of demographic and background characteristics (such as age, sex and country of origin, previous education and employment), as well as language skills, health and time spent in the UK, in influencing integration. Cheung and Phillimore (2013) recommend that integration can be promoted through improving access to good quality language training, encouraging social networks (for example family reunification) and ensuring that refugees have access to employment and financial support. While this research can contribute to the body of knowledge on integration, it is important to recognise that the usefulness of the SNR is limited by issues common in longitudinal survey data; in particular, there is a high rate of attrition, meaning that the majority of interviewees had dropped out by the later waves. This study does also touch on the link between the asylum experience and refugee outcomes, suggesting that allowing asylum seekers to choose dispersal locations with friends and family as well as actively providing protection from harassment could support integration (Cheung and Phillimore, 2013).

Zetter et al (2005, p. 176) highlight the importance of distinguishing between experience and entitlements before and after receiving permission to remain. It is argued that dispersal represents a 'key instrument of deterrence', treating asylum applicants as 'temporary' residents, and therefore keeping them 'disconnected from the modalities of integration and the support services needed for settlement after a positive status determination, such as access to housing, welfare and unemployment benefits' (Zetter et al 2005, p. 176). Employment is repeatedly identified as important for facilitating positive settlement; Phillimore (2011, p, 578) described this as 'acculturation', defined as 'the changes that happen to groups and individuals when two different cultures meet,' and employment is central to this process. Length of time spent awaiting a decision during the asylum process is also found to be a key factor, and the two may not be unrelated: the right to work is removed from asylum applicants until a positive decision is received. Having a job not only provides an income but also facilitates interaction with others, improving self-esteem and language skills. The longer individuals spend in limbo without access to settlement support, the more likely that his will have a ‘knock-on effect’ on outcomes once they receive the right to remain (Phillimore, 2011, p. 589).

It is clear that legal status determines rights to residency as well as access to employment, education and healthcare. Moreover, access to these services have been identified as ‘means and markers’ in the measurement of integration and therefore differences in legal status must be central to any analysis of the settlement process and individual outcomes (Da Lomba, 2010, p. 149). This supports the case for research which differentiates between asylum seekers and refugees, but which considers experience in both states as crucial to understanding the settlement process and individual outcomes. Da Lomba (2010) argues that integration starts on arrival, and therefore stresses the importance of experience spent as asylum seekers, which can have a long term impact on integration. It is argued that the problem is further exacerbated for individuals by a time of enforced economic inactivity, which can ‘undermine their future employability’ and therefore their prospects for integration as refugees (Da Lomba, 2010, p. 424). That said, it is important to recognise that while the restrictions on asylum seekers may limit integration, the gaining of access to many more rights on the acquisition of refugee status does not guarantee or inevitably result in the removal of non-legal barriers.

Much of the available literature on the subject of refugee settlement (e.g. Phillimore and Goodson, 2008; Ager and Strang, 2004; Home Office, 2005; Cheung and Phillimore, 2013) discusses the relevance of indicators of integration. The focus of these reports tends to be on how the integration of refugees can be conceptualised and measured; the experiences of asylum seekers as a distinct group, or as the state preceding refugee status, are often only briefly considered, and limited attention is given to the impact of dispersal on subsequent refugee settlement. The Home Office (2005) refer to dispersal, but only to broadly recognise it as the context within which refugee integration must occur, with much of the focus on methods to ensure that refugees do not leave their ‘original dispersal area’. When seeking to understand asylum seeker and refugee populations and policy implications, it is important that the particular situations and different experiences and entitlements before and after receiving a decision are central to any analysis.

While there is wide variation internationally in the migration histories, social context and policy frameworks of receiving countries, some comparison can be useful to highlight good practice and lessons that can be learned with regard to integration of refugees. Particular barriers identified in OECD (Organisation for Economic Co-operation and Development) countries include a lack of recognition for qualifications and skills, difficulty accessing the labour market and receiving appropriate healthcare (OECD, 2016). Furthermore these are often exacerbated by a lack of early provision of integration support from the state, for example with language skills training. A key recommendation in the OECD (2016) report on integration is to consider employment prospects within dispersal policies. One example is New Zealand where educational and employment opportunities are taken into account in the decision of settlement location, based on the existing skills and experience of refugees. It is suggested that dispersal policies that overlook potential employment prospects, for example in Sweden and Denmark, result in worse outcomes for those dispersed compared to those who are not dispersed, or who move on to a different location.

Another element that is crucial to the settlement and wellbeing of asylum seekers and refugees is community acceptance. Stewart (2011, p. 14) found that 'given the choice, asylum seekers will forgo NASS support in favour of moving to what they regard as more inclusive neighbourhoods'. Media reporting, and local press in particular, can set a 'framework' for how local residents react to the arrival of asylum seekers, by representing and at the same time helping to construct local identity (Finney and Robinson, 2008, p. 397). Robinson et al (2003, p. 167) go so far as to suggest that dispersal is a response to 'moral panic' about concentrations of asylum seekers or refugees in localities, and that positive media reporting and efforts to counter negative representations can tackle the real 'problem', rendering dispersal policies unnecessary.

3.3 Conclusion

Quantitative analysis of the relationship between dispersal and deprivation at the national level is limited, and although it is widely reported that there is an association, the available evidence published to date is far from conclusive. Without further analysis at the national level it is not possible to make conclusions (and subsequent recommendations) which relate to this national policy. It is also evident that a clear distinction must be made between the different legal statuses of asylum seekers and refugees; individual level analysis which investigates the relationship between experiences in the two states will allow better understanding of the settlement process. Finally, in order to assess and measure integration of refugees, it is important to carry out longitudinal analysis so that change over time can be observed and the relationship between background characteristics, settlement experience and social and economic outcomes can be thoroughly explored, with a particular focus on differences between those who were dispersed or not during their asylum application.

Chapter 4 Patterns of Dispersal and Deprivation.

4.1 Introduction

The policy of dispersal introduced through the 1999 Immigration and Asylum Act had two main aims: changing the geographic focus of settlement and restricting the flow of asylum seekers through deterrence. When assessing this policy, it is crucial to consider how these aspects have shaped the resulting geography of settlement.

It is clear from the examination of existing literature that a general assumption exists, based on qualitative reporting and local case study analysis, that there must be a positive relationship between dispersal areas and high deprivation (see for example Hynes, 2011; Phillimore and Goodson, 2006). The cumulative weight of this research is compelling, but nevertheless, quantitative evidence of an association at a national level is still notably lacking. Furthermore, there has been a tendency to focus research on 'ASRs', asylum seekers and refugees, which unfortunately masks some important differences in experiences, rights and characteristics within this population. This chapter presents a national picture of asylum seekers and deprivation, to the extent which this is possible from published asylum data, with special consideration of the different patterns observed for dispersed and non-dispersed asylum seekers.

The aim of this chapter is to describe and analyse variation in the patterns of dispersal and deprivation across England, and further understand the relationship between the two. This is achieved through mapping of the data and inequality analysis, as well as the application of cluster analysis, grouping LAs with similar characteristics in order to shed light upon how the dispersal policy has been implemented.

Two main research questions are considered:

1. How do patterns of dispersal, subsistence only support and deprivation vary across the country?
2. What is the relationship between rates of dispersal and deprivation at a Local Authority level?

In approaching the first research question, mapping techniques are used to provide an initial description of the patterns which can be observed across England. Mapping asylum support data enables comparison of the settlement geography of dispersed asylum seekers and those on 'subsistence only' support with patterns of high deprivation. Measures of inequality are also used to show variation across the country. This is followed by a more formal assessment of these relationships, using cluster analysis to address the second research question classifying LAs based on levels of deprivation and rates of dispersed asylum seekers.

The analyses in this chapter focus on two years: 2005 and 2008. There are a number of reasons for this, firstly there are limitations in the availability of data; in particular, the Indices of Deprivation which are central to this analysis have only been produced for selected years. Analysing data for two years allows comparisons to be made which highlight similarities and differences as well as ensuring that robust conclusions can be made. Finally, from a policy perspective this choice of years will allow some assessment of the impact of changes over time, for example with the increase in private provision of dispersal housing. Furthermore, the nature of the deprivation indices used here means that they are only calculated for England and are not comparable with the indices for other countries[8]. The majority of asylum seekers have been dispersed across England; in 2005, 87 of the 354 LAs[9] in England, and 90 in 2008, housed at least one asylum seeker in dispersal accommodation. Only a small number of LAs in Scotland, Wales and Northern Ireland are affected (see Appendix A), meaning that analysis of the settlement geography would provide limited additional information; therefore the analysis in this chapter will focus on LAs in England. Nevertheless, it can be noted that the vast majority of those dispersed to Scotland, Wales and Northern Ireland are housed in large cities such as Glasgow, which is the most deprived LA in Scotland, and therefore not suggesting a different pattern to those observed in England.

In this chapter, an initial description of the data sources, their availability and limitations is followed by an outline of the methodology which is to be applied. The results of the analyses are then presented, with the use of maps being central to highlighting and understanding patterns identified in the data. These results are then discussed in the context of previous research and literature alongside in-depth analysis of how the patterns identified relate to policy debate and implementation intentions and practice, as set out in Government and LA documents and reported by stakeholders. Finally, some concluding remarks identify the wider implications of these findings.

[8] Scotland (see: http://www.gov.scot/Topics/Statistics/SIMD), Wales (see: http://gov.wales/statistics-and-research/welsh-index-multiple-deprivation/?lang=en) and Northern Ireland (see: http://www.nisra.gov.uk/deprivation/nimdm_2010.htm) have their own indices of multiple deprivation.
[9] Unless stated otherwise, results are reported using the boundaries in place before the 2009 changes. These are the same for 2005 and 2008, allowing direct comparison and identification of changes occurring over this time (see Appendix B for more details).

4.2 Data

4.2.1 Asylum statistics

Immigration statistics for the United Kingdom are collected by the Home Office[10] and published quarterly in February, May, August and November[11]. These include: data on applications and decisions for main applicants and dependents; information on nationality, age and sex of applicants; and support status of applicants by LA. The official statistics report aggregated numbers rather than individual level data, and either the number of events occurring in the relevant quarter (e.g. applications) or the total population as at end of the quarter (e.g. supported in dispersal accommodation) (Home Office, 2015).

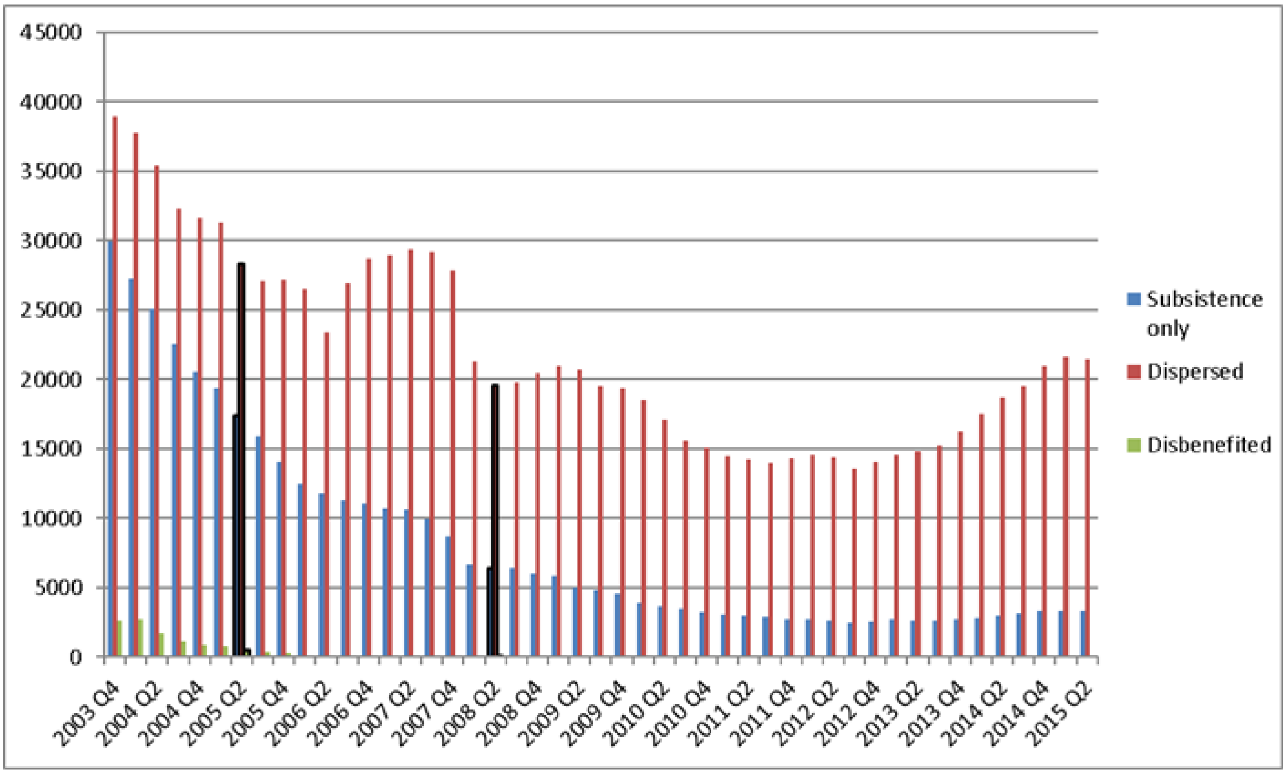
Figure 4.1 shows official asylum statistics reporting the support status of all those with applications pending at a given time each quarter since 2003. Those in receipt of housing support are recorded as 'In dispersed accommodation' (excluding those in initial accommodation); those who are able to provide themselves with accommodation (generally staying with friends or family) but require other assistance are recorded as 'In receipt of subsistence only'; those who were provided support by LAs under the old system and remained in this housing are recorded as 'Disbenefited'[12]. The data includes dependents in receipt of support but excludes unaccompanied asylum seeking children, who are supported by LAs.

Asylum statistics are a snapshot of where all asylum seekers are living at a given time and do not provide information on when individuals arrived in those areas, or how long they have been living there. Therefore the assumption is made that deprivation data corresponds to the time period when an asylum seeker is living in the reported location.

---

[10] Data on applications and appeals are taken from the Home Office's Case Information Database (CID), and on asylum support are taken from the Asylum Support System (ASYS) database (Home Office 2015, p. 60).
[11] See: https://www.gov.uk/government/collections/immigration-statistics-quarterly-release.
[12] These are very small numbers in the early years of dispersal, decreasing to zero from 2005.

Figure 4.1: Asylum seekers supported in England by type, quarterly since 2003.



Source: Author's creation from Home Office Asylum Statistics (Home Office, 2015b).

4.2.2 Local Authority population data

Mid-year population estimates for 2005 and 2008 are available from the Office for National Statistics and include all persons usually resident in each LAs[13]. These are used to calculate the proportions of asylum seekers living in each LA. The data in this chapter use geographic boundaries that were in place before the changes of the 1st April 2009.

Population density is also reported alongside results from cluster analyses. The 40 LAs with the highest population density in 2005 are identified in Appendix B. In 2008, the same LAs had the highest population density, with the exception of Bexley, which was replaced by the City of London. As expected, these are urban areas, including many London boroughs and LAs with relatively small land areas (see Appendix B).

[13] People arriving into an area from outside the UK are only included in the population estimates if their total stay in the UK is 12 months or more. Visitors and short term migrants (those resident for 3 to 12 months) are not included. Similarly, people who leave the UK are only excluded from the population estimates if they remain outside the UK for 12 months or more. Students are recorded at their term time address (Home Office, 2015).

4.2.3 English Indices of Deprivation 2007 and 2010

Deprivation can be measured in different ways, but in order to analyse whether asylum seekers are likely to be experiencing deprivation in their dispersal location, an area measure is required. This chapter analyses data from the English Indices of Deprivation 2007 and 2010 which, where possible, use data from 2005 and 2008 respectively[14]. The Indices are conceptualised as 'a weighted area level aggregation of … specific dimensions of deprivation' (DCLG, 2011, p. 8) and use 38 indicators to produce a local area score for seven domains of deprivation. This allows various aspects which contribute to deprivation to be taken into account and summarised. The main sources of data are the 2001 Census and government departments. The domains are then combined to provide an overall score and relative rank for each Lower Super Output Area (LSOA)[15]. This is presented as a simple 'area rate', i.e. population experiencing deprivation per population at risk (as for the Income and Employment domains), or by using Maximum Likelihood factor analysis to find appropriate weights for combining indicators into a single score, based on the inter-correlations between the indicators (DCLG, 2008; 2011).

It is good practice to report measures of deprivation at a low level of geography (e.g. LSOA) in order to observe spatial variation. For example, a detailed measure is needed in order to identify a small pocket of severe deprivation; however, in this analysis the priority is to compare deprivation levels with the locations of dispersed asylums seekers and this data is only published at LA level. Furthermore, LAs are the important administrative unit for this analysis as a key aim of the policy was to reduce the 'burden' of support which LAs were responsible for. Therefore a Local Authority 'district' level summary is used: 'Average score' is a population-weighted average of the combined deprivation scores for the LSOAs in a district (DCLG, 2011, p. 57). This takes into account the full range of LSOA scores across a district; it retains the fact that more deprived LSOAs may have more 'extreme' scores, which is not revealed to the same extent if the ranks are used. However, it is not possible to use this measure to compare how deprived two areas are in relation to one another; for example, it is not correct to say that an area with a score of 30 is 'twice as deprived' as an area with a score of 15. It is also important to note the potential feedback in deprivation indices: the number of asylum seekers supported is one of five indicators in the

---

[14] The majority of data refer to 2005 or 2008 respectively, but note that some indicators use data from the 2001 Census.

[15] Super Output Areas (SOA) were designed to provide better reporting of small area statistics. A LSOA covers a population of no less than 1,000 and no more than 3,000 as well as no less than 400 and no more than 1,200 households. http://www.ons.gov.uk/ons/guide-method/geography/beginner-s-guide/census/super-output-areas--soas-/index.html.

Income Deprivation Domain and will therefore contribute to a higher deprivation score; the presence of supported asylum seekers may have more impact on the index in smaller cities.

4.3 Methods

A study of inequality of settlement and deprivation levels across LAs is followed by cluster analysis to identify categories of LAs with similar characteristics. These methods are used to highlight patterns at a national level and summarise the data in a meaningful way in order to better understand how the dispersal policy is implemented, the relationships between dispersal and deprivation, and to inform further analysis.

4.3.1 Inequality analysis of dispersal

Descriptive mapping of numbers and proportions of dispersed and subsistence only asylum seekers, as well as average deprivation score, gives an initial indication of the different patterns which can be observed. This systematic presentation of the spatial distribution of asylum seekers across the country has been previously lacking as essential context for any research on this topic. The national picture of the geographic spread, as well as the number and concentration of asylum seekers residing within LA boundaries, are intrinsic to the aims of this policy and must therefore be the starting point for investigation. In order to map these data, ArcGIS (Geographical Information System) software is used to 'join' boundary data to each dataset to be presented at LA level. The results of cluster analysis are also mapped in order to fully assess the nature of the patterns observed.

Alongside initial descriptive mapping of dispersal, it is interesting to measure the equality of distribution, i.e. how 'fairly' asylum seekers are spread across LAs. Calculating the Gini coefficient gives a measure of the extent to which the distribution of asylum seekers deviates from a perfectly equal distribution (where each LA would be allocated the same number). Bootstrap methods are used to assess uncertainty and to compare Gini coefficients; 95% confidence intervals (CI) are presented alongside each summary measure. The Lorenz Curve illustrates any deviation from the line of equality and, by plotting a cumulative percentage in the order of increasing magnitude, shows where in the distribution the inequality occurs. These are applied to both absolute numbers of asylum seekers and the proportion of asylum seekers relative to LA populations. Scatter plots which show the line of linear regression also help to identify differences in years, numbers and proportions, and allows comparison of the patterns observed.

4.3.2 Cluster analysis with deprivation

The second part of the analysis further investigates spatial patterns of dispersal and deprivation by identifying similar LAs. Cluster analysis is used to classify subjects (or cases) into a number of different groups on the basis of a set of variables, resulting in similar subjects being allocated to the same group. Hierarchical clustering methods use measures of distance to decide which cases are most different (or similar). These methods are appropriate where the number of cases is relatively small and it is important to examine a range of solutions with different numbers of clusters, in order to produce the most meaningful and informative results. For these reasons, hierarchical clustering is used in this analysis, rather than k-means or two step clustering (Norusis, 2011). Hierarchical clustering is either agglomerative or divisive: agglomerative methods begin with each case in an individual cluster and then successively combine similar clusters; divisive methods begin with all cases in one cluster and successively divide into increasing numbers of clusters (Norusis, 2011). Three hierarchical methods are tested in this analysis:

1. Average linkage between groups
2. Average linkage within groups
3. Ward’s method

These methods take into account the distances of *all* cases (either within/between clusters, or to the variable mean), rather than just the nearest or furthest cases or variable means (as in the nearest neighbour, furthest neighbour and centroid methods).

A measure of distance must also be selected for hierarchical clustering: the Euclidean measure of distance is the square root of the sum of the squared differences between values over all variables; Squared Euclidean distance (the sum of the squared distances) gives greater emphasis to greater distances, therefore the resulting clusters may be different, depending on the presence of outliers within the dataset. Both of these measures are tested for the Average Linkage methods, which compare each case between or within groups, but Ward’s methods only uses the Squared Euclidean distance measure, and takes into account how much the sum of squares will increase with each potential merge of groups. The agglomeration schedule, cluster sizes, means and measures of spread (e.g. range and standard deviation) are analysed in order to present results which have an interpretable number of relatively homogenous clusters.

In this analysis, Local Authorities are the cases to be clustered by their characteristics: average deprivation score (‘population weighted average of the combined scores for the SOAs in a district’) and number of dispersed asylum seekers per 1,000 LA population. Rates of asylum seekers are used in order to account for different LA sizes. The data are standardised using z-scores to give consistent scales (each variable will have a mean of zero and a variance of one); this is important in cluster analysis which compares ‘distance’ to group similar cases.

Results are presented with summary descriptions for comparison between clusters and solutions. Thresholds for these are included in Appendix E; these are arbitrary and for the purpose of highlighting relative differences in cluster means. Individual cases may fall outside of these thresholds, as illustrated by the ranges included in tables 7.8-7.14 in Appendix E.

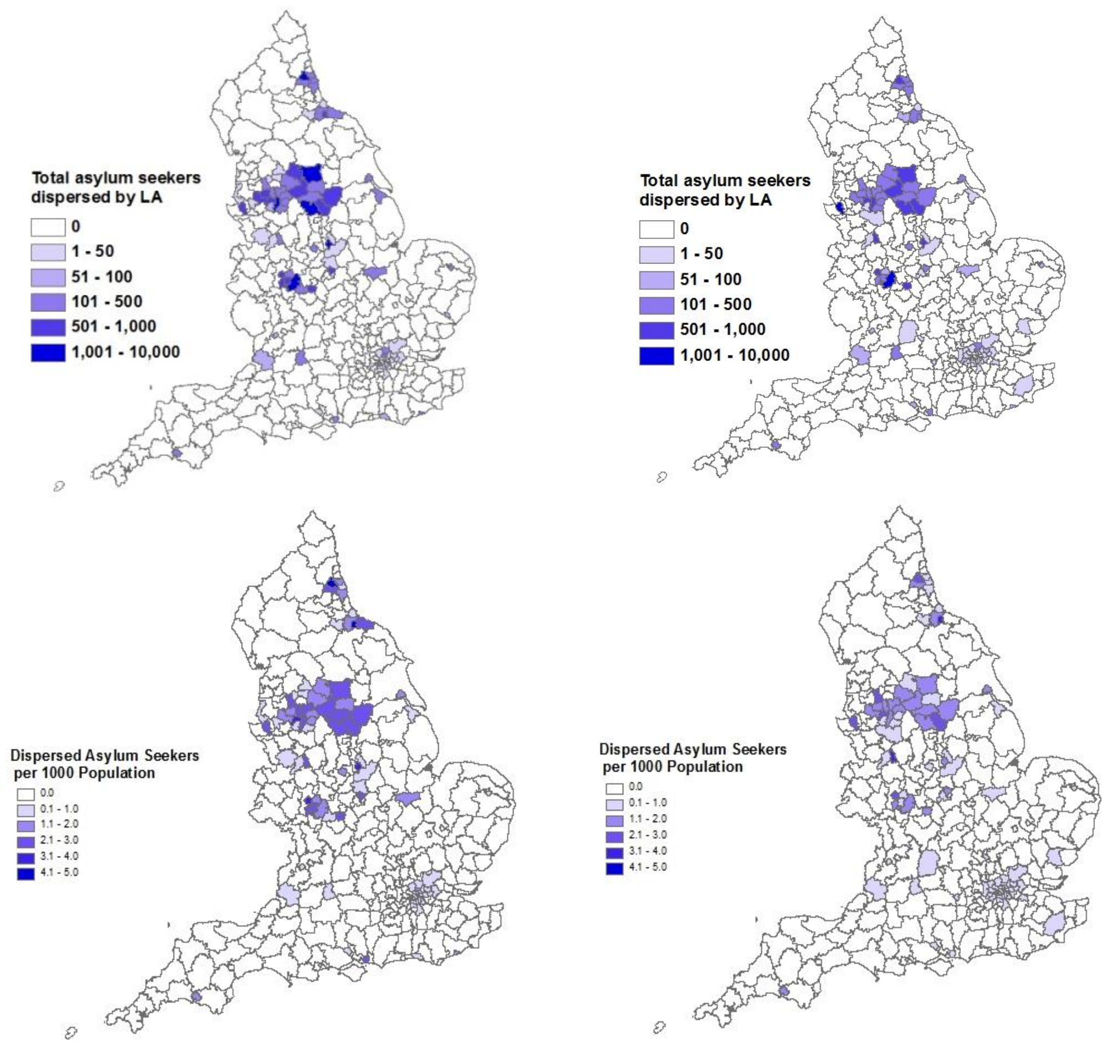
4.4 Results

4.4.1 Inequality analysis of dispersal

Maps of LAs in England for 2005 and 2008 (Figure 4.2) show the accommodation locations of asylum seekers; presenting data in this way enables the identification of patterns and comparison of different support groups and initial indications of correlation with deprivation.

Figure 4.2: Total and proportion (below) of asylum seekers dispersed by Local Authority in England, 2005 Q2 and 2008 Q2.

2005 Q2 **2008 Q2**



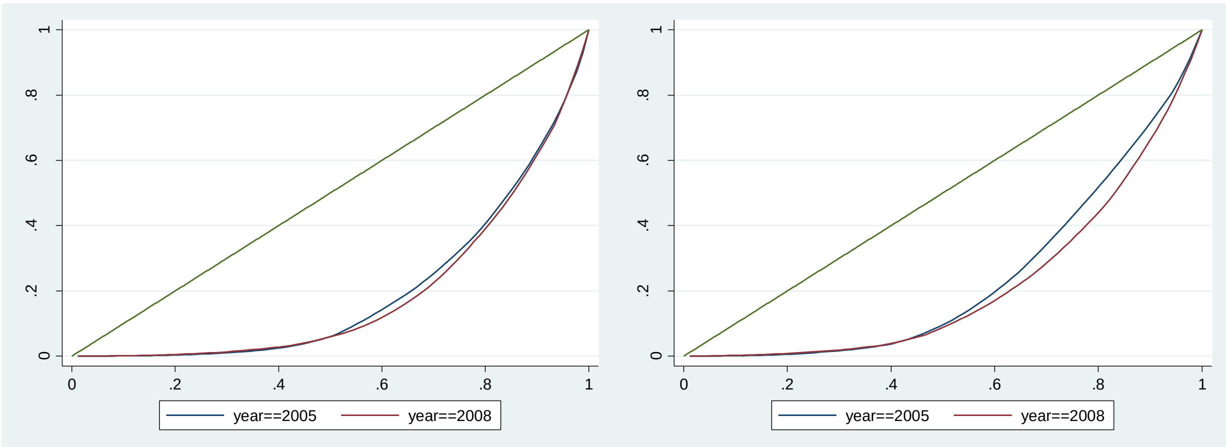
Source: Author's creation from Home Office Asylum Statistics (Home Office, 2011).

The total number of dispersed asylum seekers in each LA (with at least one asylum seeker) ranges from 2 to 1,976 in the middle of 2005 and 1 to 1,163 in the middle of 2008. In 2005, Leeds was the LA with the highest number of dispersed asylum seekers (1,976), one of six LAs with more than 1,000 residing within its boundaries (see Appendix D). This had decreased to just two LAs in 2008: Birmingham and Liverpool. While the overall number of dispersed asylum seekers decreased from 28,281 in 2005 to 19,542 in 2008, and the number in many LAs also decreased, the number located in some LAs increased considerably, for example Stoke-on-Trent rose from 494 in 2005 Q2 to 852 in 2008 Q2 (see Appendix D).

There is considerable variation in how many asylum seekers are housed across LAs (see Figure 4.2); however, it is also important to assess the distribution of asylum seekers relative to LA population size. This was considered an important element in the creation of 'cluster areas' for settlement, with more highly populated LAs permitted to receive more asylum seekers (NAO, 2014). The absolute numbers of asylum seekers is of interest as it represents the size of the population that each LA is hosting, and therefore the size of housing stock and financial cost; however, the considerable variation in the size and makeup of LAs (demographically and geographically) means that a more pertinent measure for the purpose of comparing inequality is the concentration of asylum seekers relative to the size of each LA. The proportion dispersed in each LA (with at least one) ranges from 0.008 to 4.45 per 1,000 population in 2005 and 0.006 to 3.56 per 1,000 in 2008. The LAs with the highest concentration were Newcastle-upon-Tyne in 2005 and Stoke-on-Trent in 2008 (see Appendix D).

Figure 4.3 shows the Lorenz curves for the cumulative percentage of asylum seekers in LAs (only including LAs with at least one AS) in 2005 and 2008 to more formally identify differences in the distribution of dispersed asylum seekers.

Figure 4.3: Lorenz curves for the cumulative percentage of asylum seekers (totals, left, per 1,000 population, right) dispersed in Local Authorities in 2005 Q2 and 2008 Q2 and line of equality.



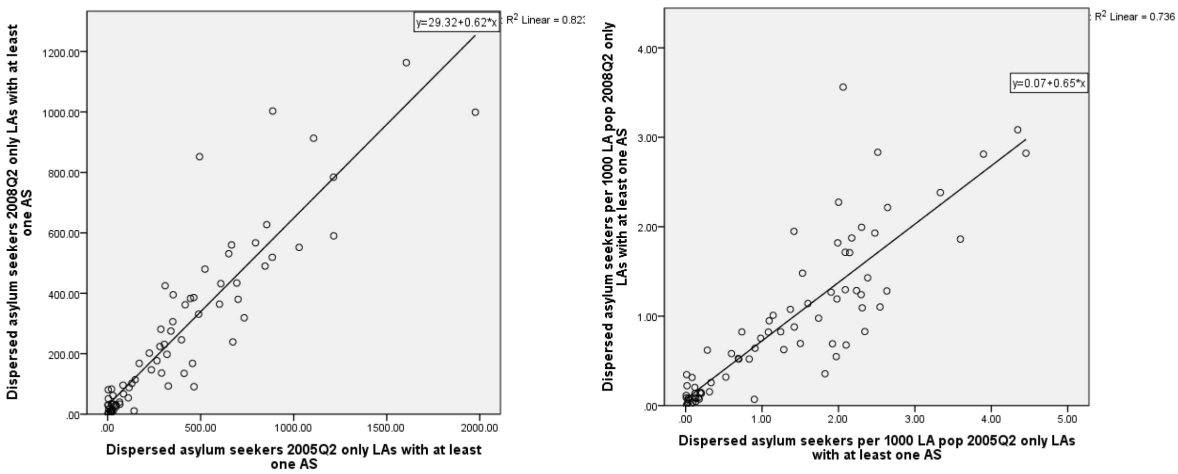
Source: Author's creation from Home Office Asylum Statistics (Home Office, 2011).

The line for 2005 is closer to the line of equality, and has a Gini coefficient of 0.604 (CI: 0.551-0.685) which is lower than 0.618 (CI: 0.569-0.681) in 2008 for absolute numbers of asylum seekers dispersed. This suggests that they were spread less equally across LAs in 2008, in spite of lower overall numbers dispersed across a greater number of LAs. That said, it is important to note the level of uncertainty, with the overlap in confidence intervals suggesting that this difference is not statistically significant.

Plotting the cumulative percentage of the *proportion* of dispersed asylum seekers relative to the LA population shows that the distribution is less unequal for both years (i.e. closer to the line of equality), compared to absolute numbers; this is also reflected in the Gini coefficients of 0.521 (CI: 0.456-0.628) for 2005 and 0.565 (CI: 0.510-0.623) for 2008. However, the difference in how equally asylum seekers were spread between the two time points is greater when LA population is taken into account.

Figure 4.4 shows the difference in number and proportions of dispersed for each LA in 2005 and 2008. While LAs with fewer dispersed asylum seekers tend to experience a smaller change in numbers, those with the highest numbers also exhibit the greatest change between the two years.

Figure 4.4: Scatter graph of Local Authorities by total number (left) and number per 1,000 population (right) dispersed in 2005 and 2008, with linear regression line.

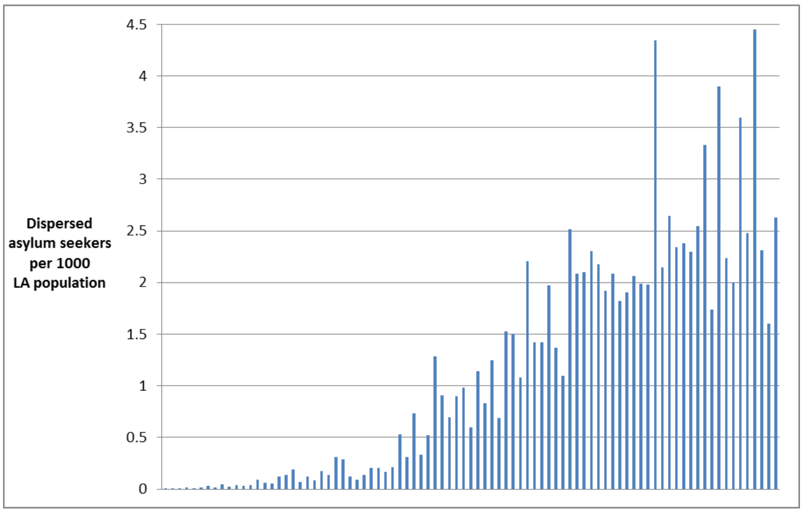


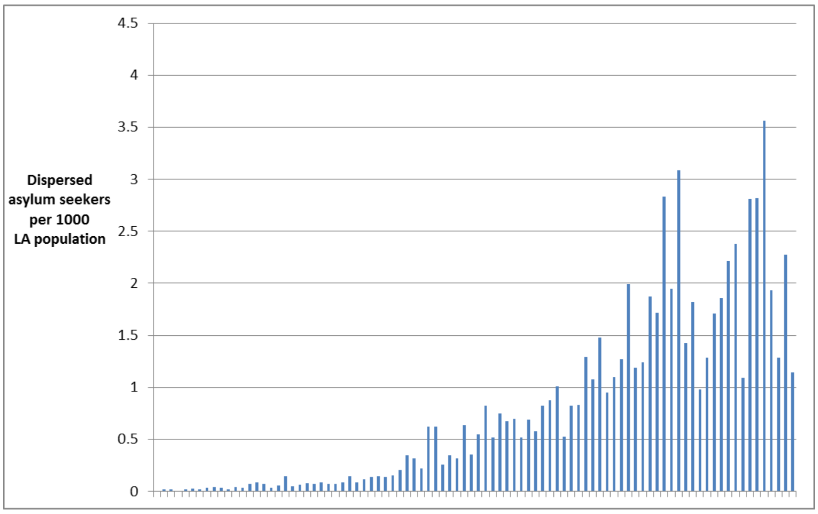
Source: Author's creation from Home Office Asylum Statistics (Home Office, 2011).

Some LAs had a relatively high *number* of dispersed asylum seekers, but when this is compared as a *proportion* of the total LA population, other areas are considerably higher due to the differences in overall size of LA populations[16]. Figure 4.5 shows the proportion of dispersed asylum seekers in order of lowest to highest absolute number dispersed for 2005 and 2008.

[16] See Appendix D for LA ranks of highest number and proportion dispersed.

Figure 4.5: Dispersed asylum seekers per 1,000 LA population, in order of lowest to highest total number dispersed (by LA), 2005 Q2 (top) and 2008 Q2 (below) (only showing LAs with at least one dispersed asylum seeker).

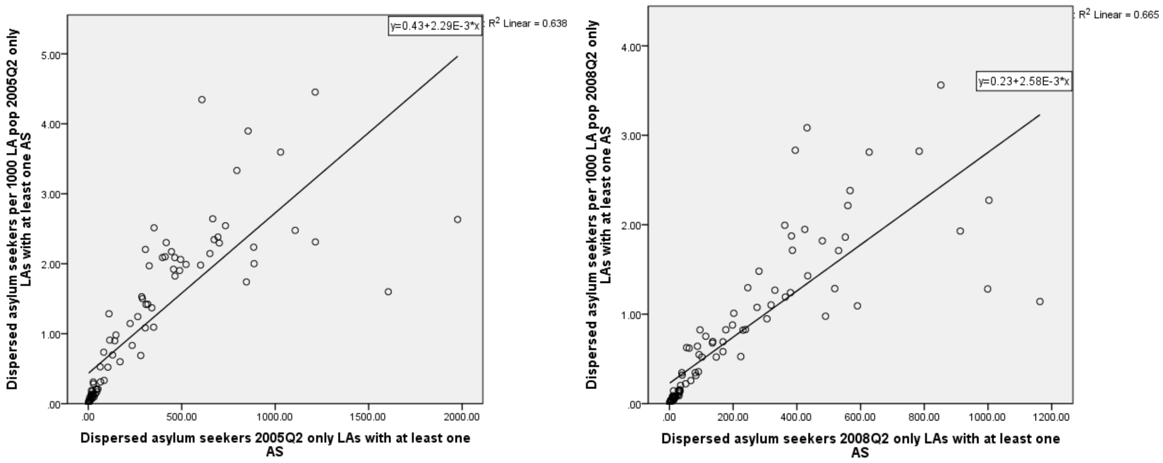




Source: Author's creation from Home Office Asylum Statistics (Home Office, 2011).

Again, Figure 4.6 shows that greater variation is observed for higher numbers and proportions of dispersed asylum seekers. Analysis of Spearmen's rank correlation showed a strong positive relationship (0.956) between absolute numbers and proportions dispersed in 2005, which was significant at the one percent level. A similar pattern is observed in 2008 but with lower maximum proportions and numbers dispersed. Spearmen's rank correlation also showed a strong positive relationship (0.965) between absolute numbers and proportions dispersed in 2008 which was significant at the one percent level.

Figure 4.6: Scatter graph of Local Authorities by total dispersed and number dispersed per 1,000 LA population, with linear regression line in 2005 (left) and 2008 (right).

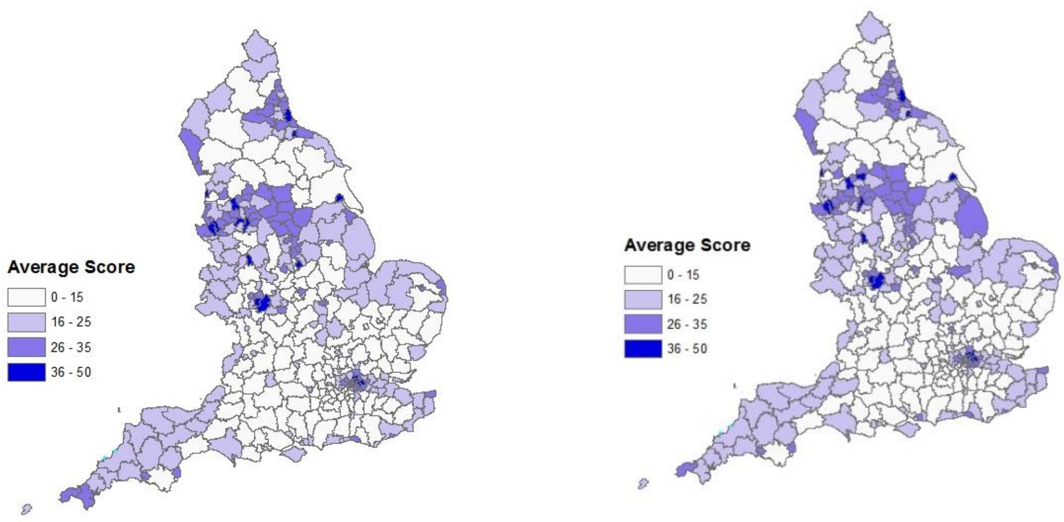


Source: Author's creation from Home Office Asylum Statistics (Home Office, 2011).

One of the main purposes of dispersal was to shift the focus of refugee settlement away from London and the South East of England; Figure 4.2 shows that in 2005 and 2008 dispersed asylum seekers were mainly being housed in LAs across the North West as well as the North East and Midlands. Maps of deprivation indices at LA level (Figure 4.7) show that these areas are also identified as having relatively high deprivation. A notable exception where high deprivation does not correlate with high levels of dispersal is London, where 812 asylum seekers were dispersed in 2005 compared to 12,456 receiving subsistence only support (see Figure 4.8); in 2008 the figures were 871 and 4,392 respectively.

Figure 4.7: Indices of Multiple Deprivation, Average score by Local Authority.

IMD 2007 (based on 2005 data) **IMD 2010 (based on 2008 data)**

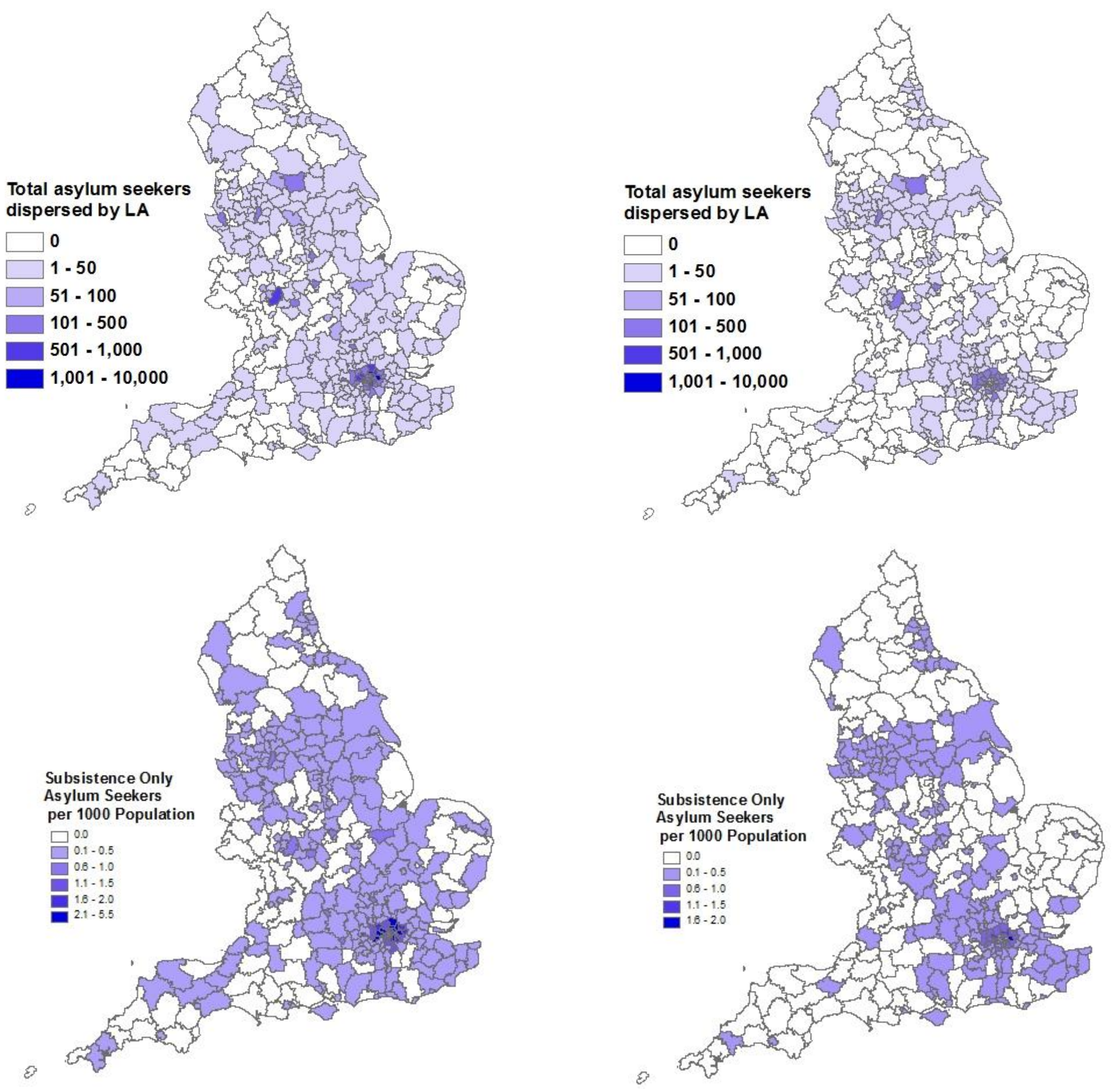


Source: Author's creation from DCLG data (DCLG, 2008; 2011).

Overall, 17,382 asylum seekers were on subsistence only support in 2005 and 6,345 in 2008. Although Figure 4.8 shows that those asylum seekers receiving SO support are more concentrated in London, it is also clear that this population are living in a much greater number of LAs, with 254 LAs having at least one SO asylum seeker in 2005, compared to 87 LAs with at least one dispersed. This was also the case in 2008: 90 LAs had at least one dispersed asylum seeker and 199 LAs had at least one receiving SO support.

Figure 4.8: Total and proportion (below) of asylum seekers on subsistence only support by Local Authority in England, 2005 Q2 and 2008 Q2.

2005 Q2 **2008 Q2**

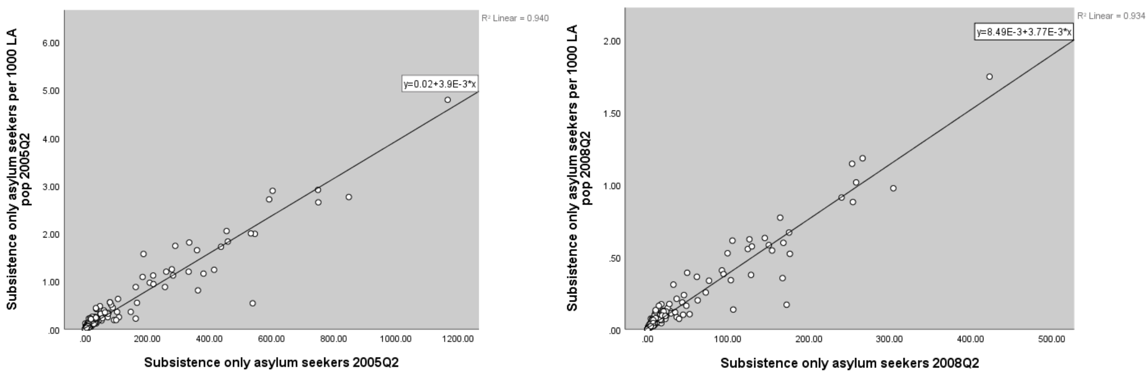


Source: Author's creation from Home Office Asylum Statistics (Home Office, 2011).

It has been suggested that this is evidence that asylum seekers actually disperse themselves better than the government policy (see for example Hynes, 2011), however, the focus of settlement in London (see Appendix D) remains a drawback of allowing individual choice:

Newham has the highest number on subsistence only support in 2005 (1,166) and 2008 (423). The capital has historically been the region of choice for new migrant communities and many of these current asylum seekers will be living with friends and family who have previously settled here. This greater inequality is reflected in higher Gini coefficients for the subsistence only population compared to those dispersed. In 2005 the Gini coefficient for the absolute number of those on subsistence only support was 0.815 (CI: 0.788-0.834) and 0.760 (CI: 0.734-0.788) for the proportion of subsistence asylum seekers in relation to the LA population; in 2008 these were 0.760 (CI: 0.738-0.801) and 0.702 (CI: 0.667-0.733) respectively. The CIs do not overlap with the dispersed figures for either year, suggesting a statistically significant difference.

Figure 4.9: Scatter graph of Local Authorities by total SO and number SO per 1,000 LA population, with linear regression line in 2005 (left) and 2008 (right).



Source: Author's creation from Home Office Asylum Statistics (Home Office, 2011).

Figure 4.9 shows that there is much less variation between absolute numbers and proportions of SO asylum seekers, compared to those dispersed, for both 2005 and 2008. Furthermore, a greater number of LAs are grouped at the lower end, reflecting patterns shown in Figure 4.8.

4.4.2 Cluster analysis with deprivation

Following from the above analyses which highlighted variation in the distribution of asylum seekers at a national level and how this relates to patterns of deprivation, cluster analysis can be used to identify LAs sharing similar characteristics and further understand the nature of this relationship.

Of the three methods tested for 2005 data, Average Linkage (between groups) produces a five-cluster solution which groups LAs in a meaningful way for clear interpretation (see 'Solution A', Table 4.1). There is no difference between Euclidean and Squared Euclidean measures of distance, and the solution is very similar to the results from the Average Linkage, within groups method. Mean population density and mean geographic size are also reported for each cluster in Appendix

E in order to further inform the discussion and interpretation of the results produced. The means reported include the LAs that have at least one asylum seeker living in dispersal accommodation (87 in 2005 and 90 in 2008) included in the cluster analyses, rather than the mean of all authorities across England.

Table 4.1: Summary of Local Authority clusters (solution A), 2005. [17]

Cluster	Number of LAs	Deprivation	Rate of dispersal	Description
1	25	Low	Low	Includes suburban towns and affluent London
2	32	Med	High	Larger geographic areas including large cities
3	15	High	Low	Mostly deprived London
4	5	High	Very high	Smaller geographic urban areas
5	10	Very high	High	Smaller geographic urban areas

Source: Author's analysis using Home Office Asylum Statistics (Home Office, 2011).

Cluster 3 can be identified as mostly London boroughs with high deprivation but few dispersed asylum seekers, as the policy intended. Clusters 4 and 5 have high deprivation and high proportions dispersed, and include LAs with large cities and relatively small geographic area. Cluster 2 is the largest cluster and many of these LAs have large geographic areas and may be less homogenous with greater variation at lower levels[18].

Ward's method applied to 2005 data produced 'Solution B' (see Table 4.2); this has the same London cluster (4) as solution A (3) and a 'high, high' cluster (5) which combines the 'high, high' clusters (4 and 5) in solution A. The repeated identification of these clusters as a result of different methods tested suggests that the findings are robust.

[17] See Appendix E for full list of LAs by cluster.
[18] See Appendix E for cluster characteristics.

Table 4.2: Summary of Local Authority clusters (solution B), 2005.[19]

Cluster	Number of LAs	Deprivation	Rate of dispersal	Description
1	17	Very low	Low	Includes suburban towns and affluent London
2	22	Med	Med	Includes large towns and more southerly cities
3	18	Med	High	Large geographic areas including large cities
4	15	High	Low	Mostly deprived London
5	15	Very high	High	Smaller geographic urban areas

Source: Author's analysis using Home Office Asylum Statistics (Home Office, 2011).

In analysis of 2008 data using the three methods described above, solutions with six clusters consistently included a 'high, low', 'mostly deprived London' cluster (18 LAs), a 'high, very high' cluster (six LAs) and a small 'very high, high' cluster (Liverpool and Manchester). There was also always a 'low, low' cluster but this varied in the number of LAs. Often, solutions with fewer clusters lost interesting detail (e.g. London cluster combined with another cluster); this was less the case though with the 'Average within group linkage' solution (using the Squared Euclidean distance measure) with four clusters, which retained a distinct London cluster and 'high, high' cluster (combining clusters 4 and 6 from solution C). The six cluster solution using 'average between groups linkage' is presented in Table 4.3 (Euclidean and Squared Euclidean distance measures produced the same results).

[19] See Appendix E for full list of LAs by cluster.

Table 4.3: Summary of Local Authority clusters (solution C), 2008.[20]

Cluster	Number of LAs	Deprivation	Rate of dispersal	Description
1	27	Very low	Low	Includes suburban towns and affluent London
2	28	Med	Med	Larger geographic areas including large cities
3	9	Med	High	Large geographic areas including large cities
4	6	High	Very high	Smaller geographic urban areas
5	18	High	Low	Mostly deprived London
6	2	Very high	High	Liverpool and Manchester

Source: Author's analysis using Home Office Asylum Statistics (Home Office, 2011).

The four cluster solution using average within groups linkage (Squared Euclidean distance) shown in Table 4.4 is a clear summary of the patterns observed, grouping LAs by their characteristics in a meaningful way. Results of analysis with a greater number of clusters give detail to inform understanding of these clusters, but solution D represents robust and stable findings that can be easily interpreted.

Table 4.4: Summary of Local Authority clusters (solution D), 2008.[21]

Cluster	Number of LAs	Deprivation	Rate of dispersal	Description
1	40	Low	Low	Large cluster. Includes suburban towns and affluent London
2	24	Med	Med	Larger geographic areas including large cities
3	18	High	Low	Mostly deprived London
4	8	High	High	Smaller geographic urban areas

Source: Author's analysis using Home Office Asylum Statistics (Home Office, 2011).

[20] See Appendix E for full list of LAs by cluster.
[21] See Appendix E for full list of LAs by cluster.

Figure 4.10 shows the geographic spread of cluster analysis results and the allocation of LAs in solutions A to D.

Figure 4.10: Solutions A and B for 2005 (top), C and D for 2008 (below).

Source: Author's analysis using Home Office Asylum Statistics (Home Office, 2011).

It is important to note that there is no generally accepted 'best' method of cluster analysis and there is a largely subjective element in the assessment and interpretation of the results (Manly and Navarro Alberto, 2017). The aim is to gain some meaningful insights into the similarities and differences that can be observed between cases, based on the selected relevant variables. Therefore the results presented in solution D have been selected as a useful summary of the most common cluster patterns which were observed from the methods tested; solutions with a greater number of (often smaller) clusters were subject to greater variation for the dispersed population. The cluster from solution D can be conceptualised as shown in Table 4.5 where cells highlighted in blue show clusters that were observed.

Table 4.5: LA clusters frequently observed for proportions dispersed and deprivation score.

		Asylum seekers (dispersed)		
		Low	Medium	High
Deprivation	Low			
	Medium			
	High			

Source: Author’s creation based on clusters using Home Office Asylum Statistics (Home Office, 2011).

This highlights the clusters that were produced in this analysis, but also shows which combinations of the variables were not observed. Table 4.5 reflects the most frequent results for dispersed asylum seekers, showing an emphasis on the diagonal cells in addition to high deprivation and low asylum seekers. While assessment of how the dispersal policy is implemented is the key focus of this study, some comparison with the subsistence only population can be useful to highlight differences in observed patterns. Cluster analysis carried out on LAs by the proportions of asylum seekers supported for subsistence only and deprivation scores shows that those who were not dispersed are consistently divided into six distinct clusters with either very high or low proportions of asylum seekers. Table 4.6 presents the clusters produced for this much larger group of LAs as a summary of the most common cluster pattern which was observed from the methods tested.

Table 4.6: LA clusters frequently observed for proportions of SO and deprivation score.

		Asylum seekers (SO)		
		Low	Medium	High
Deprivation	Low			
	Medium			
	High			

Source: Author's creation based on clusters using Home Office Asylum Statistics (Home Office, 2011).

A clear pattern is revealed which splits LAs into around 26 with high proportions of asylum seekers (consisting almost exclusively of London boroughs) and the remaining large number of LAs outside London with low proportions; and then each by low, medium and high levels of deprivation.

4.5 Discussion and conclusions

The context for this analysis is the stated aims of the 1999 Immigration and Asylum Act: to change the geographic focus of settlement and to restrict the flow of asylum seekers through deterrence. The results presented in this chapter provide evidence of the relationship between asylum support and deprivation across LAs in England. The broad findings confirm what has been suggested by practitioners and in anecdotal evidence as well as local level and qualitative research studies (e.g. Hynes, 2011; Phillimore and Goodson, 2006), but has previously been lacking the systematic quantitative analysis at a national level presented here. The separate reporting of populations by support status (dispersed or subsistence only) also gives new insight into the dispersal policy and how the nature of support provided, including limitations on location of settlement, impacts on the resulting geographic patterns.

4.5.1 Inequality analysis of dispersal

The results of the inequality analysis show that there is variation in the concentration of asylum seekers across LAs in England. Asylum seekers are not spread equally under the dispersal scheme, either by absolute numbers or as a proportion of the LA population, and those on subsistence only support show different patterns of settlement. Much of the variation observed in the number of asylum seekers dispersed to different LAs is explained by the nature of the contracts with housing providers at a local level. For example NASS agreed a contract with the West Midlands Consortium for Asylum seekers and Refugees (WMCAR)[22] in 2000 to house dispersed asylum seekers within the region. This resulted in high numbers in certain LAs within the region, such as Birmingham and Wolverhampton.

In theory, decisions about where to house asylum seekers are made based on a range of factors (NAO, 2014), such as the creation of 'language clusters', the availability of services and housing, local population density (in order to restrict the levels dispersed to one in 200) and the requirement to be outside London and its surrounding areas. In practice, the results show that the main aim of moving the focus of settlement away from the capital has been achieved; local studies and qualitative research suggest that beyond this, choice of locations has been largely based on availability of affordable accommodation (see for example Zetter et al, 2002). The maps presented in this chapter show that, with the exception of London, similar patterns are observed in the location of dispersed asylum seekers and areas of relatively high deprivation, when measured at the LA level. This is supported by the HC Home Affairs Committee (2017, para.32) which states that a key aim of COMPASS was to save money, meaning that locations with cheap housing have been 'targeted', and the resulting patterns of dispersal and deprivation observed at a national level is unsurprising.

According to the COMPASS contracts, the providers of dispersal housing must consider a range of social, housing and community cohesion factors (HC Home Affairs Committee, 2017, para. 34) when proposing properties. These factors include: 'the availability and concentration of accommodation; the capacity of local health, education and other support services; and the level of risk of increased social tension if the number of asylum seekers increases within a given area' (NAO, 2013). These are monitored by LAs, who have the right to withdraw existing consent for specific properties to be used for asylum seeker accommodation or reject new proposals if there are any specific concerns. The disconnect between decision making at a policy level and how

[22] The Consortium includes Birmingham, Coventry, Dudley, Sandwell, Solihull, Stoke-on-Trent, Walsall and Wolverhampton. Birmingham City Council leads the consortium and manages the contract with NASS.

properties are acquired by providers with the involvement of LAs has resulted in a system which has failed to produce an 'equitable distribution' across the country, but has instead resulted in settlement focussed on some of the most deprived areas (HC Home Affairs Committee, 2017, para. 43).

These findings raise further questions about whether asylum seekers can be supported effectively in dispersal areas if these tend to be highly deprived. Previous research has shown that there is regional inequality in education attainment (Commission on Inequality in Education, 2016) and healthcare (Right Care, 2015) across the country, and therefore where a person lives can impact on their wellbeing. The following chapter will investigate further whether these influence the outcomes of dispersed asylum seekers once they have been granted permission to stay.

The pattern of those asylum seekers on SO support is much more widespread, with higher concentrations in London. This shows that when asylum seekers have the 'choice' of where to live, many will consistently locate in and around London. It is clear that one of the main policy aims is being met by shifting the focus of settlement away from the capital. That said, the fact that there are a greater number of LAs accommodating a few asylum seekers suggests that, as Hynes (2001, p. 86) proposes, to some extent they 'disperse themselves more effectively than social policy interventions or efforts at institutional redistribution'.

A balance must be achieved in decision-making around this policy between 'spreading the burden' – taking pressure off local services and housing, particularly in London – and 'clustering', to allow efficient and effective provision of services and support networks of other asylum seekers e.g. 'clustering' of same languages. With regard to the first aim to 'spread the burden', it should be remembered that LAs are housing the asylum seekers, often through contracts with private providers, but are not actually responsible for them *financially*; through the creation of NASS funding was centralised in the 1999 Act. It is therefore the case that reference to a 'burden' to be spread must be about other pressures, such as service provision (e.g. education and healthcare), community cohesion and housing availability. The question of 'clustering' is more complex; it is clear from previous research that the implementation of clustering has often been inconsistent or lacking (NAO, 2014). Furthermore, the extent to which language clusters may even be beneficial is unclear: sharing the same language may be less important to individuals than for example a common religion, ethnicity or political beliefs, particularly when many asylum seekers are fleeing conflict and persecution where divisions along these lines are widespread. It is arguably more important to ensure that asylum seekers are living in locations where they are able to access essential services. The relationship between support status, social connections, access to services and resulting outcomes are addressed further through longitudinal analysis of survey data in the next chapter.

4.5.2 Cluster analysis with deprivation

It is clear that asylum seekers are not spread equally across England; cluster analysis can give further insight into the patterns observed, identifying groups of LAs with shared characteristics of dispersal and deprivation.

The results consistently produced a cluster which was made up largely of London boroughs with high average deprivation and low levels of dispersal. This is unsurprising given that one of the key aims of dispersal was to move the focus of settlement away from London. The average population density of the 'London' clusters is also considerably higher than all other clusters, and the geographic size is relatively small, as expected. That said, London is still supporting the majority of SO asylum seekers: 12,456 (72%) in 2005 Q2 and 4,392 (69%) in 2008 Q2. Furthermore, it has been suggested that dispersed asylum seekers may move from their dispersal areas, for example to London, once they have received a positive decision. Data on the refugee population is even more scarce, but Hynes (2011, p. 105) found 'numerous examples' of individual asylum seekers relocating and Stewart (2011) also provides evidence of some onward migration of refugees in Scotland. Studies of secondary migration in Europe support these findings, but in England, individuals may be more restricted; for example by the requirement for a 'local connection' when applying to move into mainstream LA accommodation would limit the ability to move to other areas if continued housing support was needed.

A number of LAs are repeatedly grouped in a 'high, high' clusters for both 2005 and 2008; these include Blackburn with Darwen, Liverpool, Manchester, Middlesbrough, Newcastle-upon-Tyne, Salford, Stoke-on-Trent and Wolverhampton. They tend to have small geographic areas and high population density and while some of these have a history of receiving migrants, resulting in established ethnic minority communities, a notable exception is Stoke-on-Trent. Dispersal to Stoke-on-Trent since 2000 as part of the West Midlands regional consortium has been controversial due to a lack of preparation for the receiving area and lack of existing support in place, as asylum seekers brought into this situation can be particularly conspicuous due to the fact that they are often housed together in one place.

It is also interesting to note that one London authority, Haringey, was included in the 2005 'high, high clusters; this was one of the few LAs in the capital to be accommodating a considerable population of dispersed asylum seekers (319). This number had decreased by around a third by 2008 Q2 and Haringey was then clustered with other highly deprived London authorities.

Liverpool and Manchester, which were allocated to their own small cluster for many of the 2008 solutions tested, had the seventh and eleventh highest proportion of dispersed asylum seekers per LA population of all LAs in 2008 Q2. They were also ranked first and fourth respectively in the 2010 Deprivation Index.

There is also a clear 'low, low' group which tends to include what can be described as 'suburban towns' as well as more affluent London boroughs. These suburban towns (e.g. Solihull, Trafford, Stockport, Rushcliffe) are often made up of desirable residential areas near to major cities. The LAs in London also had low levels of dispersed asylum seekers but lower deprivation scores than the other London areas in the 'high, low' clusters. Many LAs in this cluster house less than 30 dispersed asylum seekers, which would only be a handful of families. This may be because there are particular reasons for these individuals to be located in these areas, for example to be near to relatives already here, or near to essential health services, rather than being a result of large scale housing contracts.

Comparing patterns of dispersal and deprivation with findings for those on subsistence only support confirms that although the 'low, low' group for dispersed includes a number of affluent London boroughs, these LAs fall into the 'low, high' category for SO, with many asylum seekers settling here; this cluster is not observed for dispersed asylum seekers and is a key example of locations where the dispersal policy is intended to reduce settlement. The clusters of those receiving subsistence only support show a clear split between those living in London and those in other areas of the country. This pattern is much more pronounced than any relationship with deprivation, with clusters observed for low, medium and high deprivation with both high and low levels of asylum seekers. However, that is not to say that SO asylum seekers are not also living in deprived areas. Further investigation below LA level is necessary to determine this, but it is well known that migrants and ethnic minorities do tend to be more deprived than the general population, and those on SO support are likely to be living with friends and family in this category who are already resident.

The 'med, med' and 'med, high' clusters of dispersed asylum seekers do not show a clear correlation with deprivation, but looking at the LAs in these clusters, it is possible to identify some possible reasons for this. Some have larger geographic areas, often covering large cities; there will be variation below LA level, and where areas of high dispersal/deprivation exist, they are not identifiable in this data. That said the House of Commons (HC) Home Affairs Committee (2017, para. 38) found evidence that asylum accommodation is 'often not evenly dispersed within a LA but clustered in a few wards'. There are also LAs which include large towns or more southerly cities where there will be pockets of deprivation but often also a lack of affordable housing, similar to deprived London boroughs. The policy intention is to house asylum seekers outside London, but the resulting pattern of dispersal concentrated in the midlands, north west and north east, may be close to that of deprivation due to available vacant or low cost housing (Anie et al, 2005; HC Home Affairs Committee, 2017). Deprivation observed in the south west of the country is not associated with dispersal to these areas.

The analyses presented here have consistently identified groups of LAs with low, medium and high levels of dispersal that also have low, medium and high levels of deprivation. An additional cluster was observed with high levels of deprivation and low levels of dispersal, mostly made up of London boroughs, in line with the dispersal policy. A more complete understanding of this relationship is not possible without further disaggregation, below LA, because deprivation levels in particular vary at a lower geography (DCLG, 2011). Nevertheless, these findings provide additional support to the existing local level and qualitative research which also point towards the association between dispersal and high levels of deprivation for some areas.

Chapter 5 Dispersed asylum seekers, refugee outcomes

5.1 Introduction

It is clear from a review of the existing literature and research that specific data on refugee and asylum seekers in the UK, particularly regarding refugee characteristics, has been extremely limited. The publication of the Survey of New Refugees in 2010 by the Home Office represented a considerable attempt to rectify this, primarily in order to provide information on refugee integration and enable policy evaluation. However, to date, this data has been notably under-used by researchers to investigate specific pressing questions about the refugee experience, beyond a broad exploration of integration[23].

The aim of this chapter is to identify social and economic outcomes of refugees who were dispersed as asylum seekers and how these compare to those who were not dispersed. In order to gain a thorough understanding of this, it is important to establish which individual background and contextual characteristics are associated with being dispersed, and to then control for these when assessing differences in outcomes. This chapter also aims to understand the patterns of change over time which occur during the first 21 months of refugees' leave to remain in the UK, and how these vary depending on asylum experience.

This longitudinal analysis of individual data at a national level addresses gaps identified in the existing literature, and the focus on refugees who were dispersed as asylum seekers allows experience in these two states to linked, and differences to be highlighted.

Three main research questions are considered:

1. What are the background characteristics of dispersed and non-dispersed refugees? Which characteristics are associated with being dispersed?
2. Are baseline respondents and those remaining at wave three the same and can we assume that attrition is random?
3. What is the impact of dispersal on the social and economic outcomes of refugees, after controlling for background characteristics and any attrition bias?

The first research question is addressed using descriptive statistics and regression modelling of the baseline data to explore how a range of characteristics are associated with dispersal status. Descriptive statistics are essential for policy analysis as they show the real nature of the refugee

[23] See Cheung and Phillimore (2013) for a study of social networks and social capital.

population living in the UK. This element facilitates better understanding of the needs of refugees and also allows rigorous analysis of policy implementation and practice, as well as appropriate and effective policy recommendations. Cross sectional analysis of the baseline survey highlights which characteristics of refugees are associated with being dispersed. The relationships between individual background and contextual characteristics, such as age, education and country of origin, and the likelihood of being dispersed, are investigated using multinomial logistic regression models.

As is commonly the case with longitudinal datasets, attrition and non-response are prominent features of this survey. A thorough assessment of the differences between the baseline population and those remaining at wave three addresses the second research question. The ability of weights to mitigate any bias introduced through attrition is also investigated.

Analysis for the third research question includes modelling the data from all four waves of the survey, taking into account its longitudinal structure. This highlights which of the differences in social and economic outcomes between dispersed and non-dispersed refugees, identified in the initial descriptive analysis, are still observed when background characteristics reported in the cross sectional analysis are controlled for.

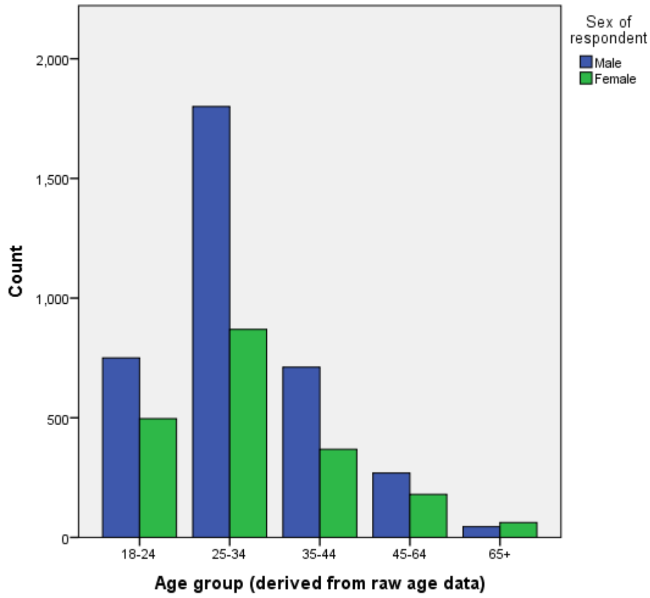
5.2 Data: Survey of New Refugees

The Survey of New Refugees was commissioned by the Home Office, primarily to provide data on the integration of refugees in Britain, and to assess the new Strategic Upgrade of National Refugee Integration Services (SUNRISE) pilot scheme which was intended to support refugees in the integration process; however, not enough participants responded to allow meaningful analysis of its impact. It is the only longitudinal dataset on refugees in England and, crucially, includes information on whether respondents were living in dispersal accommodation.

5.2.1 Study population

The Survey of New Refugees was sent to all those over the age of 18 in Britain who were granted asylum, humanitarian protection or discretionary leave to remain between 1st December 2005 and 25th March 2007. Baseline questionnaires were distributed weekly to those who had received their asylum decision since the previous week. There were 9,127 in the eligible population and 7,765 were sent the initial Baseline survey, with 5,678 of these responding within the 12 weeks necessary for inclusion in the dataset. This was followed by three subsequent questionnaires eight, 15 and 21 months later. Figure 5.1 shows that there were more males than females in every age group, except 65+, in the unweighted data.

Figure 5.1: Grouped age[24] and sex distribution of baseline respondents, unweighted.



Source: Author's creation from SNR data (Home Office, 2010e).

The majority of new refugees are in the working ages, with males aged 25-34 clearly the largest group.

5.2.2 Survey design and data collection

The initial baseline survey recorded background characteristics as well as data on current situation (e.g. housing) and outcome variables to be followed up in later waves. This includes education and employment before coming to the UK, English skills, contact with groups/organisations (including help needed/accessed), contact with friends/relatives and health.

There was also longitudinal recording of additional 'outcome' variables eight, 15 and 21 months after the baseline. The data includes some repeated measures, including for example on: accommodation and housing, help and support received, education and training, employment and benefits or other income, health and general wellbeing and experience of crime.

Surveys were distributed by post in English and a second language if there was a record that the recipient spoke one of ten languages available. If no response was received, a reminder and another copy of the survey were sent. Only those who responded to a survey were included in the next follow up. If surveys were received more than 12 weeks after distribution they were

[24] Note that age groups are not equal size.

discarded from the analysis. A form to report change of address was included with the postal survey.

5.2.3 Response rates and attrition

Longitudinal data is good for assessing the average impact of a policy intervention. SNR is a panel survey, meaning that the same individuals were interviewed in multiple waves over time. However, there are a number of potential problems, such as tracking individuals over time and maintaining cooperation of respondents which can affect response rates and therefore the analytical power of the data. Table 5.1 shows the response rates at each stage of the survey.

The baseline survey recorded data on the demographic characteristics of refugees but also on their lives before arriving in the UK. It is important to note that the length of time spent in the UK varies between individuals, as do the dates at which they completed each survey.

Table 5.1: Survey of New Refugees response rates.

Number of new refugees		9127
Number contacted		7765 (85.1%)
Number responded	baseline	5678 (73.1%)
	8 months	1840 (32.4%)
	15 months	1259 (68.4%)
	21 months	939 (74.6%)
Number responded to all previous waves	8 months	1826 (32.2%)
	15 months	1173 (64.2%)
	21 months	867 (73.9%)

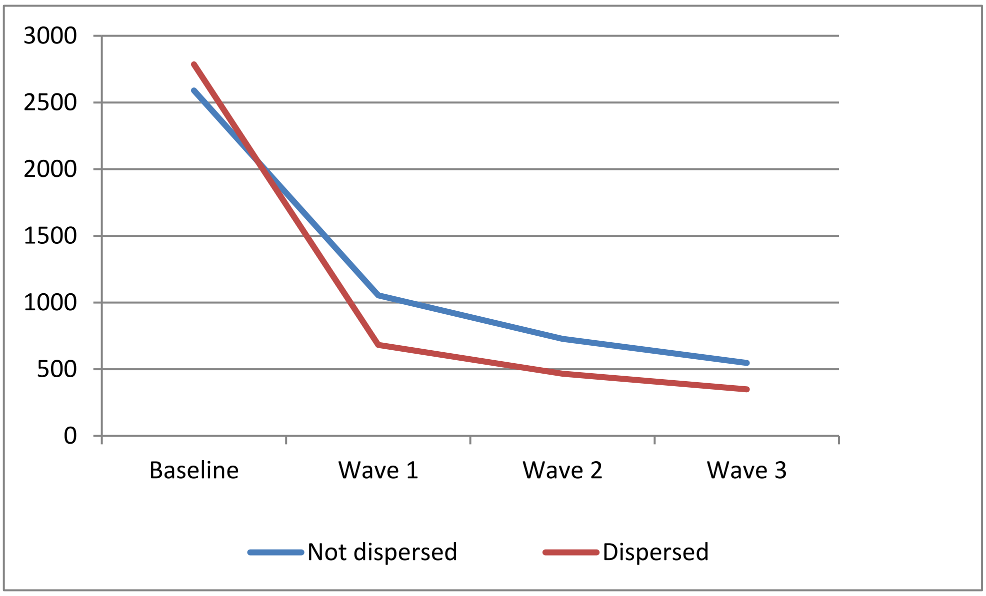
Source: SNR data (Home Office, 2010e).

Non-response through attrition is cumulative and therefore more problematic with longitudinal data and increases with additional waves. Table 5.1 shows that the largest attrition was observed between the baseline and the first follow up.

This dataset includes information on whether a refugee was living in dispersal ('NASS') accommodation at the time of the baseline survey. The baseline surveys were sent within one week of a decision, and it is therefore reasonable to assume that almost all respondents would still be living in the residence they occupied as asylum seekers. At the baseline survey 47.5

percent of respondents were living in NASS accommodation, and Figure 5.2 shows the higher rate of attrition for those dispersed.

Figure 5.2: Attrition of those living in dispersal accommodation and those who were not.



Source: Author's creation from SNR data (Home Office, 2010e).

The problem of attrition is generally a challenge in longitudinal surveys (Watson and Wooden, 2009) but has some unique elements for the refugee population: those who were living in dispersal accommodation at the baseline survey are required to move into alternative housing within a short time of their decision. This, along with the higher mobility observed among new migrants, means that there is a greater chance that they will not be followed to their new address. Characteristics associated with a higher rate of attrition therefore need to be identified so that any bias may be considered in the analysis of the longitudinal data recorded in this survey.

5.2.4 Weighting

When introducing measures to address the problem of attrition it is important to recognise the underlying assumptions. Data can only be described as missing completely at random where there is no association between the non-response and any observed or unobserved data; if the missingness can be fully explained by the observed data then this is missing at random; otherwise it can be considered missing not at random (Twisk, 2003). If it is possible to assume that the parameters which explain the missingness (of data that is missing at random or missing completely at random) are independent from the parameters of interest for analysis, then there is no bias introduced to the analysis and missingness can be ignored. This is not the case for SNR data, as shown in section 5.2.3; non-response at each wave can be expected to introduce bias and there is a cumulative effect of non-response bias over the different waves (Home Office, 2010d). For longitudinal data some of the missingness can be addressed through weights based on information about non-respondents from previous waves; this is described as 'informative missingness' (Twisk, 2003, p. 205).

A series of weights were created to address these different aspects described above (see Appendix F and Home Office, 2010d, for full details). An initial non-contact weight was generated by modelling the relationship between a set of predictor variables[25] and an outcome of whether or not a questionnaire was issued, giving a predicted probability of contact, the inverse of which provides the weight. This was combined with a non-response weight (generated using a similar method but with response as the outcome) and calibrated to give a baseline weight which, when applied to the baseline data, gives a distribution for key characteristics (the predictor variables) which is the same as the total population of new refugees granted. Similarly, a cross-sectional weight was created for each subsequent wave to try to take account of non-response bias.

Weighting of the longitudinal data involves excluding any cases that did not respond to all previous waves of the survey, before adding the longitudinal weights created for that wave. These were therefore generated only for those respondents who responded at that and all previous waves, meaning that non-. Response behaviour was again modelled using logistic regression using data from previous waves as predictors of response outcome[26]; this is informative missingness as the information from previous waves can be used to mitigate bias introduced through attrition. It is important to note that non-response (and therefore non-response bias) is cumulative over survey waves which means that in later waves, weights need to be larger and more variable, impacting on the sample efficiency and precision (Home Office, 2010d).

Finally, it should be noted that some of the data collected in the SNR is retrospective. For example, the baseline survey asks 'Have you ever needed any kind of help or support from any of these groups or organisations?' Gathering retrospective information may be subject to recall error, and the possible bias resulting from this need to be acknowledged.

---

[25] UKBA region, treatment group (i.e. whether offered and/or accepted support from SUNRISE), age (grouped), sex and continent of origin.

[26] See Home Office (2010d, p. 7) for full list of predictor variables used in weights for each wave.

5.3 Methods

The analysis of longitudinal data can provide unique opportunities; it can be particularly useful for policy analysis where the focus is on identifying impact, ideally in order to establish causality rather than just identify associations (Lynn, 2009). A range of methods for analysing longitudinal survey data are utilised here to address the three research questions. Descriptive analysis of cross-sectional data and plotting multiple waves for comparison, as well as cross-sectional and longitudinal models, will provide the most complete understanding of dispersal and refugee outcomes. This is carried out alongside a thorough assessment of the attrition observed and any bias which it introduces.

5.3.1 Cross-sectional baseline analysis

The purpose of this cross-sectional analysis is to model data from the baseline survey to identify factors associated with being dispersed. A multinomial logistic regression model is used with backwards stepwise selection and the outcome variable is ‘accommodation type’ (NASS dispersal accommodation, living with family, living with friends, other). The choice of variables is informed by previous research as well as findings of the descriptive analysis. They include demographic characteristics as well as information about individual experience before arrival in the UK and during the asylum process. Cross-sectional baseline weights which attempt to address non-contact and non-response bias will be applied to the data before analysis.

5.3.2 Investigating attrition

In order to understand how these longitudinal data can be analysed, it is essential to carry out a thorough assessment of the attrition observed and any bias that is introduced as a result. This investigation of attrition and any measures introduced to address it must take into consideration the overall aim and specific purpose of the analysis to be carried out. For this research, the aim of improving our understanding of dispersal means that the characteristics found to be associated with dispersal status are the most important variables to focus on. Therefore, descriptive statistics will be presented in order to compare the population of baseline respondents with those remaining at wave three of the survey; this allows identification of differences in those who have dropped out. In this way, it will be possible to mitigate any bias that has been introduced when it comes to longitudinal modelling of refugee outcomes. The populations will also be compared after weights provided with the dataset for the baseline survey and the longitudinal weights for wave three (described above) have been applied; if these effectively counteract any bias introduced through attrition then they may be used for the longitudinal models.

5.3.3 Longitudinal modelling

When modelling longitudinal data with repeated measures, it is important to take into account the structure of the data and apply methods that recognise the 'groups' and link the responses of individuals over time. Using a random effects model allows the residuals for the same individual to be correlated (Rabe- Hesketh and Everitt, 2004). Random-effects modelling gives the individual level effect on the outcome variables; marginal-effects models which summarise the population average effect have a different interpretation but could also be used here (Molenberghs and Verbeke, 2005). That said, the main aim of this analysis is to highlight the characteristics associated with dispersal and compare the impact of weighting; therefore this analysis will just focus on random effects models.

Variables to be included are chosen in the context of descriptive and cross-sectional results and analysis of attrition. Longitudinal weights provided with the dataset for the appropriate combination of waves (in this instance, baseline and all subsequent waves) may be applied in order to address bias introduced through attrition. Outcome variables are selected for modelling based on previous studies and literature as well as an understanding of the dispersal policy and settlement process.

Employment is reported to be one of the most important factors in successful settlement and integration (Home Office, 2005; Da Lomba, 2010) and benefit claims reflect the level of ongoing support needed. The models of economic activity exclude those who were inactive (i.e. caring for family or retired) and is recoded into the binary variable: employed (including those in part or full-time work or self-employed) and unemployed (seeking work). Anie et al. (2005) describe 'successful' dispersal as not being the victim of physical or verbal attack; while this is a limited definition, it is modelled here for further investigation of the concept. Health was selected as a useful self-reported measure of wellbeing and also following the identification in the literature of local health services as an important aspect of deprivation. For this analysis, health is recoded into the binary variable: very good or good and fair, bad or very bad. Background variables selected for inclusion in the models are age (grouped), sex, country of origin and accommodation.

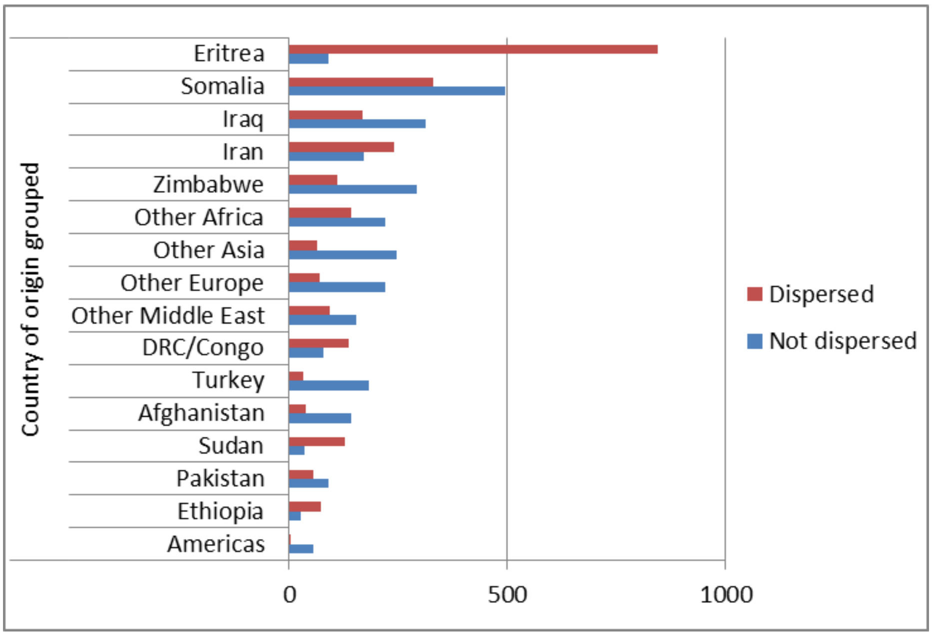
5.4 Results

5.4.1 Cross-sectional baseline analysis

Understanding the dataset and the characteristics of respondents provides essential context for subsequent more complex analysis. Initial descriptive analyses presented in Figures 5.3, 5.4 and

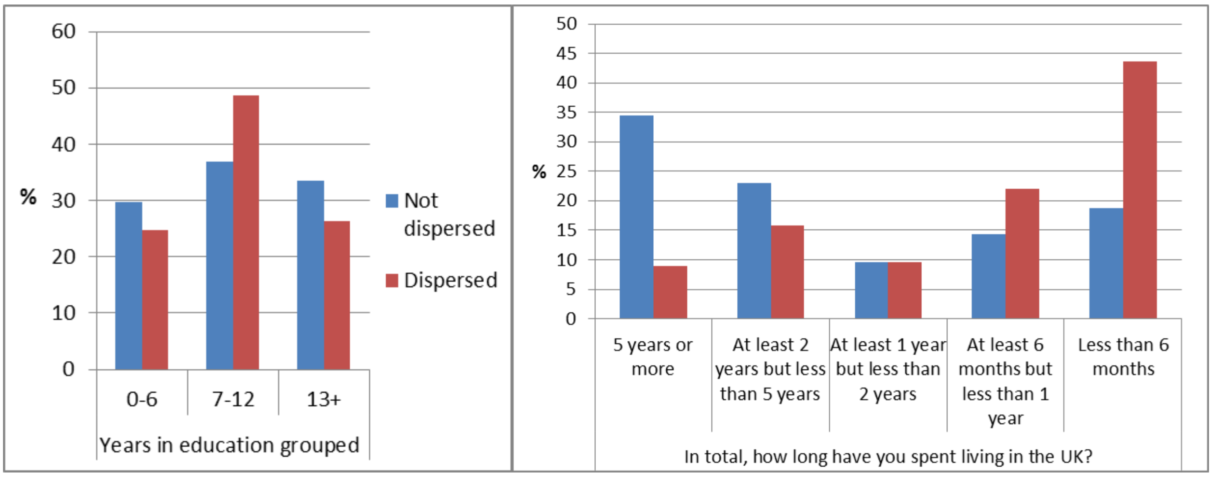
5.5 allow a comparison of respondents by different background characteristics and highlights important variables to explore further.

Figure 5.3: Number dispersed/not dispersed by country of origin, grouped.



Source: Author's creation from SNR data (Home Office, 2010e).

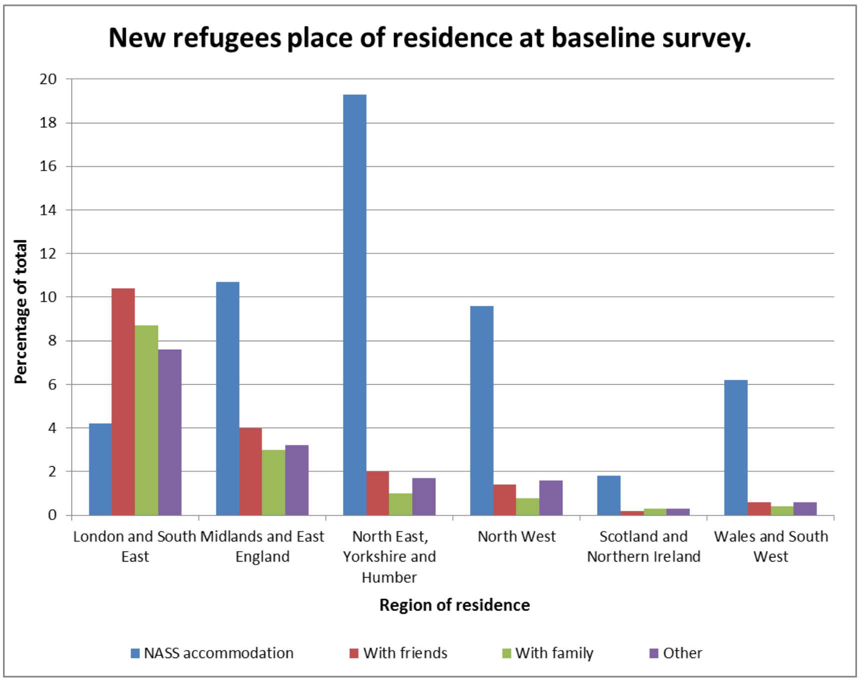
Figure 5.4: Percentage of those dispersed/not dispersed by years in education and by



time spent in the UK.

Source: Author's creation from SNR data (Home Office, 2010e).

Figure 5.5: Region of residence by accommodation.



Source: Author's creation from SNR data (Home Office, 2010e).

The descriptive analyses suggest that there are differences in the background characteristics of those refugees who were dispersed as asylum seekers and those who were not. For example, a high proportion of refugees from Eritrea experience dispersal; those living in the UK for two years or more are less likely to be dispersed; lower levels of dispersal are seen among those living in the London and the South East.

In order to provide a more formal assessment of whether the differences observed in the descriptive analyses are important, the data is modelled using multinomial logistic regression and the results are presented along with formal statistical tests in Appendix F. Cross-sectional baseline weights are applied to the data for this analysis to try to address any non-response and non-contact bias.

The models produced include background and contextual characteristics that are identified as important from the literature and initial exploratory analysis, or are considered important to control for as they could influence these relationships. Length of residence and ability to speak English are likely to be highly correlated; ability to speak English was not significant at the five percent level in a model including both variables (see Appendix F, Table 7.16) and a second model including an interaction between these two variables was also tested but was also not found to be significant and therefore removed, leaving just length of residence. The model presented here (and table 7.17) shows the characteristics associated with living in NASS dispersal accommodation, compared to living with family, found to be significant at the five percent level.

This indicates that these results are unlikely if there is *not* an association between the explanatory variables and the outcome (being dispersed).

Characteristics positively associated with being dispersed:

- Age (particularly 25-34 and 35-44 years)
- Country of origin: Particularly Eritrea and Ethiopia (also Sudan, DRC/Congo, 'Other Middle East', 'Other Africa', Iran, 'Other Europe' and Iraq)
- Needed help or support from organisations[27]

Characteristics negatively associated with being dispersed:

- Better health (particularly very good health)
- Longer time spent in the UK (particularly more than five years)
- Having friends or relatives in the UK
- Meeting with friends more regularly
- Meeting with relatives more regularly
- Living with a partner
- Living in London and the South East or the Midlands and East England

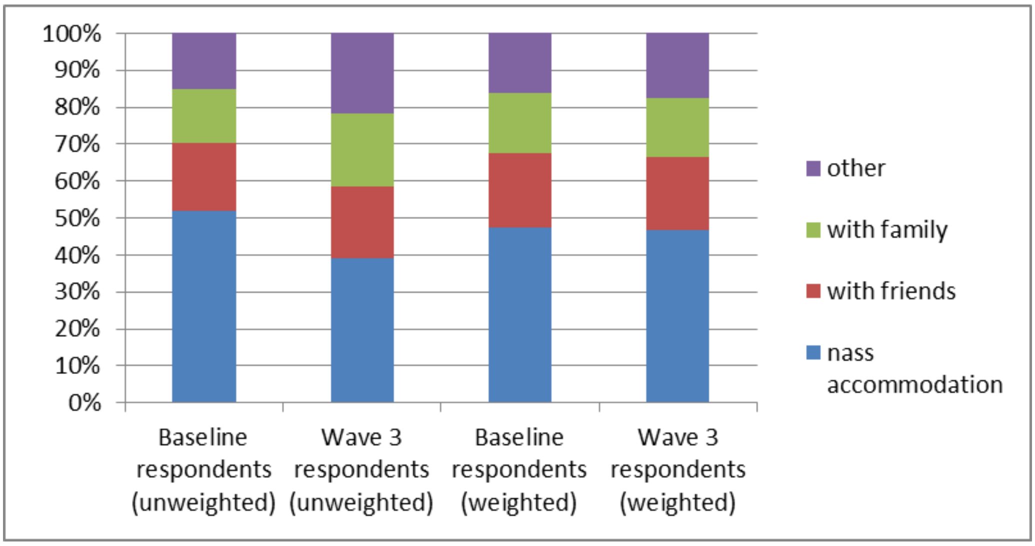
These results show that background characteristics related to time before arrival in the UK, as well as indicators of experience during the asylum application process, are associated with dispersal status. Understanding the baseline data can provide useful context for exploring the nature of attrition over subsequent waves of the survey.

5.4.2 Investigating attrition

Comparing bars one and two shows the bias introduced when the respondents who drop out are different from those who remain; any reduction in the difference observed in bars three and four shows the extent to which the weights successfully address this bias. The unweighted data in Figure 5.6 shows that those living in NASS (dispersal) accommodation at the time of decision were more likely to drop out by wave three. The weighted data however is closer to the patterns of accommodation observed at the baseline.

[27] This includes the following help or support: financial, legal advice, transportation, information, meeting people, food/clothing, help finding work/housing, language or emotional support.

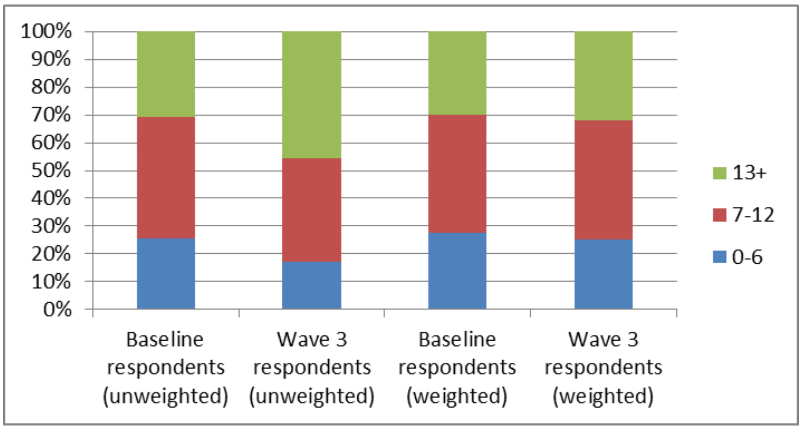
Figure 5.6: Baseline accommodation of baseline and wave three respondents.



Source: Author's creation from SNR data (Home Office, 2010e).

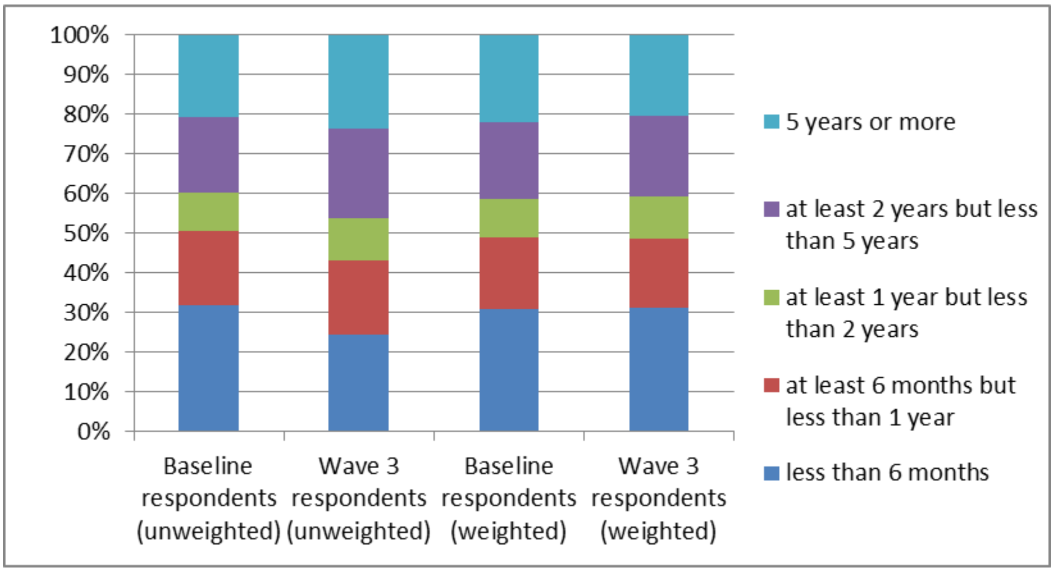
A similar pattern is observed in Figures 5.7 and 5.8 where the greater likelihood to drop out of those with fewer years in education or shorter time spent in the UK is addressed through the use of weights.

Figure 5.7: Years in education reported at baseline of baseline and wave three respondents.



Source: Author's creation from SNR data (Home Office, 2010e).

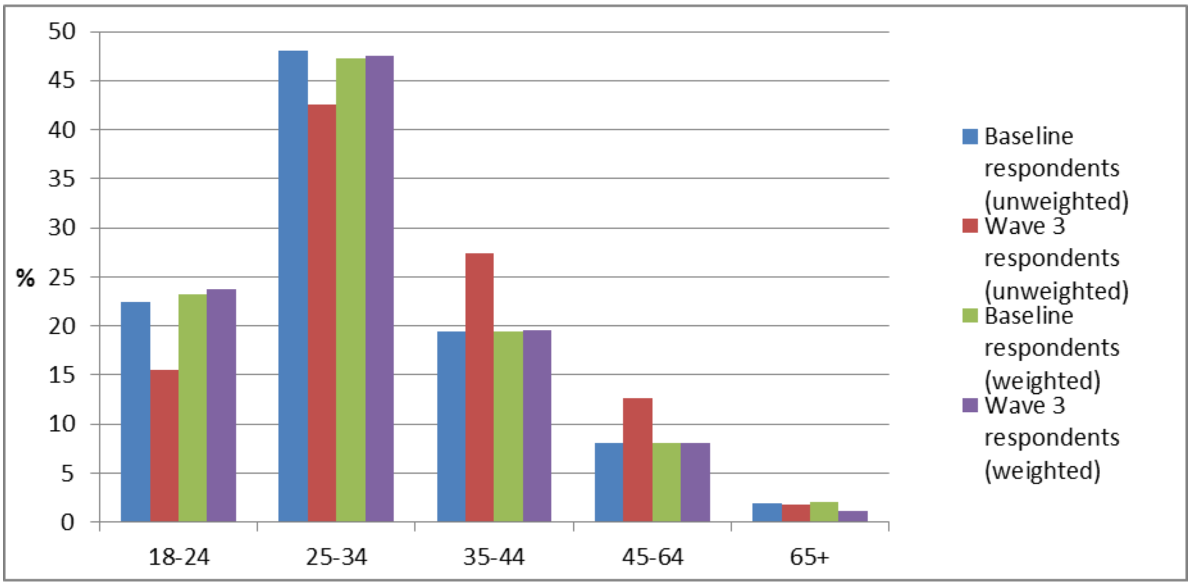
Figure 5.8: Length of residence reported at baseline of baseline and wave three respondents.



Source: Author's creation from SNR data (Home Office, 2010e).

Age is reported at the baseline (see Figure 5.9) so will remain static over time. In reality, respondents will age over time but the purpose here is to investigate how attrition varies and therefore the static measure is appropriate. The bars showing weighted data are notably more similar than the unweighted bars for every age group except 65+ (which is a very small proportion of respondents).

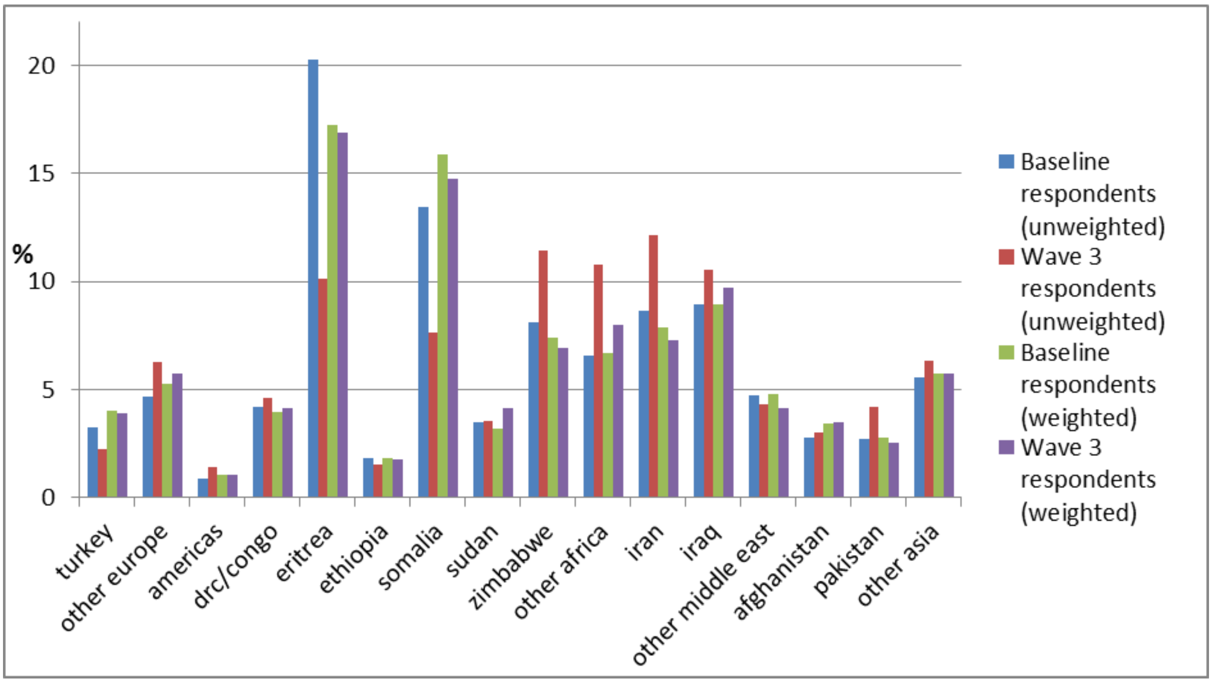
Figure 5.9: Age at baseline of baseline and wave three respondents.



Source: Author's creation from SNR data (Home Office, 2010e).

The patterns observed in Figure 5.10 are more complex with a greater number of categories, but for most countries of origin the weighted bars show that some attrition bias has been addressed.

Figure 5.10: Country of origin of baseline and wave three respondents.



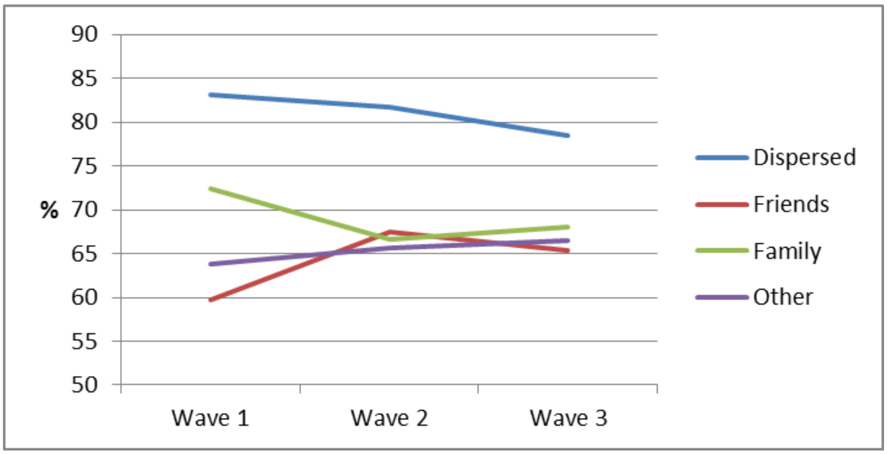
Source: Author's creation from SNR data (Home Office, 2010e).

Overall, the results presented here show that for almost every variable, the difference in the baseline and wave three populations is reduced by applying the two weights; this reduces the bias introduced by differences in propensity to drop out. Differences in likelihood of dropping out related to background characteristics need to be addressed in any subsequent analysis of outcomes, in order to establish whether variation observed over time is a result of differences in those who drop out, or can be explained by particular characteristics.

5.4.3 Longitudinal modelling

Descriptive graphs showing change over time by accommodation status highlight different patterns in a selection of outcome variables that are observed. Those refugees who were dispersed report higher rates of benefit claims (see Figure 5.11) than those in other accommodation categories. This corresponds to what we know about dispersal: those requiring housing support are inherently more reliant on the state for essential resources.

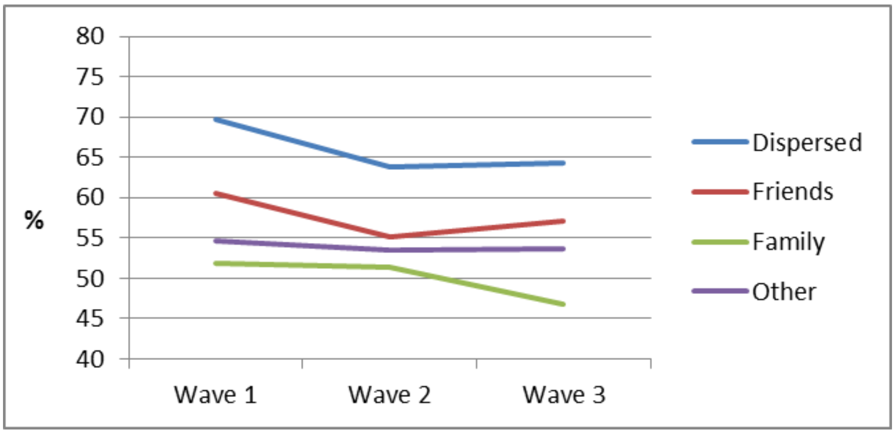
Figure 5.11: Percentage receiving benefits.



Source: Author's creation from SNR data (Home Office, 2010e).

Figure 5.12 also highlights dispersed refugees as having a higher percentage perceiving their job to be below their skill level. This pattern suggests an unfulfilled economic potential among this group. It is important to note however that these descriptive graphs are presented for initial exploratory analysis only and the uncertainty of the results will increase over time as the sample size decreases due to attrition; there is also greater potential for non-response bias.

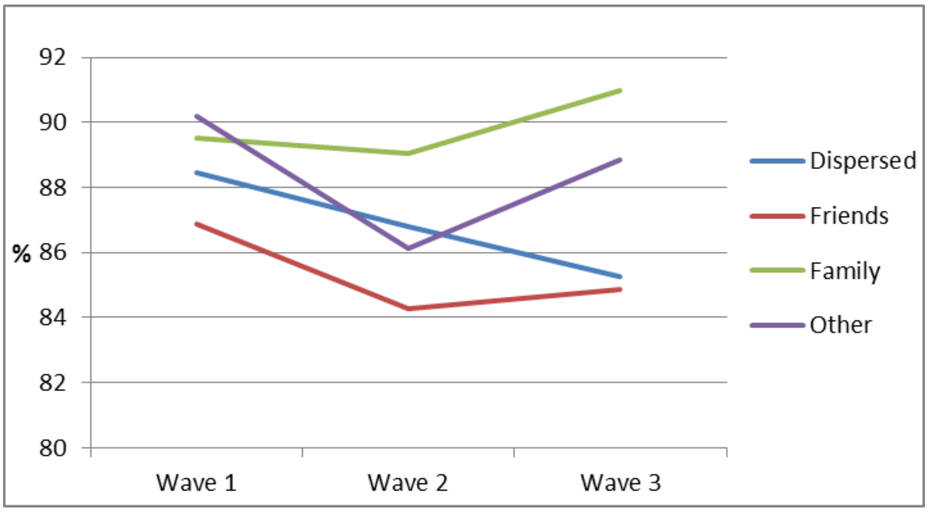
Figure 5.12: Percentage perceiving their job to be below their skill level.



Source: Author's creation from SNR data (Home Office, 2010e).

The measure of whether respondents want to stay in their current location can be seen as a summary indicator of satisfaction and wellbeing. Figure 5.13 shows that while the percentage of dispersed refugees wanting to stay decreases over time, all other groups increase at wave three. Further investigation is required to identify whether this relationship is still apparent after controlling for other possible explanatory variables.

Figure 5.13: Percentage wanting to stay in the same town or city.



Source: Author's creation from SNR data (Home Office, 2010e).

In order to assess whether these observed differences over time are still evident when other variables are controlled for, longitudinal modelling of the data can be useful. Random effects logistic regression models are used to explore the relationship between background characteristics and observed outcome variables, taking into account the data structure. The first models were run on unweighted data, with no attempt to reduce non-response bias. Table 5.2 shows which of the independent variables were significantly associated with the outcome when modelling the unweighted and weighted data. The second model is run on the same data (those who responded to all four waves) but with longitudinal weights applied to attempt to mitigate any bias introduced through attrition, using information on non-respondents from previous waves. Results of all models are presented to show the impact of weighting; including all four independent variables in each model is interesting to explore the associations between these and the different outcomes. Furthermore, the strength of background literature and the preceding analyses means that all of these are theoretically relevant and justifies keeping those that are not statistically significant in this context.

Table 5.2: Models of unweighted data and with longitudinal weights to address non response bias due to attrition.

Model of outcome	Independent variables: significant associations at the 5% level in **bold**	
	Unweighted	Weighted
Receiving benefits	**Age, sex, country of origin, accommodation**	**Age, sex, country of origin, accommodation**
Economically active (employed)	Age, **sex**, country of origin, **accommodation**	**Age**, sex, country of origin, **accommodation**
Victim of attack	Age, sex, country of origin, accommodation	Age, sex, country of origin, accommodation
Health (very good or good)	**Age, sex,** country of origin, **accommodation**	**Age, sex**, country of origin, accommodation
Want to stay in same town/city	**Age**, sex, **country of origin**, accommodation	Age, sex, country of origin, accommodation
Perception that job is below skill level	**Age**, sex, country of origin, **accommodation**	Age, sex, **country of origin**, **accommodation**

Source: Author's analysis using SNR data (Home Office, 2010e).

While weighting can account for some of the attrition bias, it cannot remove it all together as only selected variables are included in the weighting models and there may also be unobserved factors that are related to the propensity to drop out. Therefore caution is required in the interpretation of results, with or without weighting, as reported associations between background and outcome variables may be influenced by other factors.

5.5 Discussion and conclusions

5.5.1 Background characteristics associated with dispersal

Understanding and identifying differences between characteristics of dispersed and non-dispersed refugees not only increases our understanding of the composition of these populations but also means that they can then be taken into account and controlled for in subsequent modelling and longitudinal analysis. This will contribute to more robust conclusions about the impact of dispersal and enable us to move beyond simple description of associations to consider the transition from experience as an asylum seeker to integration and outcomes as a refugee (Da Lomba, 2010; Stewart, 2011).

Significant differences are observed in background characteristics between those who are dispersed and those living with friends, family or another living arrangement (‘other’). This is supported by previous research (such as Cheung and Phillimore, 2013) which has highlighted the importance of existing networks and communities, particularly as the availability of friends or family to live with on arrival directly impacts the need to rely on dispersal accommodation. Therefore, it is unsurprising that country of origin is significantly associated with being dispersed; some nationalities, such as Pakistan, have large communities already living in the UK, but those with a relatively more recent arrival history (for example Eritrea) may not be sufficiently established to provide accommodation to the newest arrivals. The age of refugees may also be related to whether they have links to networks and whether they are arriving with family members, so affecting their need for accommodation.

Length of residence before receiving a decision also varies by dispersal status and is important to consider with regard to subsequent outcomes; Zetter et. al (2005, p. 176) suggest that an extended period as an asylum seeker keeps people ‘disconnected’ from accessing support and services required for integration. Longer time spent in the UK is associated with a lower probability of being dispersed; some of these individuals may have moved from dispersal accommodation to live with friends or family as they have made connections since arrival, or there may have been changes in policy implementation.

One of the main aims behind the introduction of the dispersal policy in 1999 was to move the focus of settlement away from London and the South East. It is clear from these results that the patterns of settlement for dispersed and non-dispersed refugees are different: those retaining a choice of location remain more likely to settle in London and the South East but also the Midlands and East England.

Identifying differences in baseline characteristics of those who were dispersed and not dispersed allows a better understanding of the population affected by this policy and gives useful context for longitudinal analysis of outcomes over time.

5.5.2 Attrition

Attrition observed in this dataset is higher than that of many large surveys in the UK (Lynn, 2009). The differences observed in the characteristics of respondents responding to the baseline survey and wave three show that we cannot assume that they have dropped out randomly, without any relation to the background variables. Therefore, the longitudinal data available for the remaining population at wave three is not representative of the whole refugee population surveyed. The purpose of weights included in the dataset is to address any non-contact and non-response bias.

Initially applying the baseline weight is useful as this means that the distribution of the baseline data matches the total population of new refugees for a set of key variables (Home Office, 2010d). The Survey of New Refugees is effectively a census of the whole population of refugees who received their decision between 1st December 2005 and 25th March 2007, therefore the aim is to achieve a population of respondents that is as close as possible to the sampling frame. In fact, the information on country of origin used for the baseline weight is more detailed than that presented in the survey data and is therefore likely to be better at taking account of variation (Home Office, 2010d).

The results show that a range of variables are related to whether respondents drop out. Those who were dispersed, those with fewer years in education and those aged 18-34 were particularly likely to stop responding to subsequent waves of the survey. These characteristics may be related to the greater mobility of young adults and those needing to leave dispersal accommodation. However, for almost every variable presented here, the differences observed in the characteristics of respondents at the baseline and wave three have been reduced for the weighted data. This shows that they are useful for addressing bias introduced through attrition for the variables presented. For countries that have a population register alongside the Labour Force Survey, for example in Norway, information from the register may be used to reduce bias due to non-response and increase the accuracy of estimates (Thomsen and Zhang, 2001). In this longitudinal dataset, the fact that attrition is cumulative over time means that bias is more likely with each wave and the amount of intervention (i.e. the size and variation of the weights) required to mitigate this increases, reducing the efficiency of the sample (Home Office 2010d). Furthermore, although weights can be introduced to address bias on observed characteristics, any unobserved bias that cannot be addressed will also increase with attrition over time.

5.5.3 Dispersal and social and economic outcomes

The results presented here support findings in previous studies which suggest that some indicators of social and economic outcomes are associated with dispersal status (Cheung and Phillimore, 2013; Home Office, 2010b). Models run on weighted data attempt to address any bias that is introduced when factors associated with whether an individual responds to the survey are also associated with the outcome variable of interest.

The results of the longitudinal models presented here show that accommodation is related to levels of employment among refugees. Actively participating in the workforce is widely considered crucial for integration (Ager and Strang, 2004; Home Office, 2005) and Phillimore et al (2003) highlight some of the barriers to education and employment that refugees face. Higher levels of unemployment are related to higher deprivation levels (DCLG, 2008; 2011) and therefore this relationship is unsurprising in the context of the patterns of dispersal and deprivation identified in Chapter 4. Those seeking employment are eligible for certain benefits, but the economically inactive population and those in work may also qualify.

The characteristics of individuals and their situations are intrinsic to many benefits which are paid in the UK, such as the state pension, child tax credits or housing benefit. This is reflected in the findings that age, sex, country of origin and accommodation[28] were all significantly associated with whether a respondent received benefits. Once a decision has been received and refugees have moved out of dispersal accommodation they must be supported by the LA. Clearly the higher levels of benefit claims among those who were dispersed is directly related to this. This also exacerbates the pressure on services in these areas as LAs do not receive any additional funding (HC Home Affairs Committee, 2017).

No significant associations were found between the variables included in the models and whether individuals were the victim of attack. That said, while this measure may be useful for contributing to a picture of how successful settlement is, it is a limited indicator of integration or wellbeing more broadly. Furthermore, it is not possible to assess geographic variation across dispersal areas which might show some interesting patterns (Anie et. al, 2005). Greater geographic detail would also be useful for analysing whether respondents want to stay in the same town or city; a lack of services and the presence of crime have previously been linked with experiencing social exclusion (Hynes, 2011), and while different locations may rate better or worse for this indicator, there is no observable broader pattern in relation to the variables included here.

The focus on dispersal to low value properties (albeit outside London) along with the inability to broaden dispersal to a wider number of LAs means that greater pressure has been put on services such as schools and healthcare. Middlesbrough is one example of where dispersal is placing 'enormous strain' on already deprived areas (HC Home Affairs Committee, 2017, para. 38) and if refugees remain living here over time, this can be expected to have an increasing impact on their wellbeing and outcomes. That said, health is a subjective, self-reported measure and can fluctuate over time; the findings from modelling weighted data show that while age and sex are important predictors of health outcomes, accommodation is not. Whether these relationships change beyond 21 months following a decision is not possible to establish from this dataset.

The ability to confidently apply findings from the analysis of this dataset to today's refugee population is becoming increasingly difficult. This data represent a specific cohort arriving

[28] NASS accommodation (i.e. dispersed), with friends, with family, other.

between defined dates and for specific reasons based on the context of conflict and persecution taking place in certain countries under certain regimes. These refugees now received their decisions at least ten years ago and arrived even earlier and the world has changed over the last decade; the drivers causing migrants to move, as well as the ability for individuals to undertake the journey and choose the UK as a destination are not static. Furthermore the economic, social and policy context of the UK has moved on and therefore understanding the processes and mechanisms which cause the patterns which are observed in the data are at best out of date but also arguably now of limited relevance or usefulness.

While the survey of new refugees remains the most comprehensive source of information on the refugee population in the UK, and dispersal in particular, it is crucial that our understanding moves on in order to make cogent and useful recommendations for the successful settlement of current and future asylum seekers and refugees.

Chapter 6 The potential for combining data

6.1 Introduction

The limited availability of data is a key challenge faced by those researching asylum seeking and refugee populations in the UK and has been a recurring theme of this research thus far, both in the review of existing studies and literature but also as it has often restricted the scope and depth of new analyses presented here. In recent years there has been an increasing focus on how to 'unleash the potential' of existing data for both research purposes and policy analysis (HM Government, 2012; BIS, 2013, p. 4). Policymakers have repeatedly stated the intention that data should be made available and that legislation should allow for research in the public interest whilst also protecting individuals' data.

Therefore, the aim of this third substantive chapter is to explore additional data sources that exist and to assess the feasibility of methods for combining data to contribute to an increased understanding of the dispersal policy and its impact. The potential to maximise available information could be useful for researchers as well as government statisticians and policymakers. In making any recommendations it is important to link up needs, in terms of data scarcity and gaps in knowledge, with the possibilities for data commissioning and suggestions for new approaches to analysis which can address those needs.

The following research questions address the three main elements of this challenge:

1. How can combining datasets on asylum seeker and refugee populations help us further explore dispersal?
2. What are the trade-offs between additional information gains and the errors of estimation that are introduced?
3. What are the policy and research implications of findings from analysis of combined data on dispersal?

The first research question directly addresses the limited availability of data on refugees and asylum seekers in the UK. The additional 'Whyuk' variable added to the quarterly Labour Force Survey (LFS) in 2010, available through the UK Data Service (UKDS) Secure Lab, is identified as a key source for further analyses, although this dataset does not include a dispersal indicator. This dataset is utilised to assess the feasibility of a variety of statistical analytical methods for combining datasets which could be applied to existing sources in order to maximise the information currently published on asylum seekers and refugees. Three possible options for combining sources are considered: individual data linkage is presented with a discussion of the potential for analysis of a resulting dataset, subsequently, an illustration of how information on individuals can be 'borrowed' shows how successfully dispersal status can be predicted and finally a method that utilises cell structures to combine information on aggregates has potential to allow further analysis of outcomes.

Recommendations for methodological improvements and suggestions for additional, targeted data collection are made based on analysis of information gains and errors involved in the second research question. An in-depth discussion of the policy implications of the findings from the first two research questions then offers suggestions for expanding our knowledge and understanding of dispersal.

6.2 Data: alternative sources to analyse dispersal

The datasets analysed in previous chapters are relatively easy for researchers to access as they are published either on the government website or by signing up to the UKDS. As we have seen, the most comprehensive available source for the analysis of dispersal, in spite of its limitations, is the Survey of New Refugees. In the UK, additional data on asylum seekers and refugees collected in the LFS may be accessed via the UK Data Service with 'Secure Researcher' accreditation, which is required for teams involved in research. This data is available through the UKDS virtual 'Secure Lab,'[29] meaning that researchers must apply for access and undergo training in data security in order to work with the data in a secure virtual environment.

6.2.1 Limitations identified from the analysis of the Survey of New Refugees

Analysis of data from the SNR in Chapter 5 indicates that a number of background characteristics (such as country of origin and region of residence) are associated with living in dispersal accommodation at the time of decision (between 1st December 2005 and 25th March 2007). Differences in subsequent outcomes for those who were dispersed are also highlighted. In spite of these findings, it is clear that the high rate of attrition in the survey, particularly for those who were dispersed, limits the ability to make robust conclusions from longitudinal analysis. Furthermore, the fact that the SNR is a one-off survey means that findings are increasingly out of date.

The identification of characteristics associated with dispersal has clear policy use, as described in Chapter 5, but also has the potential to be exploited further in the analysis of additional datasets

[29] See: https://discover.ukdataservice.ac.uk/catalogue/?sn=6727&type=Data%20catalogue

where an indicator of dispersal experience is not recorded. In order to best exploit this information, harmonisation of variables across both datasets is essential. It is not possible to harmonise all desired variables adequately for inclusion in analysis across multiple sources, particularly where the original data is not available. For this analysis, where the purpose is to illustrate potential methods which might be useful, the ability to harmonise is a requirement for the inclusion of variables.

6.2.2 Labour Force Survey and Annual Population Survey

The Labour Force Survey (LFS) is the largest household survey in the UK and primarily gathers information to provide labour market and economic indicators at a national level[30], such as rates of employment. Approximately 41,000 households are surveyed each quarter, and each household is surveyed over five consecutive quarters (ONS, 2016b). The addition of a ‘sample boost’ to increase the number of respondents to a minimum level in each local area produces the Annual Population Survey (APS) dataset. The current structure of the APS includes waves one and five of the quarterly LFS plus the Local Labour Force Survey (LLFS) for England, Wales and Scotland[31].

Since 2010, the APS has asked those born abroad what was their main reason for migration. The survey also records the year in which the respondent first arrived but does not report current status. This means that it is not possible to identify whether an individual is a refugee or is still in the process of seeking asylum. There may also be cases where the initial reason for migration was not to seek asylum (e.g. they arrived to study or to join family) but a subsequent ‘in-country’ application was made and a change in status was granted. That said, the number of people who switch status is a very small proportion of the refugee population and can be disregarded for this research.

The APS dataset for Jan-Dec 2014 includes 16,001 respondents who arrived in the UK since 2000 when the dispersal policy was first introduced (see Table 6.1); 805 of these came to seek asylum. This group of asylum seekers and refugees can therefore be considered as ‘at risk’ of dispersal.

[30] England, Wales, Scotland and Northern Ireland.
[31] See ONS (2016b) p. 20 for diagram illustrating structure of LFS waves that make up the APS.

Table 6.1: Main reason for migration by period of arrival of those aged 18+, APS 2014.

Main reason for coming to UK	Year of arrival		
	Before 2000	2000 to 2014	Total
Employment	2,984	6,346	9,330
Study	1,621	2,464	4,085
Get married/form civil partnership	1,041	931	1,972
As a spouse/dependent of UK citizen	4,143	2,079	6,222
Spouse/dependent of someone coming to UK	3,137	2,103	5,240
Seeking asylum	674	805	1,479
Visitor	458	378	836
Other	1,373	895	2,268
Total	15,431	16,001	31,432

Source: APS data (ONS, 2018).

Utilising variables which are as close as possible to those recorded in the SNR will enable close comparison and analysis of changes occurring over time. While the APS can provide a wealth of additional and relatively up-to-date information on a sample of the refugee population, it does not record whether an individual was dispersed; therefore this analysis explores whether combining it with other sources of data can contribute to an increased understanding of the dispersal policy and its impact.

6.3 Methods: exploiting existing data through combining sources, a feasibility study

In the context of limited data availability it is essential that existing datasets are fully utilised, not only to potentially provide immediate new insights but also to maximise the return on resources which have been put into the collection and production of such data.

6.3.1 Linking data on individuals (Administrative Data Research Network)

Aside from the (often expensive and time-consuming) collection of new data, another possibility for increasing the analytical potential of what is currently available is data linkage. Where it is possible to link data at the individual level, this is the ideal method for utilising information from multiple sources. The process matches the same individual (or 'case') in two or more datasets using a unique identifier, such as social security numbers, or a combination of non-unique identifier variables, such as name, sex and race (Gomatam et. al. 2001). The resulting dataset includes variables from both datasets for all linked cases.

The Administrative Data Research Network (ADRN)[32] is intended to facilitate researchers' access to the necessary mechanisms required for data linkage, where it can be shown that the research has the potential to benefit society. The negotiations for access to datasets for linkage are carried out on a 'case-by-case' basis, with each data owner approached by ADRN on behalf of the researcher for a specific project. Subject to receiving these permissions, the process should result in an anonymised, linked dataset being made available for analysis within a secure environment.

The possibility of creating a dataset linking the SNR with data from the 2011 Census and the English Indices of Deprivation in order to analyse the relationship between dispersal, deprivation and outcomes was pursued at length for this project. As described previously, the SNR includes an indicator of whether an individual was living in dispersal accommodation during their application for asylum; the primary aim of linking this with census data would be to add information on the deprivation levels of the locations to which they were dispersed. This could provide a considerable improvement in understanding of the dispersal policy, the choice of dispersal locations and the impact of dispersal and deprivation levels on participants' subsequent outcomes. Beyond this, there would be the potential to use data collected in the 2011 Census to analyse refugee outcomes at a later time point (the SNR ended in 2010), for example employment, housing status and health. This would provide further scope for analysis of the experience of dispersed asylum seekers after receiving a positive decision, as well as informing an assessment of the policy aims and implementation. However, the fact that consent to link SNR data had not been sought from respondents at the time of collection meant that ADRN considered the proposal unfeasible. The intention of policymakers has been to facilitate the re-use of existing data, for example through amending the Freedom of Information Act 2000; in spite of this, at the time of writing, the requirement for expressed, informed consent to have been given by each individual is currently still a major barrier to data linkage and as a result, it is necessary to consider in detail what can be achieved with the data that is currently available.

6.3.2 Borrowing information on individuals

Record linkage and individual level matching are not currently possible for the analysis of dispersal; therefore, it is important to consider in what other ways it might be possible to 'borrow' information in order to further our understanding of this policy and its impact. Unlike the

[32] The ADRN is funded by the Economic and Social Research Council (ESRC). See: www.adrn.ac.uk

SNR, the APS does not include an indicator of whether an individual was dispersed during their asylum application. Therefore, the aim is to use information from the SNR about which characteristics are associated with dispersal to predict the probability that a respondent in the APS (whose main reason for migration was for international protection) was dispersed during their application.

This is achieved in a two-step process:

1. SNR data is used to model 'dispersal' as an outcome using a range of explanatory variables.
2. The model coefficients are then be applied to the same variables in the APS dataset, providing a predicted probability of dispersal for each individual.

Finally, the effect of dispersal on a range of 'outcomes' are analysed in each dataset and compared to assess robustness of the method as well as the implications for its possible usefulness. It is important to consider whether differences are an effect of dispersal, a result of methodology and error in predicting dispersal, or other factors e.g. APS being at a later time point. Matching has been widely used for assessing the effect of a 'treatment' on a population (Stuart, 2010); often in the context of an unknown outcome. In this instance it is the treatment group (those who have been dispersed) rather than the outcomes of interest which is unknown in the APS.

Before any analysis is carried out, individuals who were born in the UK or reported that their main reason for migration was *not* for international protection are removed from the APS dataset. Those who reported arriving before the dispersal policy was first implemented in 2000 are also removed.

Cross sectional analysis of the SNR baseline survey carried out in Chapter 5 is used to inform the choice of characteristics used in the initial SNR model of whether respondents were dispersed, although some changes are made as the priority in this instance is for the model coefficients to be transferable to the APS data. Furthermore, in this analysis the accommodation indicator in the SNR which includes four categories[33] is recoded into a binary variable of dispersal (1=yes 0=no). Therefore, it will be possible to model probability of dispersal for individuals in the APS where this information is not recorded, using coefficients from the SNR model showing the relationship between the explanatory variables and this outcome.

[33] With friends, with family, NASS accommodation and other – see Chapter 5.

6.3.3 Borrowing information on aggregates

Combining these two datasets can allow information on the aggregate structure of the SNR data on those who were dispersed and not dispersed to be applied to the APS data, in order to estimate and compare dispersal status with outcomes observed in the APS. Of the two datasets, the APS has a rich set of variables which would be useful in analysis of dispersal but no dispersal indicator. The SNR shares some variables in common with the APS but also has a dispersal indicator. The aim is to 'augment' the first dataset with information about dispersal by utilising the 'overlap' to apply cell structures from the second dataset.

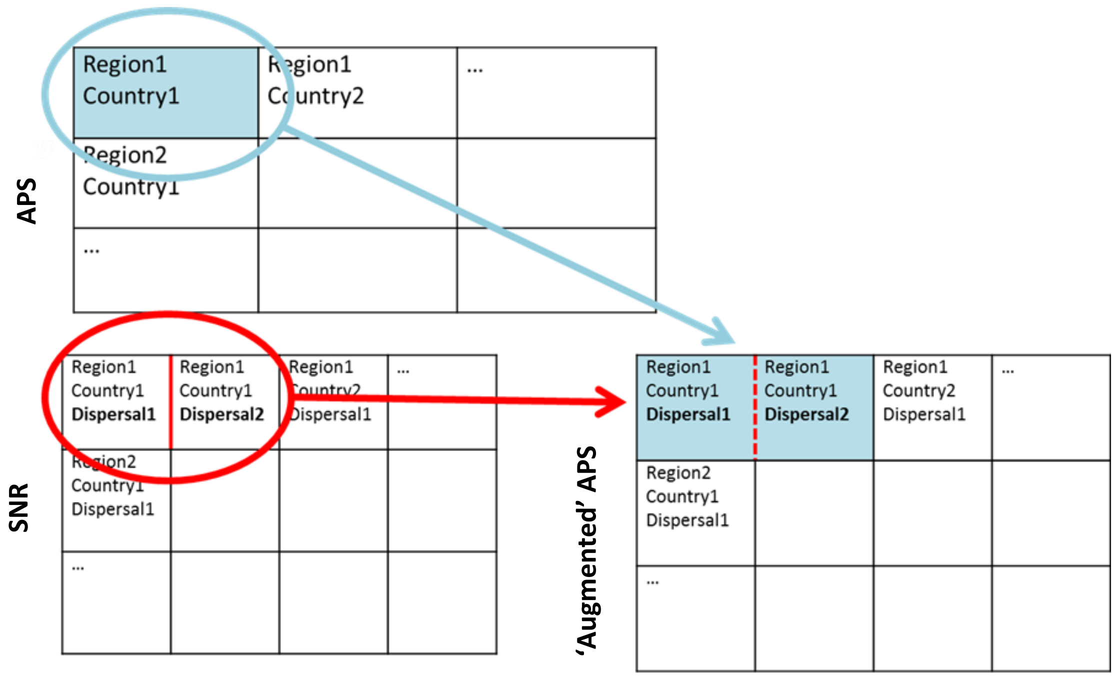
Statistical matching or 'data fusion' methods can be used here where the variable of interest is not jointly observed, but by combining sources in the context of certain assumptions, some conclusions may be drawn (Rassler, 2004). It must be assumed that the variables that are not shared in the two datasets are probabilistically independent conditionally on the shared variables (D'Orazio et al, 2006). The aim is to make some inference about the variables of interest from the fusion distribution constructed using statistical matching at the macro level (D'Orazio et al, 2006). That said, the fusion distribution should be seen as as a 'pseudo estimate of the target distribution', as it is not possible to empirically verify the underlying assumption, according to Zhang (2015, p. 784). Furthermore, uncertainty in statistical matching is inevitable; the lack of joint observations of the variable of interest means that identification uncertainty will exist, separate from sampling uncertainty, even with infinite observations (Zhang, 2015). D'Orazio et al (2006, p. 138) note that it is the fact that many distributions 'are compatible with the available partial information, … [which leads to] the so-called identification problem.' An awareness of the underlying assumptions and uncertainty are essential context for analysis which seeks to combine sources in this way.

Before the analysis is carried out, variables that are available in both datasets are identified and assessed for feasibility of harmonisation. Ensuring that variables can be made sufficiently similar in terms of categories is essential and limits the selection. For this analysis, region and country of origin are two variables that are found to be sufficiently similar in both datasets and have been found to be important for whether an individual is dispersed in previous analysis. The following methodology has been applied to the data in order to reach the presented results.

1. <u>Calculating rescaled counts (P_{rcd}):</u> SNR data is arranged to give cell counts (C_{rcd}) by region, country and dispersal indicator. The APS data is arranged by region and country, then the cell structure from the SNR is used to apportion each APS cell into 'dispersed' and 'not dispersed'. This gives predicted counts (P_{rcd}).

In Figure 6.1, blue shading represents the cell total from the APS and the red line represents how the total is split by the dispersal proportions observed in the SNR.

Figure 6.1: Diagram conceptualising rescaled counts (P_{rcd}).



Source: Author's own creation.

2. Calculating offset: An 'offset' can be used when combining data to impose the structure of one dataset on another, for defined variables (Raymer et al., 2007; Yildiz and Smith, 2015). This can be conceptualised as auxiliary information which will augment a dataset that lacks some information of interest; in this case adding information on the dispersal structure (in relation to region and country) to the APS, and therefore .

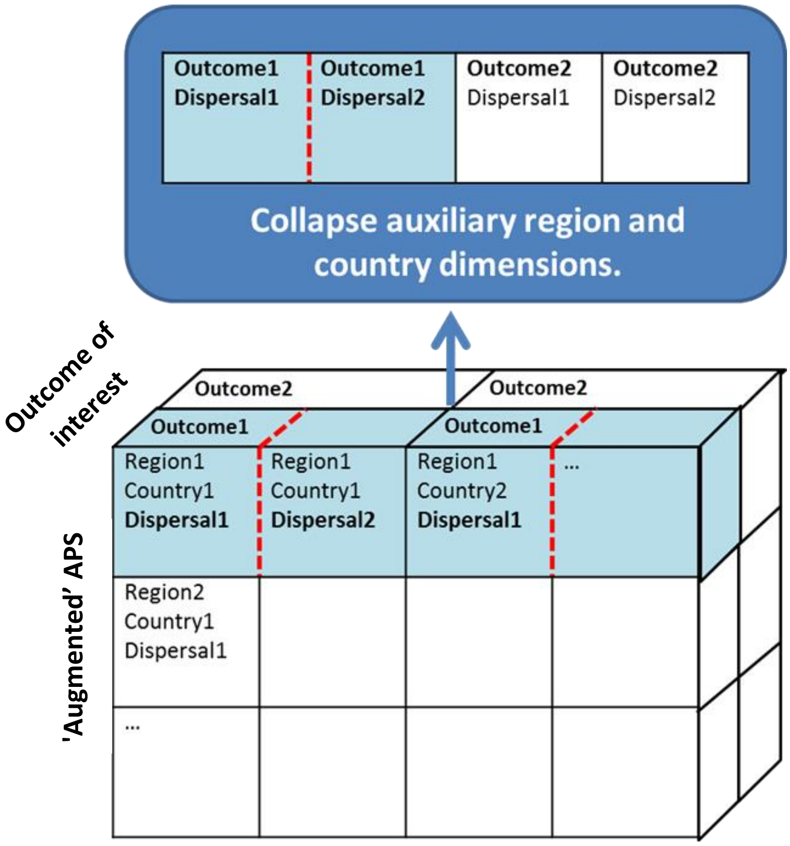
C_{rcd}= SNR data by region of residence, country of origin and dispersal status.

$$\ln(C_{rcd}) = offset$$

3. Poisson regression model for counts with offset: modelling predicted counts based on region, country of origin and offset gives predicted counts with a Poisson regression model (M_{rcd}).

Predicted counts (P_{rcd} or M_{rcd} for without or with the Poisson regression model respectively) are then apportioned to categories of APS variable of interest (economic activity, benefits, health, housing) for dispersed and not dispersed asylum seekers. Finally, the data is summed by region and country to give simple contingency tables of predicted dispersal by categories of the APS variable of interest, as illustrated in Figure 6.2.

Figure 6.2: Diagram conceptualising how outcome variable can be added.



Source: Author's own creation.

The Chi-square test of independence can be applied to these contingency tables to address the question of whether those who were dispersed have significantly different outcomes to those who were not dispersed. Therefore the null hypothesis is that there is no association between dispersal experience and the outcome variable.

6.3.4 Harmonisation

In order for both methods of 'borrowing' information to be carried out, it is important to harmonise variables across the SNR and APS; this will mean that the methods can be applied consistently and will allow direct comparison of the results. For most variables, the APS data had to be collapsed to coincide with the shorter list of categories in the published SNR dataset. See Appendix H for details.

The original APS data reports single year of which is the most useful format for analyses, but the SNR has age grouped into five categories[34] and the APS is therefore recoded to correspond with

[34] Note that these categories are not of equal size and for presentation 65+ is combined with 45-64 to avoid disclosure.

these. Age can be problematic as it inherently must change over time; however, the fact that ages are grouped and are also reported for a period means that it is not possible to ameliorate this. Age at arrival may also be important, but for the reasons described is also difficult to isolate in these data. Due to the fact that APS data are held within the UKDS Secure Lab, measures to limit risk of disclosure are required which in turn limits how the results are presented. These measures are vital not only to conform to data protection legislation but also to ensure that confidence in data holders and researchers is maintained (ADRN, 2016). Therefore, following the requirements for disclosure control, all figures which are lower than ten have to be removed; this may refer to an individual cell but there are also instances where results from a complete table (or variable within a table) which would be desirable to include in full can only be described in the text for data protection reasons.

Country of origin is reported as individual countries in the APS, whereas the SNR reports eleven countries and groups the remaining into continents (plus 'Middle East'). Some information is lost by harmonising APS data with the SNR categories; however, as described above, the priority for this analysis is to 'borrow' information from the SNR which is only possible if variables are consistent. Furthermore, the individual countries reported are those that report the highest number of refugees, and are therefore most important for this analysis. As described above, measures taken to minimise risk of disclosure mean that the country of origin variable is often not included in the results tables but instead patterns and findings are discussed in the text. Table 6.2 and Figure 6.4 show how year of arrival (reported as single years in the APS) can be grouped into five categories of length of residence for this analysis, to correspond with the SNR categories, shown in Figure 6.3[35].

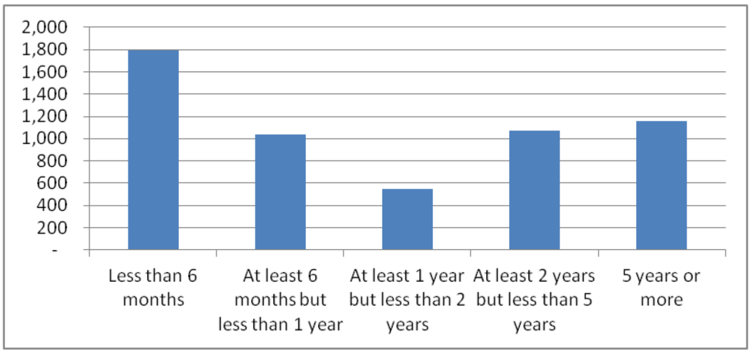
Table 6.2: Length of residence and closest corresponding time period, SNR and APS

SNR Length of residence	SNR arrival period	Approximate APS arrival period
Less than 6 months	1st June 2005 - 25th March 2007	2014
6 months to 1 year	1st Dec 2004 – 25th Sept 2006	2013
1 year to 2 years	1st Dec 2003 – 25th March 2006	2012
2 years to 5 years	1st Dec 2000 – 25th March 2005	2008-2011
5 years or more	Before 25th March 2002	2000-2007

Source: SNR and APS data (Home Office, 2010e; ONS, 2018).

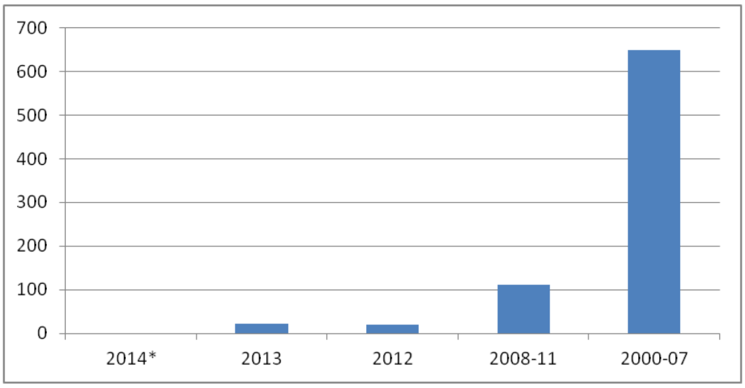
[35] Note that these categories are not of equal length.

Figure 6.3: Length of residence at time of decision, SNR



Source: Author's creation from SNR data (Home Office, 2010e).

Figure 6.4: Year of arrival of those seeking asylum aged 18+, APS



Source: Author's creation from APS data (ONS, 2018).
*low values removed due to disclosure control

It should be noted that length of residence in the SNR reports length of time before decision, whereas the APS does not record time of decision and may include individuals who received refugee status many years before the survey, or those who are still awaiting a decision.

Figure 6.4 shows that the majority of those arriving since 2000 whose main reason for migration was to seek asylum arrived between 2000 and 2007. This is partly due to it being the longest time period but also the fact that this was a time of high numbers arriving each year. Year of arrival periods do not match but the aim of this analysis is to borrow information on the asylum seeking population from one dataset and apply it to the second dataset, rather than to match the same individuals in two different sources.

Tables 6.3 and 6.4 show the distribution of frequencies and percentage of asylum seekers and refugees in the APS who have arrived since 2000 after variables have been recoded to harmonise with categories as reported in the SNR.

Table 6.3: Sex of those seeking asylum aged 18+ who arrived since 2000, APS

Sex	Frequency	Percent
Male	443	55.0
Female	362	45.0
Total	805	100

Source: APS data (ONS, 2018).

Table 6.4: Region of residence of those seeking asylum aged 18+ arriving since 2000, APS

Region (recode categories)	Frequency	Percent
London and South East	321	39.9
Midlands and East England	176	21.9
North East (NE), Yorks. and Humber	115	14.3
North West	108	13.4
Scotland and Northern Ireland (NI)	23	2.9
South West and Wales	62	7.7
Total	805	100

Source: APS data (ONS, 2018).

The APS variable reporting UK region of residence at the time of the survey is grouped to correspond with the SNR reported regions, broadly corresponding to Government Office Regions.

6.4 Results

6.4.1 Borrowing information on individuals: Predicting probability of dispersal

Data is modelled using the logit command in order to predict the probability of a 'positive' dispersal outcome based on age, sex, region of residence, country of origin and length of residence. These are all categorical variables and therefore dummy variables were created.

Model coefficients are then used to predict the probability of dispersal on the SNR dataset to test how good the model is for prediction. These are recoded as 1 for predicted probability greater than 0.5 and as 0 for values less than 0.5. Predicted outcome is subsequently cross-tabulated with known dispersal indicator in Table 6.5 as a test of how well the model predicts dispersal outcome for the SNR.

Table 6.5: Known and predicted outcome for model 1 (dispersed or not), SNR

	Known outcome		
Predicted outcome	No 0	Yes 1	**Total**
No 0	1,922	512	2,434 (48.1%)
Yes 1	416	2,207	2,623 (51.9%)
Total	2,338 (46.2%)	2,719 (53.8%)	5,057

Source: Author's analysis using SNR data (Home Office, 2010e).

With the first model the dispersal outcome is correctly predicted for 81.6 percent of respondents (highlighted green). Model coefficients were then used to predict the probability of dispersal on the APS dataset, effectively 'borrowing' information from the SNR (see Appendix I). These are then recoded 1 if predicted probability is 0.5 or more and 0 if it is less than 0.5.

Table 6.6 shows that with this model only 20 percent of respondents are predicted to have been dispersed, compared to 52 percent in the SNR, which is itself slightly lower than the known value (54 percent).

Table 6.6: Predicted dispersal outcome in the APS for three models based on SNR data.

	Counts			**Rounded percentages**		
Variable	**No**	**Yes**	**Total**	**No**	**Yes**	**Total**
Model 1	607	152	759	80%	20%	100%
Model 2	481	278	759	63%	37%	100%
Model3	475	284	759	63%	37%	100%

*Age groups 45-64 and 65+ combined for presentation due to disclosure control

Source: Author's analysis using SNR and APS data (Home Office, 2010e; ONS, 2018).

If we let y=/0 denote dispersed/not dispersed and x denote the cross-classified groups according to the selected covariates, the marginal distributions of y=1 in sample B and A are given by:

$f(y=1 \mid B) = sum_x\, f(y=1 \mid x, B\ f(x \mid B)$

$f(y=1 \mid A) = sum_x\, f(y=1 \mid x, B\ f(x \mid A)$

where f(y=1 | x, A) is replaced by f(y=1 |x, B) in sample matching estimation. Therefore it is clear that the difference in the marginal distributions f(x|B) and f(x|A) could contribute to the differences between the observed marginal distribution f)y=1|B) and the predicted f(y=1|A).[36]

One reason that the predicted percentage dispersed in the APS with this model is so much lower than in the SNR may be that it under-predicts the number dispersed in the earliest arrival group (2000-2007). As Table 6.2 shows, length of residence in the APS was recoded to correspond with length of residence as reported in the SNR rather than time period, as the range of dates in which the SNR was completed made this impractical and of limited analytical value. This allows the potential effect of length of residence in the UK to be accounted for in the model, but not any period effect. Therefore, if the rate of dispersal is related to the year in which an asylum seeker arrived, then it is not correct to 'borrow' this information from the SNR in this way and apply it to predict probability of dispersal in the APS.

There is greatest potential for problems with the earliest arrival group (2000-2007 in the APS) as the corresponding SNR group of longest residence (five years or more) refers to those arriving before the 25th of March 2002. The dispersal policy was implemented from 2000 which means that those arriving before this time could not have been dispersed. Therefore, there is a clear possibility that using this group to inform predicted probability of dispersal in the APS may be unhelpful. When the earliest arrival group (five years or more) is excluded, 60.2 percent of respondents in the SNR were dispersed. Table 6.2 shows that this remaining group corresponds more closely to the arrival period 2000-07 in the APS when period of arrival rather than length of residence is prioritised.

The pattern observed in the SNR is reflected in the APS model, however, the lack of correspondence of time period means that the individuals in the APS were not exposed to the rates of occurrence of arrival observed in the SNR. For these reasons it is useful to produce a second model, shown in Table 6.7, without length of residence to see if this improves how well the outcome is predicted.

[36] I would like to thank Li-Chun Zhang and Beata Nowok for their suggested notation here.

Table 6.7: Known and predicted outcome for model 2 (dispersed or not), SNR

	Known outcome		
Predicted outcome	No 0	Yes 1	**Total**
No 0	1,843	612	2,455 (48.1%)
Yes 1	432	2,215	2,647 (51.9%)
Total	2,275 (46.2%)	2,827 (53.8%)	5,102

Source: Author's analysis using SNR data (Home Office, 2010e).

The second model correctly predicts the dispersal outcome for 79.5 percent of respondents (highlighted green), which is slightly lower that in model 1. The results of this model (see Table 6.6), show that 36.6 percent of respondents are predicted to have been dispersed, compared to 52 percent in the SNR, and 20 percent with model 1.

Analysis of the SNR baseline survey in Chapter 5 found that age and sex were both significantly associated with dispersal outcome. However, when these variables are included in the logistic regression model with a binary outcome and a reduced number of explanatory variables in models one and two, they are not found to have a statistically significant effect at the five percent level (see Appendix I). Therefore a third model (Table 6.8) is tested to assess its ability to correctly predict dispersal outcome when age and sex are removed.

Table 6.8: Known and predicted outcome for model 3 (dispersed or not), SNR

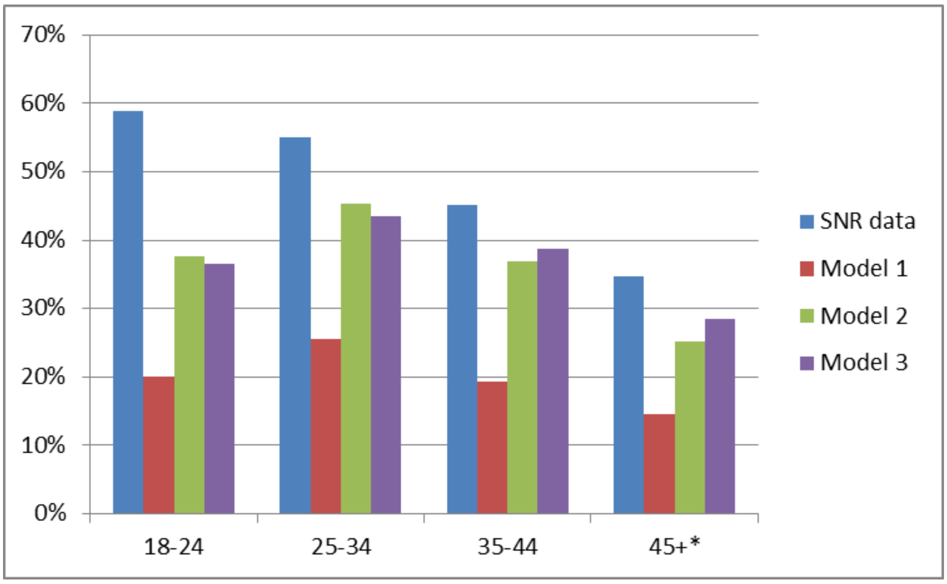
	Known outcome		
Predicted outcome	No 0	Yes 1	**Total**
No 0	1,865	626	2,491 (48.2%)
Yes 1	416	2,262	2,678 (51.8%)
Total	2,281 (44.1%)	2,888 (55.9%)	5,169

Source: Author's analysis using SNR data (Home Office, 2010e).

With model 3 the dispersal outcome is correctly predicted for 79.8 percent of respondents (highlighted green), which is slightly higher that with model 2. The results of this model (see Table 6.6), show that 37.4 percent of respondents are predicted to have been dispersed. This is lower than the levels reported in the SNR, but the APS population represents different cohorts arriving due to different drivers in a different economic and political context in the UK, but this analysis is based on small numbers.

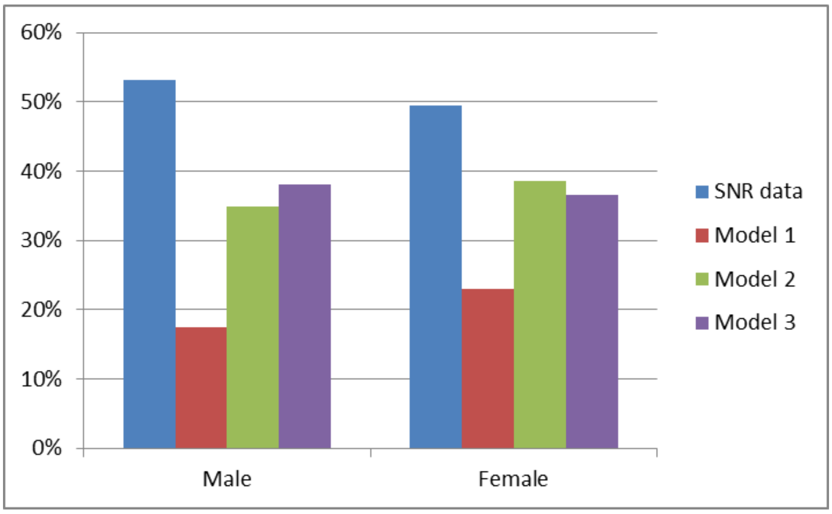
Figures 6.5 and 6.6 show the age and sex breakdown of dispersed refugees in the SNR as well as of those predicted to be dispersed with each model. Model 1 which includes length of shows very different age and sex distributions compared to the SNR. Once this is excluded, models 2 and 3 show patterns closer to those observed in the SNR and the variation could reasonable be explained by differences in the cohort covered.

Figure 6.5: Percentage dispersed by age in the SNR and predicted for the APS.



Source: Author's creation from analysis of SNR and APS data (Home Office, 2010e; ONS, 2018).

Figure 6.6: Percentage dispersed by sex in the SNR and predicted for the APS.



Source: Author's creation from analysis of SNR and APS data (Home Office, 2010e; ONS, 2018).

The models presented here consistently predicted dispersal status correctly for between 79 and 82 percent of respondents when tested on SNR data. Removing explanatory variables from models 1 and 2 only had a small effect on how well the model predicted dispersal; therefore, model three provides the simplest way of borrowing information on dispersal from the SNR and applying it to the APS. The model predicted lower levels of dispersal in the APS than were observed in the SNR.

Region and country of origin alone are able to correctly predict dispersal status 80 percent of the time in the SNR; this reflects the findings in Chapter 5 that showed these variables as being important, and can therefore be utilised to augment the APS at the aggregate level.

6.4.2 Borrowing information on aggregates

Where datasets share common variables, it is possible to utilise this overlap to borrow information on the data structure, predicting how the second dataset would look assuming these structures apply.

For this analysis we are interested in $f(z \mid y=1,A)$ where y=1/0 denotes whether an individual is dispersed/not in dataset A and z denotes the characteristic of interest, e.g. economic activity. Under the assumption of conditional independence given the relevant shared covariates x, i.e.

$$f(z \mid y=1, x, A) = f(z \mid x, A)$$

a ‘fusion distribution’ of interest in dataset A is given by

$$f(z \mid y=1,A) = \mathrm{sum}_x\, f(z,x \mid y=1,A) = \mathrm{sum}_x\, f(z \mid x,y=1,A)\, f(x \mid y=1,B) = \mathrm{sum}_x\, f(z \mid x,A)\, f(x \mid y=1,B)$$

wher the unobserved $f(x \mid y=1,A)$ is replaced by $f(x \mid y=1,B)$. This reflects the assumptions described in section 6.3.3 and furthermore shows how, by varying the ‘imputed’ $f(x \mid y=1,A)$, for example by using different sources or extreme scenarios, the associated identification uncertainty could be illustrated[37].

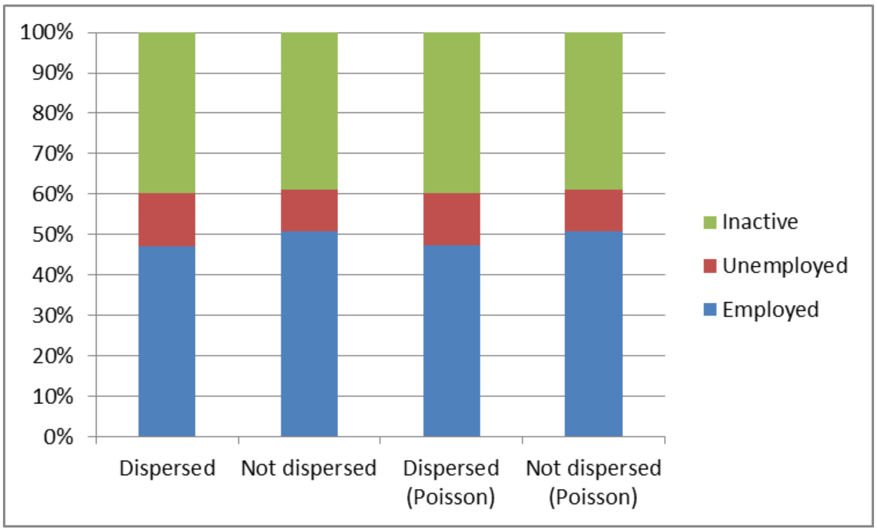
The results presented here show APS outcome variables by dispersal status, based on the patterns observed in the relationships between dispersal, region and country of origin in the SNR. Each outcome is presented twice, once showing results from restructuring and rescaling the data, and once after restructuring and rescaling the data using a Poisson regression model with offset. Full contingency tables and Chi square tests are presented in Appendix J.

Figures 6.7, 6.8 and 6.9 show slightly higher levels of benefit claims and poor health and lower levels of employment among dispersed refugees than those no dispersed. This reflects findings from analysis in Chapter 5 and results of analysis borrowing information on individuals in this

[37] I would like to thank Li-Chun Zhang and Beata Nowok for their suggested notation here.

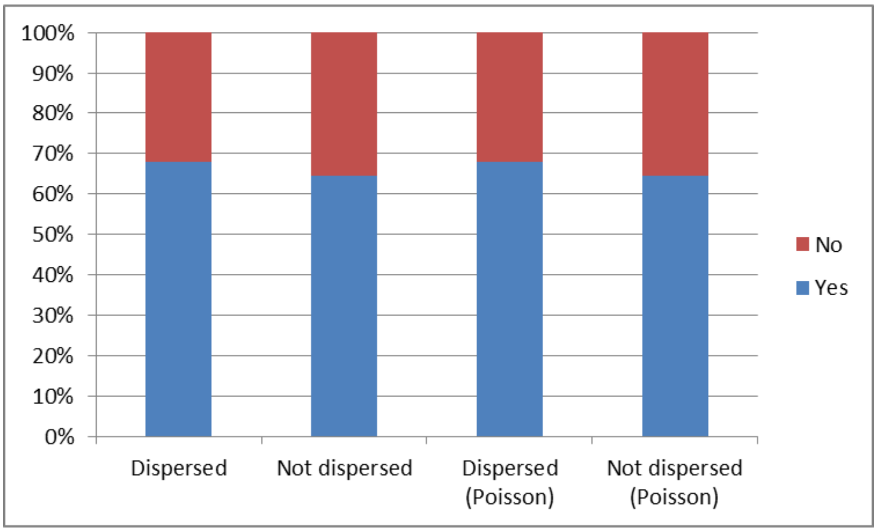
Chapter. That said, the observed difference is not as pronounced as might be expected, based on results from the preceding analyses.

Figure 6.7: Predicted percentage dispersed and not dispersed in the APS by economic activity, without and then with Poisson regression model.



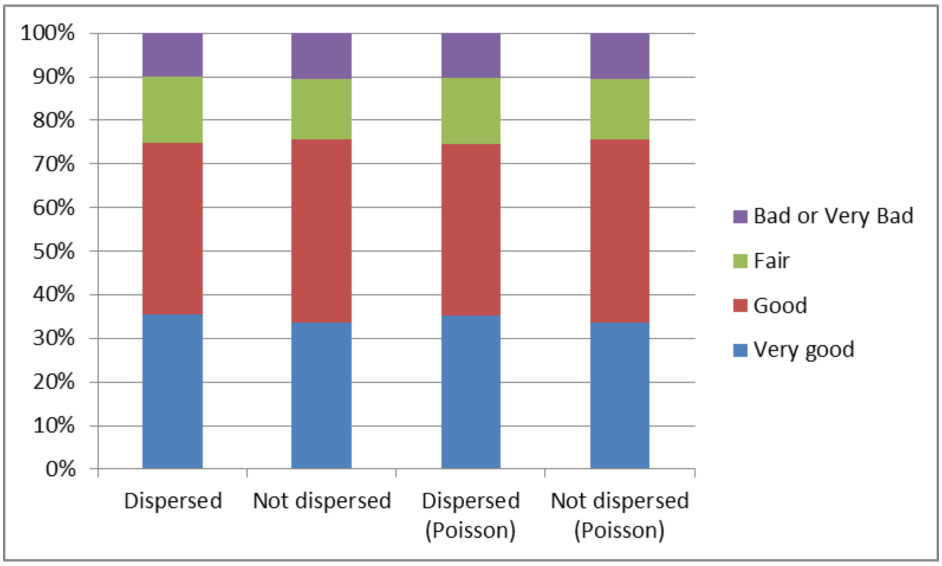
Source: Author's creation from analysis of SNR and APS data (Home Office, 2010e; ONS, 2018).

Figure 6.8: Predicted percentage dispersed and not dispersed in the APS by benefits claims, without and then with Poisson regression model.



Source: Author's creation from analysis of SNR and APS data (Home Office, 2010e; ONS, 2018).

Figure 6.9: Predicted percentage dispersed and not dispersed in the APS by health status, without and then with Poisson regression model.

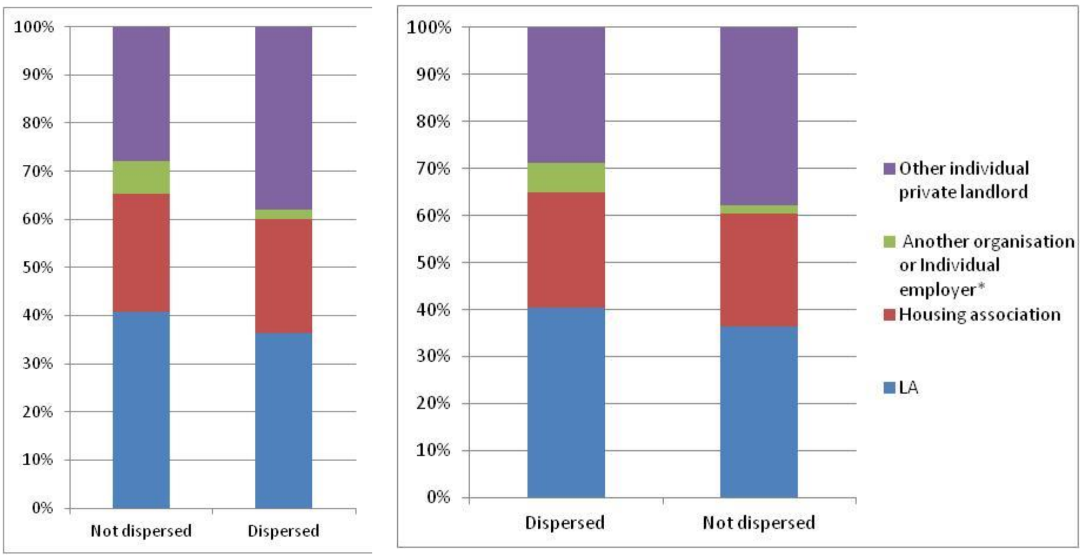


* 'Bad' and 'Very Bad' combined for presentation due to disclosure control requirements.

Source: Author's creation from analysis of SNR and APS data (Home Office, 2010e; ONS, 2018).

The majority of asylum seekers and refugees reported that they were renting and therefore this analysis presents the results of reported landlord. The higher proportion of LA housing show in Figure 6.10 among the dispersed population is feasible, based on knowledge of the policy which requires transition from dispersal accommodation on to LA support after receiving a decision.

Figure 6.10: Predicted percentage dispersed and not dispersed in the APS who are renting by housing landlord, without and then with Poisson regression model.



*Categories 'Another organisation' and 'Individual employer' combined for presentation due to disclosure control requirements.

Source: Author's creation from analysis of SNR and APS data (Home Office, 2010e; ONS, 2018).

The Chi square results in Appendix J show that, of the four outcome variables presented here, the only variable which is significantly associated with dispersal status at the one percent level is housing tenure. Economic activity, benefit claims and health status do not vary significantly by dispersal. Furthermore, in this analysis, the results with and without using the Poisson regression model with offset are remarkably similar. Although there is slight variation in the numbers estimated for individual cells, the difference is not enough to change the overall observed patterns.

These results present an illustration of how data may be combined to add an indicator of dispersal from the SNR to the APS, which does not report this information. The constraints of this analysis, including the use of only region and country of origin as the 'overlapping' variables, means that a more complex approach which includes a greater number of variables may produce different findings; here disclosure concerns restrict the possible analysis. It is also important to remember that those who were not dispersed are not necessarily a homogenous group; many may have been living with friends or family but some may have had the means to support themselves or have been housed under different schemes. Previous analysis of the SNR suggests that significant differences in outcome variables by dispersal status are observed.

6.5 Discussion and conclusions

The widely recognised issues with limited availability of data for the analysis of the asylum seeking and refugee population and dispersal in particular mean that it is essential to assess and recommend which additional methods, processes or data collection can provide the best options for researchers and policymakers to fill gaps in knowledge in this area. A form of 'cost-benefit' analysis is required in order to assess feasibility; the ideal scenario would be to recommend a course of action which would result in a massive information gain with minimal resources. The trade-off between the additional information gained by applying selected methods for combining data and the errors they introduce must also be considered. Finally, a discussion of the current options being pursued or considered, along with suggestions for how these opportunities can be maximised, offers practical suggestions for moving forward.

6.5.1 Combining datasets to explore dispersal

With so few sources of data available for the study of dispersal, in theory, combining datasets can open up a whole new range of variables to be analysed and topics to be explored (Lynn, 2009). If through combining datasets it is possible to in effect add an indicator of dispersal to existing large scale, regular and timely collections of data, without in fact commissioning new surveys or even questions within a survey, this could provide a considerable gain in information with minimal effort and additional cost, relatively speaking. It would improve our understanding in two key areas which have been highlighted as currently limited by the lack of sufficient data: to understand outcomes at different time points and to analyse additional variables to those available in the SNR.

Nevertheless, as with all quantitative analysis, the issues of access to and availability of relevant data must first be addressed. As described above, the ADRN has been created in order to facilitate access to linked datasets. A crucial element of any opening up of access to personal data, and of data held or collected by governments in particular, is the need to protect individuals' rights in the existing law. While the importance of consent and the appropriate legal framework was being considered at an early stage of discussions (BIS, 2013, p. 12), the issue remains complex and often confusing for all parties. In addition to establishing the availability of, and gaining access to, relevant datasets, it is also important to consider the content and coverage of sources in relation to the research questions to be answered. Appropriate population coverage, relevant time periods and key variables must be identified. Specific to this topic, it is clearly crucial that datasets include a satisfactory sample of asylum seekers and refugees as well as an indicator of dispersal.

The nature of all the methods to combine datasets described in this chapter means that variables must be consistent across multiple sources in order for any such processes to be carried out. Where possible variables can be harmonised, for example by regrouping categories as with age in the APS to match the SNR presented here. Nevertheless, the number of useable variables may limit the choice of methods for combining the datasets. The rich potential of combining datasets is widely acknowledged (Harron et al., 2017) but limitations in practice may mean that robust conclusions cannot be drawn; therefore a thorough analysis of the errors involved in the process of combining datasets is crucial to any work in this area.

6.5.2 Trade-offs between information gains and errors of estimation

The levels of uncertainty and error introduced during the process of combining datasets varies depending on the methods used. The review of selected methods presented here lays the groundwork by assessing their feasibility for contributing to the analysis of dispersal by weighing up the expected information gains against the error that they introduce.

The extent to which combining datasets can successfully expand our understanding of certain topics depends in the first instance on the quality and content of the original datasets. Herzog et al. (2007, p. 7) describe the five most important properties of data quality: relevance, accuracy, timeliness, comparability and completeness. When combining two or more datasets, the initial groundwork to ensure compatibility of populations covered including time periods and geography, is crucial. If the data is of poor quality then the ability to produce reliable findings and come to robust conclusions, and ultimately to suggest policy recommendations, is hampered. Survey data include particular challenges with regard to response rates and error introduced through survey design and data collection (Calderwood and Lessof, 2009). As described in Chapter 5, these issues can be magnified with attrition over multiple waves and, crucially, any bias existing in the original data will be carried through to the analysis of combined data.

When carrying out individual level linkage the particular challenges of balancing sensitivity and specificity is essential and well documented (Gomatam et al., 2002; Lynn, 2009), the choice being between ensuring that the highest possible number of matches are identified and ensuring that the fewest false positives are included. All that being said, individual level linkage along with a transparent and comprehensive assessment of uncertainty remains the best method for combining data and increasing the utility of existing datasets. Decisions made at this stage about what level of error to accept during data linkage depends on the intended use of the linked dataset. However, when two or more datasets are combined by ‘borrowing information’ on individuals (using model coefficients) or aggregates (utilising data structures), a different set of assumptions and possible sources of error and bias must be addressed.

Analysis of uncertainty when borrowing information on individuals using model coefficients is approached differently from individual linkage. The aim here is to produce results which show an estimation of what a population and its characteristics or outcomes may look like if the patterns observed in a second dataset are applied to it. In the case of dispersal, the probability that an individual was dispersed is predicted based on a selection of variables and is limited to the use of those variables which can be harmonised. Here, if we ignore issues with data collection, the uncertainty exists 1) within the initial model, 2) in the difficulty in identifying the same population and 3) in the way that probabilities are transformed into a binary indicator of dispersal.

Clearly, only the amount of information which is explained by the model can be ‘borrowed’ and applied to the second dataset. Furthermore, the choice of a cut-off for the predicted probability (here set at 0.5) to allocate individuals as dispersed rather than not dispersed is an arbitrary number. Those falling just above or below the threshold are most likely to be incorrectly categorised, and will be most susceptible to changing group if a different range of explanatory variables or cut-off is applied. That said, it is important to be explicit that the aim of this method is not directly to show replicability or to confirm robustness of findings from analysis of the SNR; for example, observing similar patterns in the APS for variables by dispersal as those observed in the SNR does not necessarily confirm the findings. What is more helpful with this method is to test the model on the original SNR dataset and compare predicted outcomes with known dispersal category. Frequently achieving a correct prediction level of around 80 percent is encouraging, particularly with a limited selection of explanatory variables.

A similar principle is relevant when 'borrowing information on aggregates'; the choice of variables (whether based on informed decisions or harmonisation limitations) limits the information to be gained from the cell structures and existing error will be carried across to the second dataset. With this method it is not possible to provide an explicit measure of how well dispersal is predicted which hampers the ability to make robust conclusions.

Individual level data linkage has the greatest potential in terms of considerable information gains along with the ability to measure and assess the errors involved, and therefore to some extent to mitigate those errors. However, issues with regard to access and subsequently considerations of data quality and compatibility mean that this is not always a viable option. The method presented to 'borrow information on individuals' allows an assessment of how well the model predicts dispersal but is limited to variables which can be harmonised across both datasets. Finally, 'borrowing information on aggregates' can offer a relatively simple way to investigate additional outcome variables, but the error and uncertainty involved is not measureable.

Possible further gains could be achieved through targeted data collection specifically for the purpose of combining datasets. For example, adding questions to existing surveys, such as the LFS or census, could allow identification of refugees or provide additional variables that overlap to expand the analysis possible as a result. Furthermore, targeted sample boosts for the population of interest may increase the sample and reduce uncertainty involved in combining datasets.

6.5.3 Policy and research implications

It is clear from the analysis presented here that the greatest scope for information gains with limited additional cost and introducing minimal error is possible through individual level data linkage. Large, high quality datasets such as administrative databases and government surveys that are linked using unique identifiers or a selection of personal information are likely to result in a linked dataset with a reasonable and measurable level of uncertainty, particularly for hard-to-reach groups (Harron et al., 2017; Lynn, 2009). If for example an indicator of dispersal from the SNR could be successfully added to the APS dataset, then this would provide timely and regular information on the policy. The issue of consent means that the linked dataset proposed for this project is not feasible. However, in the UK there is explicit provision in the Data Protection Act (DPA) for the right to confidentiality of personal information to be balanced against 'an overriding public interest in disclosure' (ADRN, 2016, p. 7). Although what constitutes public interest is not defined, it may be hoped that as the field of data linkage continues to grow and case-by-case decisions build precedents the existing frameworks to facilitate linkage can fulfil their intended purpose. An alternative method which can effectively address issues with consent and disclosure is to create synthetic datasets for analysis (Raab et al., 2016). Nevertheless, with the political will and a commitment to make the most of the large investments of time and resources which have already been made, individual level data linkage that can greatly increase our knowledge on the asylum seeking and refugee populations, and dispersal in particular, should be achievable.

In the meantime, the feasibility assessment of other methods for combining data has shown that a creative and rigorous approach to the analysis of this population and policy area has the potential to provide additional insights and add to the existing knowledge, which currently tends to be local level, qualitative and anecdotal (for example Hynes, 2011; Phillimore et al., 2004; Goodson et al., 2005).

While this research has focussed on how to achieve realistic and feasible aims within the scope of existing data availability or frameworks, it is clearly appropriate to also make more ambitious suggestions for data gains where there is such an acute and enduring lack of high quality data and the questions that need answering continue to be of such social and political concern. The simplest way to achieve the largest information gain is through additional data collection but it is important to consider how this could be carried out with the lowest demand on additional resources. Initially, keeping with the Government intention of 'using and re-using' existing data, alongside the 'migration mainstreaming' and harmonisation agenda championed by Eurostat (Knauth, 2011; Nowok and Willekens, 2011), means that continually re-assessing whether increasing publication of existing data held by government could help, either for analysis in isolation or by linking with other data. For example, the expansion of the regular releases of Home Office Asylum Statistics to include more information on key characteristics of asylum seekers such as age, sex and nationality as well as a greater temporal breakdown is a positive step (Bijak et al., 2013). A further improvement would be to publish additional detail on asylum seekers by support status (i.e. dispersed or subsistence only); reporting this data by LA and nationality would increase the potential for analysis of dispersal and deprivation where understanding geographic variation is so important.

The possibility of including an additional census question on reason for migration has been explored as part of the ONS Census Transformation Programme (ONS, 2016). One major benefit of including a question on 'reason for migration' in the census would be the ability to identify asylum seekers and refugees, allowing more informative individual level analysis of this group, as long as a sufficient response rate was achieved. A better understanding of the refugee population in particular, with the notable geographic detail which the census reports, is essential to the effective allocation of resources and implementation of policy in an efficient, fair and transparent way. Using data on socio-economic characteristics and outcomes of this group alongside data on year of arrival would also enable researchers to identify cohorts or 'waves' of refugee arrivals and analyse how they have fared over time. This would not directly provide additional data on dispersal but with data linkage methods the information provided by this question could in practice prove to have considerable potential for further research on this topic.

Currently, census data can only be linked to other data such as the SNR using proxy measures for refugee status (such as nationality/country of birth and year of arrival) or with methods to borrow information as illustrated above, but this additional question would considerably improve the ability to link data effectively for this population. It should be noted that this question would not be asking *current* status and therefore could not distinguish between refugees (or those with different permissions to stay) and asylum seekers; this would affect measures such as employment rates because of differences in right to work. Nevertheless, linking data on refugees (identified with this question) from the census to other datasets which have an indicator of whether an individual was dispersed as an asylum seeker, would allow the assessment of the dispersal policy and its impact much more effectively.

Chapter 7 Final conclusion and recommendations

This research has addressed the important and pressing questions around the extent to which dispersal has achieved its aim of 'spreading the burden' of asylum seekers and taking pressure off LAs in and around London. Furthermore, it has explored how settlement patterns compare to levels of deprivation across the country and considered the effect of dispersal on outcomes for refugees. Additional analysis has illustrated the feasibility of methods to combine available data in order to maximise its potential. This chapter will summarise the main findings of this work, highlighting how these contribute to a better understanding of dispersal and can feed into policy development moving forward.

Following the review of existing literature on the topic of dispersal, it is clear that qualitative research and anecdotal evidence have led to a general assumption that the locations of dispersal areas largely correspond with areas of high deprivation (Anie et al, 2005; Phillimore and Goodson, 2006; Hynes, 2011). Nevertheless, there has been a distinct lack of quantitative studies that can provide a national picture based on rigorous demographic methods. The results presented in Chapter 4 describe the variation in patterns of dispersal, subsistence only support and deprivation across LAs, presenting formal assessment of the inequality observed.

Initial findings confirm that settlement locations of dispersed asylum seekers are different from those on subsistence-only support, and reflect the policy aim to move settlement away from London. However, the evidence instead shows that high concentrations of asylum seekers are being dispersed to a smaller number of LAs (than those who are not dispersed), often to locations with high levels of deprivation. In contrast, it has been suggested that the recent settlement of refugees as part of the Syrian Vulnerable Persons Resettlement Scheme, initiated in 2014 and increased from the end of 2015, is achieving a broader geographic spread (HC Home Affairs Committee, 2016; 2017); the majority of LAs that have volunteered for the VPRP do not have any existing dispersal accommodation (HC Home Affairs Committee, 2017). The Scheme resettles refugees identified as particularly vulnerable directly from camps within the region and provides additional funding for services in participating LAs. According to Glasgow City Council, Syrians are getting a 'gold standard of service from all the detailed pre-planning to the arrival and the ongoing support to assist integration,' which is not available to those applying through the standard route (HC Home Affairs Committee, 2017, para. 42).

Another notable gap identified in the current understanding of dispersal was the lack of analysis of refugees' outcomes taking into account individuals' experiences of dispersal. This is addressed in Chapter 5 through analysis of the Survey of New Refugees. As the SNR is the only dataset which includes an indicator of dispersal and data that follows individuals over time, making full use of this source can provide information unavailable anywhere else. The main findings which show the relationship between dispersal status and receipt of benefits, as well as levels of employment, fit with what we know about how the system works; those living in dispersal accommodation at the time of decision then move into LA housing and continue to receive support within the deprived areas that they are located. Whether these patterns persist beyond the 21 months covered by the SNR cannot be known from existing sources. A thorough investigation of the considerable levels of attrition shows that those dispersed are more likely to drop out, and uncertainty around results of longitudinal analysis will therefore inevitably increase, in spite of attempts to address bias through weighting.

The House of Commons Home Affairs Committee (2017, para. 116) suggests that the standard of support provided under the Syrian scheme should be made available to all refugees as it offers 'prospective benefits of a reduction in overall costs through reduced reliance on welfare and other support services'. Additional funding for education and healthcare in settlement areas, for example, will directly impact on refugees' wellbeing and outcomes. Nevertheless, the Home Office 'does not accept that the support given to successful asylum seekers should be the same as the refugees brought to the UK under the Vulnerable Persons Resettlement Scheme' (HM Government, 2017). Deterring migrants from travelling to the UK to claim asylum remains a central topic of the Government's discourse and context for policy development. Increasingly settling refugees directly from camps is intended to discourage potential asylum applicants from travelling to the UK (idem). While this may be the Government's aim, there will be implications for co-operation and shared management of routes of arrival as a result of leaving the EU. Border controls in Northern France[38] are key to managing the arrival of asylum seekers and although commitment to this has been recently reiterated[39], future collaboration remains uncertain.

If the Syrian VPRP scheme is to be confidently recommended as a model to replace COMPASS for all refugee settlement in 2019, then additional research is urgently needed. The numbers currently settled under the new scheme are much smaller than those supported through dispersal and subsistence-only support. Additional analyses of this recent settlement, and the relevant data to enable this, are needed to show whether it is more successful. In particular, it is crucial to

---

[38] See Treaty of Le Touquet between the Government of theUnited Kingdom of Great Britain and Northern Ireland and the Government of the French Republic concerning the Implementation of Frontier Controls at the Sea Ports of both Countries on the Channel and North Sea, 4 February 2003: http://treaties.fco.gov.uk/docs/pdf/2004/TS0018.pdf.

[39] See for example a joint statement by the governments of France and the UK in November 2017: https://www.gov.uk/government/news/joint-statement-by-the-governments-of-france-and-the-uk.

obtain evidence for whether the Home Affairs Committee's (2017) suggestion that targeted funding of local services, along with a greater number of LAs voluntarily involved, results in a more sustainable and efficient system of support.

The final and arguably most pressing issue identified in this field of research is the existence of enduring knowledge gaps as a result of limited available data. For example, there remains very little understanding of more recent arrivals, their experience as asylum seekers and how this relates to subsequent outcomes as refugees; the SNR has not been repeated and a proposed 'Survey of Migrants,' which was explored in depth, has not materialised (Smith et al., 2011). A large-scale survey of migrants that includes particular targeting of asylum seekers and questions on experience of the immigration system, as well as attitudes and outcomes, would be expensive, but would clearly help to fill a number of evidence gaps that cannot be addressed with currently available data.

Instead, the current focus appears to be on data linkage and the desire to extract the full potential from existing sources (HM Government, 2012; BIS, 2013). This research has contributed a systematic review of the feasibility of combining datasets on the refugee and asylum seeking population in order to further our understanding of dispersal and settlement processes. The results presented in Chapter 6 show that augmenting existing datasets by adding an indicator of dispersal has the potential to greatly increase the number of variables, and therefore the topics, available for analysis. The main limitation identified here is the ability to link datasets in such a way that the error introduced is not only measurable, but also not so large as to undermine the value of any findings. A further limitation to data linkage more generally is the need to address issues around informed consent and ethics (ADRN, 2016).

While the collection of new datasets currently looks unlikely, the potential to add additional questions on migration (and more specifically refugee settlement) can be a cost-effective way to add to the knowledge base on this topic. It was the Eurostat agenda of 'mainstreaming migration statistics' (Knauth, 2011) that prompted the inclusion of a reason for migration question in the LFS, and this principle can be applied to other regularly collected datasets in the UK. A two-pronged approach which adds to the available data, for example through an additional census question, and facilitates data linkage by addressing the existing barriers described above, would create a better environment for research to considerably improve our understanding of the asylum and refugee populations and processes. This is a realistic proposition in the context of the current schemes and consultations which show that the political will exists, and the opportunity should be seized by those who recognise the importance of creating evidence-based policy.

Understanding the successes and failures of the current dispersal system is now extremely pertinent as the government considers what will follow the ending of COMPASS contracts in 2019. Accessing adequate data and applying rigorous and innovative methods can provide policymakers with the evidence required to deliver 'a sustainable, efficient and high quality end-to-end asylum accommodation and support system which works for all parties, and which effectively safeguards the vulnerable' (HM Government, 2017). This aim can be achieved by learning from existing sources as well as continued analysis to assess the impacts and outcomes of asylum and refugee settlement policy moving forward.

Appendix A– Dispersal in Northern Ireland, Scotland and Wales

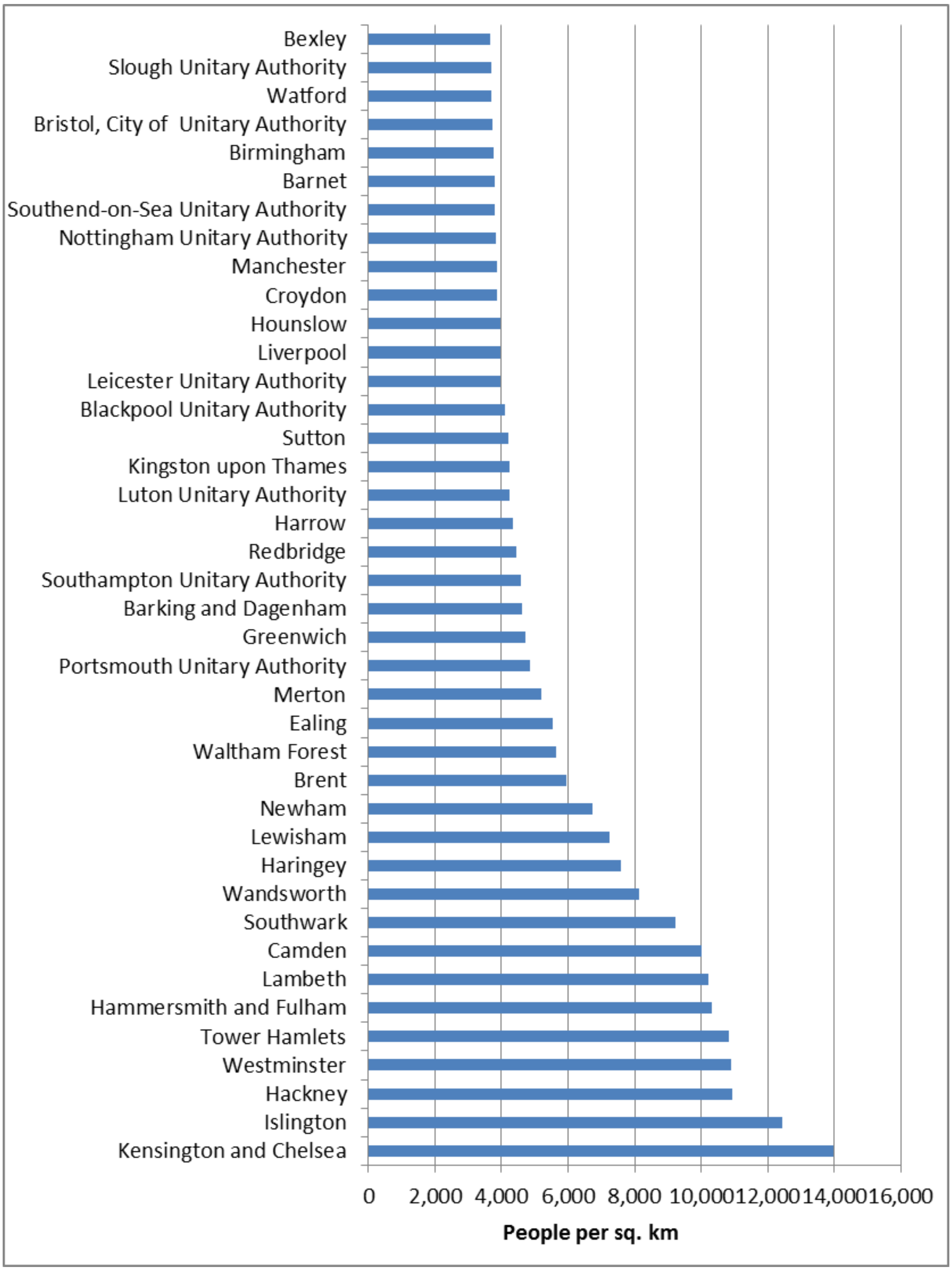
Table 7.1: Local Authorities in Northern Ireland, Scotland and Wales with dispersed asylum seekers in 2005 quarter two and 2008 quarter two.

Quarter	Local Authority	Dispersed
2005Q2	**Total Northern Ireland**	**119**
	Belfast	114
	Coleraine	2
	Newry and Mourne	2
	Newtownabbey	1
	Total Scotland	**5641**
	Glasgow	5641
	Total Wales	**2284**
	Cardiff	980
	Newport	368
	Swansea	893
	Wrexham	43
2008Q2	**Total Northern Ireland**	**231**
	Belfast	208
	Newtownabbey	20
	Lisburn	3
	Total Scotland	**2713**
	Edinburgh	2
	Glasgow	2707
	North Lanarkshire	2
	South Lanarkshire	2
	Total Wales	**1563**
	Cardiff	904
	Newport	258
	Swansea	386
	Wrexham	15

Source: Home Office Asylum Statistics (Home Office, 2015b).

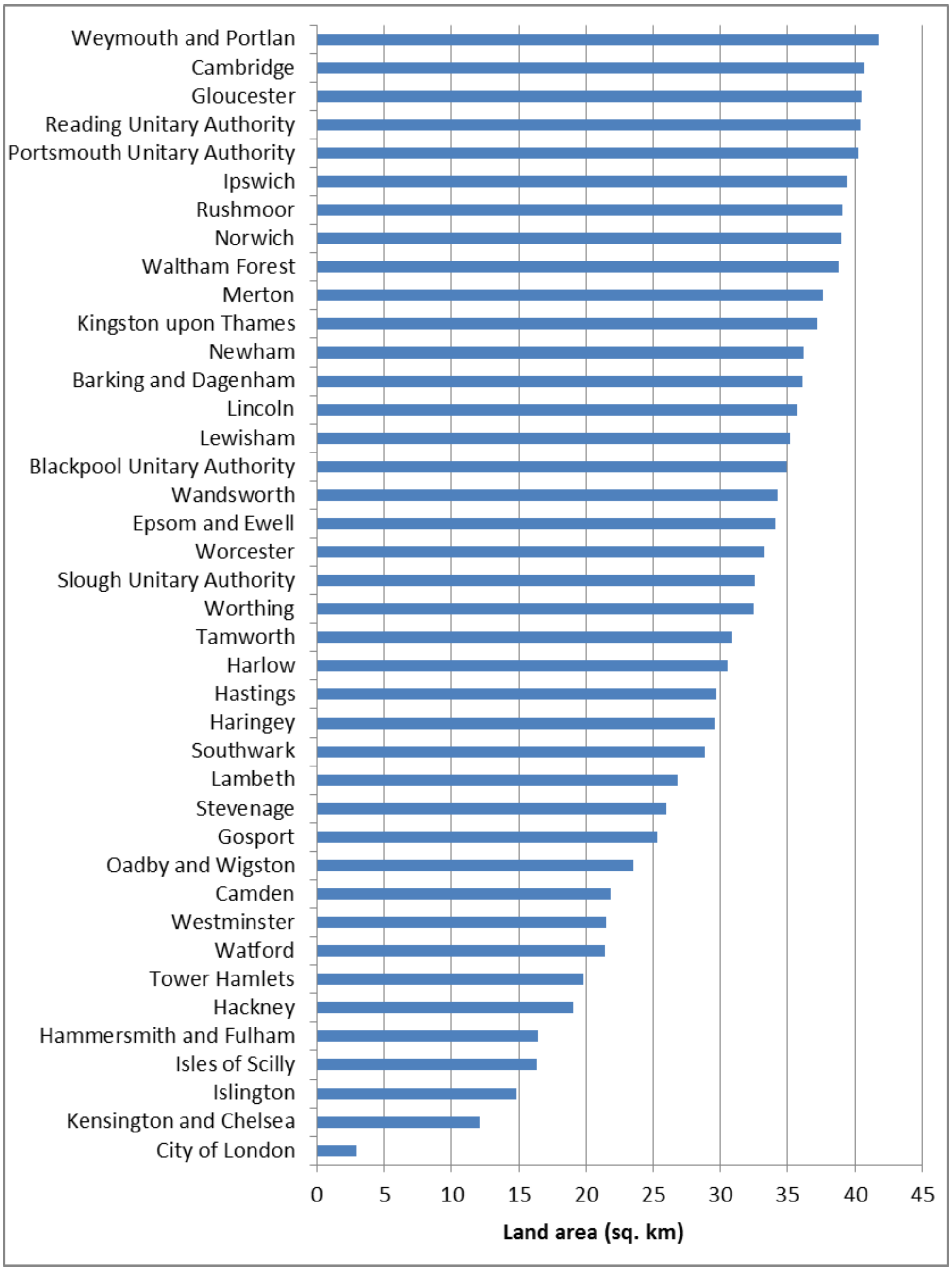
Appendix B – Local Authority statistics

Figure 7.1: Local Authorities with highest population density in 2005.



Source: Author's creation using ONS (2010) Population Density Estimates data.

Figure 7.2: Local Authorities with lowest Land area (based on 2001 Local Authority boundaries).



Source: Author's creation using ONS (2010) Population Density Estimates data.

Appendix C - Boundary data

Geographic data on LA boundaries is required for mapping. This is provided through the UK Data Service[40] in two data files: Unitary Authorities and District Authorities. The boundary data relates to 2001, before the changes which took place in 2009 when Unitary authorities were created in Cornwall, Durham, Northumberland, Shropshire and Wiltshire; Cheshire was split into two new unitary authorities, 'Cheshire East' and 'Cheshire West and Chester'; Bedfordshire was split into two new unitary authorities, 'Bedford Borough' and 'Central Bedfordshire'[41].

Table 7.2: Local Authority changes in April 2009.

Local Authorities before April 1st 2009	New Unitary Authorities from April 1st 2009
Alnwick **Berwick-upon-Tweed** **Blyth Valley** **Castle Morpeth** **Tynedale** **Wansbeck**	Northumberland
Bridgnorth **North Shropshire** **Oswestry** **Shrewsbury and Atcham** **South Shropshire**	Shropshire
Caradon **Carrick** **Kerrier** **North Cornwall** **Penwith** **Restormel**	Cornwall
Chester **Ellesmere Port and Neston** **Vale Royal**	Cheshire West and Chester
Chester-le-Street **Derwentside** **Durham** **Easington** **Sedgefield** **Teesdale** **Wear Valley**	County Durham
Congleton **Crewe and Nantwich** **Macclesfield**	Cheshire East
Kennet **North Wiltshire** **Salisbury** **West Wiltshire**	Wiltshire
Mid Bedfordshire **South Bedfordshire**	Central Bedfordshire Local
Bedford	Bedford

[40] See: http://census.edina.ac.uk/
[41] See: http://www.ons.gov.uk/ons/guide-method/geography/beginner-s-guide/administrative/our-changing-geography/local-government-restructuring/index.html.

Appendix C

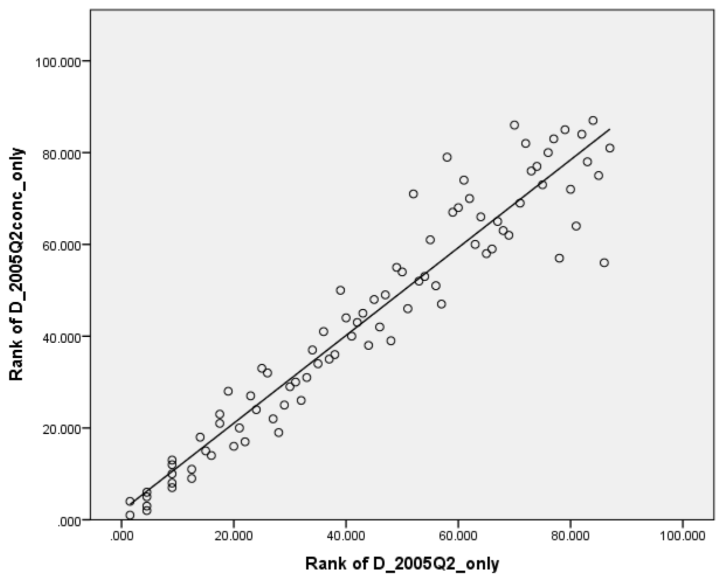
Appendix D– Rank of Local Authorities by support status

Table 7.3: Rank of top twenty Local Authorities by number and proportion of asylum seekers dispersed, 2005 Q2.

Rank	LAs by highest number dispersed		LAs by highest proportion dispersed (per 1000 population)	
1	Leeds	1976	Newcastle upon Tyne	4.45
2	Birmingham	1605	Middlesbrough Unitary Authority	4.35
3	Sheffield	1215	Salford	3.90
4	Newcastle upon Tyne	1214	Nottingham Unitary Authority	3.59
5	Manchester	1107	Wolverhampton	3.33
6	Nottingham Unitary Authority	1029	Rotherham	2.64
7	Liverpool	887	Leeds	2.63
8	Kirklees	885	Doncaster	2.54
9	Salford	855	Blackburn with Darwen Unitary Authority	2.51
10	Bradford	846	Manchester	2.48
11	Wolverhampton	795	Leicester Unitary Authority	2.38
12	Doncaster	734	Sandwell	2.34
13	Dudley	701	Sheffield	2.31
14	Leicester	694	Bury	2.30
15	Sandwell	673	Dudley	2.30
16	Rotherham	666	Kirklees	2.24
17	Coventry	651	Redcar and Cleveland	2.20
18	Middlesbrough	608	Rochdale	2.17
19	Wigan	601	Coventry	2.15
20	Bolton	523	Portsmouth	2.10

Source: Author’s creation using Home Office Asylum Statistics (Home Office, 2011).

Figure 7.3: Scatter graph of Local Authorities by rank of total number and rank of number dispersed per 1000 LA population in 2005, with linear regression line.



Appendix D

Source: Author's creation from analysis of Home Office Asylum Statistics (Home Office, 2011).

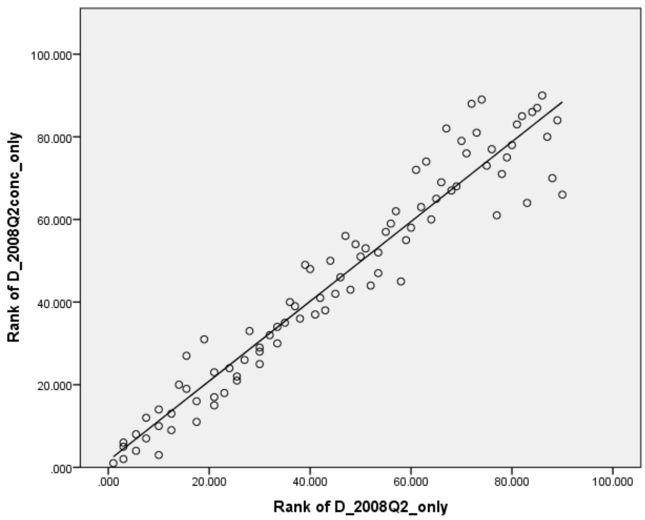
For ease of interpretation and comparison in Table 7.4 the LA with the lowest number or proportion of AS has been ranked 1 up to the LA with the highest being ranked N where N=total number of Las with at least one AS. Where there is a tied rank, each case is given a mean of the shared ranks.

Table 7.4: Rank of top twenty Local Authorities by number and proportion of asylum seekers dispersed, 2008 Q2.

Rank	LAs by highest number dispersed		LAs by highest proportion dispersed (per 1000 population)	
1	Birmingham	1163	Stoke-on-Trent Unitary Authority	3.56
2	Liverpool	1003	Middlesbrough Unitary Authority	3.08
3	Leeds	999	Blackburn with Darwen Unitary Authority	2.83
4	Manchester	913	Newcastle upon Tyne	2.82
5	Stoke-on-Trent Unitary Authority	852	Salford	2.81
6	Newcastle upon Tyne	784	Wolverhampton	2.38
7	Salford	627	Liverpool	2.27
8	Sheffield	590	Rotherham	2.21
9	Wolverhampton	567	Bury	1.99
10	Rotherham	560	Oldham	1.95
11	Nottingham Unitary Authority	552	Manchester	1.93
12	Coventry	531	Rochdale	1.87
13	Kirklees	519	Nottingham Unitary Authority	1.86
14	Bradford	490	Bolton	1.82
15	Bolton	480	Barnsley	1.71
16	Leicester	434	Coventry	1.71
17	Middlesbrough	432	Stockton-on-Tees	1.48
18	Oldham	425	Leicester	1.43
19	Blackburn with Darwen	395	Gateshead	1.30
20	Barnsley	386	Kirklees	1.29

Source: Author's creation using Home Office Asylum Statistics (Home Office, 2011).

Figure 7.4: Scatter graph of Local Authorities by rank of total number and rank of number dispersed per 1000 LA population in 2008, with linear regression line.



Source: Author's creation from analysis of Home Office Asylum Statistics (Home Office, 2011).

Table 7.5: Rank of top twenty Local Authorities by number and proportion of asylum seekers on 'subsistence only' support, 2005 Q2.

Rank	LAs by highest number on SO support		LAs by highest proportion supported on SO (per 1000 population)	
1	Newham	1166	Haringey	5.04
2	Haringey	1132	Newham	4.79
3	Ealing	848	Brent	2.91
4	Enfield	750	Hackney	2.89
5	Brent	749	Ealing	2.76
6	Hackney	603	Waltham Forest	2.71
7	Waltham Forest	592	Enfield	2.65
8	Lambeth	546	Hounslow	2.05
9	Birmingham	539	Southwark	2.00
10	Southwark	533	Lambeth	1.99
11	Redbridge	459	Redbridge	1.83
12	Hounslow	455	Islington	1.81
13	Lewisham	437	Barking and Dagenham	1.74
14	Croydon	415	Lewisham	1.72
15	Barnet	381	Harrow	1.65
16	Manchester	363	Slough	1.57
17	Harrow	360	Greenwich	1.25
18	Islington	335	Croydon	1.24
19	Wandsworth	333	Wandsworth	1.20
20	Barking and Dagenham	290	Camden	1.20

Source: Author's creation using Home Office Asylum Statistics (Home Office, 2011).

Table 7.6: Rank of top twenty Local Authorities by number and proportion of asylum seekers on 'subsistence only' support, 2008 Q2.

Rank	LAs by highest number on SO support		LAs by highest proportion supported on SO (per 1000 population)	
1	Newham	423	Newham	1.74
2	Ealing	304	Haringey	1.18
3	Haringey	266	Waltham Forest	1.14
4	Brent	258	Brent	1.01
5	Enfield	254	Ealing	0.97
6	Waltham Forest	253	Redbridge	0.91
7	Redbridge	240	Enfield	0.88
8	Barnet	176	Hackney	0.77
9	Lewisham	175	Lewisham	0.67
10	Birmingham	172	Hounslow	0.63
11	Lambeth	168	Merton	0.62
12	Manchester	167	Barking and Dagenham	0.61
13	Hackney	164	Lambeth	0.60
14	Southwark	154	Hillingdon	0.58
15	Hillingdon	150	Harrow	0.57
16	Hounslow	145	Greenwich	0.55
17	Harrow	129	Southwark	0.54
18	Croydon	128	Islington	0.53
19	Merton	126	Barnet	0.52
20	Greenwich	124	Camden	0.41

Appendix D

Source: Author’s creation using Home Office Asylum Statistics (Home Office, 2011).

Appendix E – Cluster Analysis results

Table 7.7: Thresholds for summary descriptions of cluster mean deprivation score and dispersal rate.

Summary description	Mean IMD Average Score	Mean dispersed AS per 1000 LA population
Very low	<17	-
Low	17-24	<0.8
Med	24-31	0.8-1.7
High	31-37	1.7-2.8
Very high	37+	2.8+

Appendix E

Table 7.8: Local Authorities by cluster (solution A), 2005.

Cluster 1
Barnet
Brighton and Hove Unitary Authority
Charnwood
Crewe and Nantwich
Croydon
Darlington Unitary Authority
Epping Forest
Gedling
Gloucester
Havering
Hounslow
Ipswich
Kensington and Chelsea
Merton
Newcastle-under-Lyme
Redbridge
Rushcliffe
Sefton
Solihull
South Gloucestershire Unitary Authority
Southampton Unitary Authority
Stockport
Swindon Unitary Authority
Trafford
Wandsworth
Cluster 2
Barnsley
Bolton
Bradford
Bristol, City of Unitary Authority
Bury
Calderdale
Coventry
Derby Unitary Authority
Doncaster
Dudley
Enfield
Gateshead
Hastings
Kirklees
Leeds
North East Lincolnshire Unitary Authority
North Tyneside
Norwich
Oldham
Peterborough Unitary Authority
Plymouth Unitary Authority
Portsmouth Unitary Authority
Redcar and Cleveland Unitary Authority
Rotherham
Sheffield
South Tyneside
Stockton-on-Tees Unitary Authority
Sunderland
Tameside
Wakefield
Walsall
Wigan
Cluster 3
Barking and Dagenham
Burnley
Camden
Hackney
Hammersmith and Fulham
Hartlepool Unitary Authority
Hyndburn
Islington
Lambeth
Lewisham
Newham
Pendle
Southwark
St. Helens
Waltham Forest
Cluster 4
Middlesbrough Unitary Authority
Newcastle upon Tyne
Nottingham Unitary Authority
Salford
Wolverhampton
Cluster 5
Birmingham
Blackburn with Darwen Unitary Authority
Haringey
Kingston upon Hull, City of Unitary Authority
Leicester Unitary Authority
Liverpool
Manchester
Rochdale
Sandwell
Stoke-on-Trent Unitary Authority

Table 7.9: Local Authority characteristics by cluster (solution A), 2005.

Cluster	Number of LAs	Mean (and range): IMD2007 Average Score	Mean (and range): dispersed AS per 1,000 LA population	Mean population density	Mean geographic size
1	25	18.79 (17.43)	0.22 (0.82)	2832	159
2	32	27.57 (10.79)	1.71 (2.04)	1920	201
3	15	34.09 (18.03)	0.11 (0.30)	6136	57
4	5	35.45 (7.58)	3.92 (1.12)	2906	82
5	10	38.16 (13.08)	2.09 (1.09)	3497	114
Total	87	27.84 (38.84)	1.18 (4.45)	3147	147

Source: Author’s creation from analysis of Home Office Asylum Statistics, IMD data and ONS Population Density Estimates (Home Office, 2011; DCLG, 2008; ONS, 2010; 2010a).

Appendix E

Table 7.10: Local Authorities by cluster (solution B), 2005.

Cluster 1
Barnet
Charnwood
Crewe and Nantwich
Croydon
Epping Forest
Gedling
Havering
Merton
Newcastle-under-Lyme
Redbridge
Rushcliffe
Solihull
South Gloucestershire Unitary Authority
Stockport
Swindon Unitary Authority
Trafford
Wandsworth
Cluster 2
Brighton and Hove Unitary Authority
Bristol, City of Unitary Authority
Calderdale
Darlington Unitary Authority
Enfield
Gloucester
Hastings
Hounslow
Ipswich
Kensington and Chelsea
North East Lincolnshire Unitary Authority
North Tyneside
Norwich
Oldham
Plymouth Unitary Authority
Sefton
South Tyneside
Southampton Unitary Authority
Stockton-on-Tees Unitary Authority
Sunderland
Tameside
Wakefield
Cluster 3
Barnsley
Bolton
Bradford
Bury
Coventry
Derby Unitary Authority
Doncaster
Dudley
Gateshead
Kirklees
Leeds
Peterborough Unitary Authority
Portsmouth Unitary Authority
Redcar and Cleveland Unitary Authority
Rotherham
Sheffield
Walsall
Wigan
Cluster 4
Barking and Dagenham
Burnley
Camden
Hackney
Hammersmith and Fulham
Hartlepool Unitary Authority
Hyndburn
Islington
Lambeth
Lewisham
Newham
Pendle
Southwark
St. Helens
Waltham Forest
Cluster 5
Birmingham
Blackburn with Darwen Unitary Authority
Haringey
Kingston upon Hull, City of Unitary Authority
Leicester Unitary Authority
Liverpool
Manchester
Middlesbrough Unitary Authority
Newcastle upon Tyne
Nottingham Unitary Authority
Rochdale
Salford
Sandwell
Stoke-on-Trent Unitary Authority
Wolverhampton

Table 7.11: Local Authority characteristics by cluster (solution B), 2005.

Cluster	Number of LAs	Mean (and range): IMD2007 Average Score	Mean (and range): dispersed AS per 1000 LA population	Mean population density	Mean geographic size
1	17	16.39 (13.18)	0.21 (0.82)	2184	196
2	22	26.41 (10.57)	0.80 (1.52)	2897	118
3	18	27.36 (10.58)	2.17 (0.90)	1744	248
4	15	34.09 (18.03)	0.11 (0.30)	6136	57
5	15	37.26 (15.61)	2.70 (3.03)	330	103
Total	87	27.84 (4.45)	1.18 (38.84)	3147	147

Source: Author's creation from analysis of Home Office Asylum Statistics, IMD data and ONS Population Density Estimates (Home Office, 2011; DCLG, 2008; ONS, 2010; 2010a).

Table 7.12: Local Authorities by cluster (solution C), 2008.

Cluster 1
Ashford
Barnet
Blaby
Colchester
Croydon
Dartford
Epping Forest
Gedling
Harrow
Havering
Hillingdon
Hounslow
Kensington and Chelsea
Macclesfield
Merton
Newcastle-under-Lyme
Oxford
Reading
Redbridge
Richmond upon Thames
Rushcliffe
South Gloucestershire
Stockport
Swindon
Trafford
Wandsworth
West Oxfordshire
Cluster 2
Bradford
Bristol
Calderdale
Darlington
Derby
Doncaster
Dudley
Ealing
Enfield
Gateshead
Gloucester
Ipswich
Kirklees
Leeds
Luton
North Tyneside
Norwich
Peterborough
Plymouth
Portsmouth
Sheffield
South Tyneside
Southampton
Stockton-on-Tees
Sunderland
Tameside
Wakefield
Wigan
Cluster 3
Barnsley
Bolton
Bury
Coventry
Leicester
Nottingham
Oldham
Rochdale
Rotherham
Cluster 4
Blackburn with Darwen
Middlesbrough
Newcastle upon Tyne
Salford
Stoke-on-Trent
Wolverhampton
Cluster 5
Barking and Dagenham
Birmingham
Brent
Greenwich
Hackney
Haringey
Hartlepool
Hastings
Islington
Kingston upon Hull
Lambeth
Lewisham
Newham
North East Lincolnshire
Sandwell
Southwark
Walsall
Waltham Forest
Cluster 6
Liverpool
Manchester

Table 7.13: Local Authority characteristics by cluster (solution C), 2008.

Cluster	Number of LAs	Mean (and range): IMD2010 Average Score	Mean (and range): dispersed AS per 1,000 LA population	Mean population density	Mean geographic size
1	27	16.28 (15.70)	0.12 (0.51)	2751	192
2	28	25.95 (11.71)	0.82 (1.43)	2398	185
3	9	30.01 (12.19)	1.84 (0.79)	2152	156
4	6	34.51 (7.88)	2.91 (1.18)	2392	94
5	18	34.53 (13.59)	0.35 (1.26)	5831	67
6	2	42.29 (2.32)	2.10 (0.34)	4018	112
Total	90	26.11 (35.84)	0.79 (3.56)	3201	153

Source: Author’s creation from analysis of Home Office Asylum Statistics, IMD data and ONS Population Density Estimates (Home Office, 2011; DCLG, 2011; ONS, 2010; 2010b).

Table 7.14: Local Authorities by cluster (solution D), 2008.

Cluster 1
Ashford
Barnet
Blaby
Bristol
Colchester
Croydon
Darlington
Dartford
Derby
Ealing
Enfield
Epping Forest
Gedling
Gloucester
Harrow
Havering
Hillingdon
Hounslow
Ipswich
Kensington and Chelsea
Luton
Macclesfield
Merton
Newcastle-under-Lyme
North Tyneside
Norwich
Oxford
Peterborough
Portsmouth
Reading
Redbridge
Richmond upon Thames
Rushcliffe
South Gloucestershire
Southampton
Stockport
Swindon
Trafford
Wandsworth
West Oxfordshire
Cluster 2
Barnsley
Bolton
Bradford
Bury
Calderdale
Coventry
Doncaster
Dudley
Gateshead
Kirklees
Leeds
Leicester
Nottingham
Oldham
Plymouth
Rochdale
Rotherham
Sheffield
South Tyneside
Stockton-on-Tees
Sunderland
Tameside
Wakefield
Wigan
Cluster 3
Barking and Dagenham
Birmingham
Brent
Greenwich
Hackney
Haringey
Hartlepool
Hastings
Islington
Kingston upon Hull
Lambeth
Lewisham
Newham
North East Lincolnshire
Sandwell
Southwark
Walsall
Waltham Forest
Cluster 4
Blackburn with Darwen
Liverpool
Manchester
Middlesbrough
Newcastle upon Tyne
Salford
Stoke-on-Trent
Wolverhampton

Table 7.15: Local Authority characteristics by cluster (solution D), 2008.

Cluster	Number of LAs	Mean (and range): IMD2010 Average Score	Mean (and range): dispersed AS per 1000 LA population	Mean population density	Mean geographic size
1	40	19.03 (18.45)	0.25 (0.82)	2937	160
2	24	28.13 (12.19)	1.37 (1.46)	1804	224
3	18	34.53 (13.59)	0.35 (1.26)	5831	67
4	8	36.46 (13.71)	2.71 (1.63)	2798	99
Total	90	26.11 (35.84)	0.79 (3.56)	3201	153

Source: Author's creation from analysis of Home Office Asylum Statistics, IMD data and ONS Population Density Estimates (Home Office, 2011; DCLG, 2011; ONS, 2010; 2010b).

Appendix F – SNR baseline model: SPSS output

Table 7.16: Cross-sectional baseline model of accommodation with ability to speak English, SNR

Parameter Estimates

Where are you currently living? [a]		B	Std. Error	Wald	df	Sig.	Exp(B)	95% Confidence Interval for Exp (B) Lower Bound	95% Confidence Interval for Exp (B) Upper Bound
NASS accommodation	Intercept	.564	.760	.549	1	.459			
	Q6_Childcount	.010	.068	.022	1	.882	1.010	.884	1.155
	[Q31_health=1]	-.844	.358	5.555	1	.018	.430	.213	.867
	[Q31_health=2]	-.465	.352	1.748	1	.186	.628	.315	1.251
	[Q31_health=3]	-.306	.352	.760	1	.383	.736	.370	1.466
	[Q31_health=4]	-.197	.378	.271	1	.602	.821	.392	1.723
	[Q31_health=5]	0[b]	.	.	0	.	.	.	.
	[Q21_contact_nationalgrps=1]	-.116	.238	.237	1	.627	.891	.559	1.419
	[Q21_contact_nationalgrps=2]	.323	.194	2.769	1	.096	1.381	.944	2.020
	[Q21_contact_nationalgrps=3]	.068	.179	.147	1	.701	1.071	.755	1.520
	[Q21_contact_nationalgrps=4]	-.057	.164	.122	1	.727	.944	.685	1.302
	[Q21_contact_nationalgrps=5]	0[b]	.	.	0	.	.	.	.
	[Q8_time_UK=1]	.844	.216	15.290	1	.000	2.327	1.524	3.552
	[Q8_time_UK=2]	.545	.212	6.628	1	.010	1.724	1.139	2.611
	[Q8_time_UK=3]	.391	.224	3.044	1	.081	1.478	.953	2.293
	[Q8_time_UK=4]	.447	.172	6.749	1	.009	1.563	1.116	2.190
	[Q8_time_UK=5]	0[b]	.	.	0	.	.	.	.
	[Q1_gender=1]	-.027	.141	.035	1	.851	.974	.738	1.285
	[Q1_gender=2]	0[b]	.	.	0	.	.	.	.
	[Q2_age_group=1]	1.092	.396	7.611	1	.006	2.980	1.372	6.474
	[Q2_age_group=2]	1.325	.387	11.741	1	.001	3.761	1.763	8.024
	[Q2_age_group=3]	1.460	.398	13.484	1	.000	4.308	1.976	9.393
	[Q2_age_group=4]	1.075	.408	6.954	1	.008	2.931	1.318	6.519
	[Q2_age_group=5]	0[b]	.	.	0	.	.	.	.
	[Q18_spk_English=1]	-.292	.265	1.215	1	.270	.747	.444	1.255
	[Q18_spk_English=2]	-.005	.217	.001	1	.982	.995	.650	1.522
	[Q18_spk_English=3]	-.065	.193	.114	1	.735	.937	.641	1.368
	[Q18_spk_English=4]	0[b]	.	.	0	.	.	.	.
	[Q30_meet_friends=1]	1.551	.356	18.932	1	.000	4.714	2.345	9.478
	[Q30_meet_friends=2]	1.303	.334	15.178	1	.000	3.679	1.910	7.084
	[Q30_meet_friends=3]	1.260	.326	14.980	1	.000	3.525	1.862	6.671
	[Q30_meet_friends=4]	1.157	.322	12.897	1	.000	3.180	1.691	5.978
	[Q30_meet_friends=5]	1.019	.337	9.169	1	.002	2.772	1.433	5.362
	[Q30_meet_friends=6]	0[b]	.	.	0	.	.	.	.
	[Friends_relatives_recoded=1]	-1.528	.368	17.272	1	.000	.217	.106	.446
	[Friends_relatives_recoded=2]	0[b]	.	.	0	.	.	.	.
	[Q24_need_help=1]	.509	.120	17.913	1	.000	1.663	1.314	2.105
	[Q24_need_help=2]	0[b]	.	.	0	.	.	.	.
	[Q9_origin=1]	-.141	.423	.111	1	.739	.869	.379	1.989
	[Q9_origin=2]	.913	.390	5.497	1	.019	2.493	1.162	5.349
	[Q9_origin=3]	-.723	.784	.850	1	.356	.485	.104	2.257
	[Q9_origin=4]	1.730	.474	13.348	1	.000	5.642	2.230	14.275
	[Q9_origin=5]	3.007	.480	39.194	1	.000	20.234	7.892	51.877
	[Q9_origin=6]	3.058	.780	15.361	1	.000	21.295	4.614	98.289
	[Q9_origin=7]	-.002	.363	.000	1	.995	.998	.489	2.034
	[Q9_origin=8]	1.599	.516	9.610	1	.002	4.946	1.800	13.591
	[Q9_origin=9]	-.328	.391	.704	1	.401	.720	.335	1.550
	[Q9_origin=10]	1.229	.383	10.283	1	.001	3.419	1.613	7.247
	[Q9_origin=11]	.965	.380	6.432	1	.011	2.623	1.245	5.528
	[Q9_origin=12]	.852	.375	5.156	1	.023	2.344	1.124	4.889
	[Q9_origin=13]	1.287	.441	8.508	1	.004	3.621	1.525	8.597
	[Q9_origin=14]	.372	.438	.722	1	.395	1.451	.615	3.425
	[Q9_origin=15]	.527	.453	1.355	1	.244	1.694	.697	4.115
	[Q9_origin=16]	0[b]	.	.	0	.	.	.	.
	[Region=1]	-3.027	.293	106.806	1	.000	.048	.027	.086
	[Region=2]	-1.066	.300	12.629	1	.000	.344	.191	.620
	[Region=3]	.490	.326	2.252	1	.133	1.632	.861	3.093
	[Region=4]	-.053	.335	.025	1	.874	.948	.492	1.829
	[Region=5]	.188	.437	.184	1	.668	1.206	.512	2.843
	[Region=6]	0[b]	.	.	0	.	.	.	.
	[Q4_partner=1]	-1.115	.154	52.578	1	.000	.328	.243	.443
	[Q4_partner=2]	0[b]	.	.	0	.	.	.	.
	[Q10_religion=1]	-.062	.403	.023	1	.878	.940	.427	2.070
	[Q10_religion=2]	.423	.359	1.390	1	.238	1.527	.755	3.087
	[Q10_religion=3]	1.305	.729	3.210	1	.073	3.689	.885	15.386
	[Q10_religion=4]	-.021	.615	.001	1	.973	.979	.293	3.266
	[Q10_religion=5]	.630	7774.726	.000	1	1.000	1.877	.000	.[c]
	[Q10_religion=6]	.425	.337	1.597	1	.206	1.530	.791	2.960
	[Q10_religion=7]	-.849	1.223	.482	1	.488	.428	.039	4.703
	[Q10_religion=8]	0[b]	.	.	0	.	.	.	.
	[Q28_meet_relatives=1]	-1.464	.261	31.521	1	.000	.231	.139	.386
	[Q28_meet_relatives=2]	-1.468	.234	39.274	1	.000	.230	.146	.365
	[Q28_meet_relatives=3]	-1.088	.203	28.852	1	.000	.337	.227	.501
	[Q28_meet_relatives=4]	-.800	.179	19.851	1	.000	.450	.316	.639
	[Q28_meet_relatives=5]	-.415	.228	3.319	1	.068	.660	.423	1.032
	[Q28_meet_relatives=6]	0[b]	.	.	0	.	.	.	.
	[Q14_Emp_recoded=1]	.396	.273	2.101	1	.147	1.485	.870	2.537
	[Q14_Emp_recoded=2]	.422	.277	2.327	1	.127	1.526	.887	2.625
	[Q14_Emp_recoded=3]	.052	.329	.025	1	.874	1.054	.553	2.007
	[Q14_Emp_recoded=4]	-.018	.289	.004	1	.950	.982	.557	1.730
	[Q14_Emp_recoded=5]	.140	.283	.244	1	.621	1.150	.660	2.005
	[Q14_Emp_recoded=7]	.747	.708	1.112	1	.292	2.110	.527	8.448
	[Q14_Emp_recoded=8]	0[b]	.	.	0	.	.	.	.
With friends	Intercept	-1.451	.772	3.532	1	.060			
	Q6_Childcount	-.488	.086	32.547	1	.000	.614	.519	.726
	[Q31_health=1]	.276	.373	.547	1	.460	1.317	.634	2.736
	[Q31_health=2]	.413	.367	1.268	1	.260	1.512	.736	3.105
	[Q31_health=3]	.395	.367	1.158	1	.282	1.484	.723	3.048
	[Q31_health=4]	-.170	.400	.182	1	.670	.843	.385	1.846
	[Q31_health=5]	0[b]	.	.	0	.	.	.	.
	[Q21_contact_nationalgrps=1]	.370	.230	2.593	1	.107	1.448	.923	2.271
	[Q21_contact_nationalgrps=2]	.410	.193	4.484	1	.034	1.506	1.031	2.201
	[Q21_contact_nationalgrps=3]	.349	.177	3.859	1	.049	1.417	1.001	2.006
	[Q21_contact_nationalgrps=4]	.440	.160	7.559	1	.006	1.553	1.135	2.125
	[Q21_contact_nationalgrps=5]	0[b]	.	.	0	.	.	.	.
	[Q8_time_UK=1]	.069	.214	.104	1	.747	1.071	.704	1.630
	[Q8_time_UK=2]	-.266	.211	1.587	1	.208	.766	.506	1.160
	[Q8_time_UK=3]	-.193	.224	.742	1	.389	.825	.532	1.279
	[Q8_time_UK=4]	.041	.165	.061	1	.805	1.042	.754	1.440
	[Q8_time_UK=5]	0[b]	.	.	0	.	.	.	.

a. The reference category is: With family.
b. This parameter is set to zero because it is redundant.
c. Floating point overflow occurred while computing this statistic. Its value is therefore set to system missing.

Likelihood Ratio Tests

Effect	Model Fitting Criteria: -2 Log Likelihood of Reduced Model	Likelihood Ratio Tests: Chi-Square	df	Sig.
Intercept	7740.93	.000	0	.
Q6_Childcount	7798.41	57.477	3	.000
Q31_health	7783.91	42.979	12	.000
Q21_contact_nationalgrps	7769.44	28.509	12	.005
Q8_time_UK	7872.13	131.199	12	.000
Q1_gender	7751.58	10.651	3	.014
Q2_age_group	7789.08	48.147	12	.000
Q18_spk_English	7776.92	35.996	9	.000
Q30_meet_friends	7790.31	49.382	15	.000
Friends_relatives_recoded	7772.05	31.122	3	.000
Q24_need_help	7759.13	18.201	3	.000
Q9_origin	8081.94	341.011	45	.000
Region	8685.09	944.164	15	.000
Q4_partner	7862.54	121.608	3	.000
Q10_religion	7779.20	38.276	21	.012
Q28_meet_relatives	7843.38	102.455	15	.000
Q14_Emp_recoded	7775.55	34.624	18	.011

The chi-square statistic is the difference in -2 log-likelihoods between the final model and a reduced model. The reduced model is formed by omitting an effect from the final model. The null hypothesis is that all parameters of that effect are 0.

a. This reduced model is equivalent to the final model because omitting the effect does not increase the degrees of freedom.

Table 7.17: Cross-sectional baseline model of accommodation, SNR.

Parameter Estimates

Where are you currently living? [a]		B	Std. Error	Wald	df	Sig.	Exp(B)	95% Confidence Interval for Exp (B) Lower Bound	95% Confidence Interval for Exp (B) Upper Bound
NASS accommodation	Intercept	1.480	.738	4.020	1	.045			
	Q6_Childcount	.013	.068	.036	1	.850	1.013	.887	1.157
	[Q24_need_help=1]	.502	.119	17.700	1	.000	1.652	1.308	2.087
	[Q24_need_help=2]	0[b]	.	.	0	.	.	.	.
	[Q21_contact_ nationalgrps_recoded=1]	.045	.124	.132	1	.716	1.046	.820	1.335
	[Q21_contact_ nationalgrps_recoded=2]	0[b]	.	.	0	.	.	.	.
	[Q2_age_group=1]	1.078	.390	7.639	1	.006	2.938	1.368	6.311
	[Q2_age_group=2]	1.292	.381	11.502	1	.001	3.641	1.725	7.685
	[Q2_age_group=3]	1.430	.392	13.293	1	.000	4.180	1.938	9.019
	[Q2_age_group=4]	1.061	.402	6.970	1	.008	2.890	1.314	6.354
	[Q2_age_group=5]	0[b]	.	.	0	.	.	.	.
	[Q1_gender=1]	-.042	.139	.092	1	.762	.959	.730	1.260
	[Q1_gender=2]	0[b]	.	.	0	.	.	.	.
	[Q9_origin=1]	-.108	.419	.067	1	.796	.897	.395	2.039
	[Q9_origin=2]	.831	.385	4.650	1	.031	2.295	1.079	4.885
	[Q9_origin=3]	-.887	.787	1.272	1	.259	.412	.088	1.924
	[Q9_origin=4]	1.673	.471	12.608	1	.000	5.329	2.116	13.419
	[Q9_origin=5]	2.958	.478	38.355	1	.000	19.257	7.552	49.106
	[Q9_origin=6]	3.045	.776	15.400	1	.000	21.014	4.592	96.168
	[Q9_origin=7]	-.039	.360	.011	1	.915	.962	.475	1.949
	[Q9_origin=8]	1.564	.512	9.318	1	.002	4.780	1.751	13.051
	[Q9_origin=9]	-.479	.371	1.666	1	.197	.619	.299	1.282
	[Q9_origin=10]	1.142	.374	9.313	1	.002	3.132	1.505	6.521
	[Q9_origin=11]	.936	.378	6.130	1	.013	2.549	1.215	5.345
	[Q9_origin=12]	.814	.372	4.790	1	.029	2.257	1.089	4.679
	[Q9_origin=13]	1.216	.439	7.669	1	.006	3.373	1.427	7.975
	[Q9_origin=14]	.299	.435	.471	1	.493	1.348	.574	3.165
	[Q9_origin=15]	.499	.449	1.231	1	.267	1.646	.682	3.972
	[Q9_origin=16]	0[b]	.	.	0	.	.	.	.
	[Q30_meet_friends=1]	1.537	.355	18.733	1	.000	4.652	2.319	9.332
	[Q30_meet_friends=2]	1.351	.333	16.481	1	.000	3.860	2.011	7.410
	[Q30_meet_friends=3]	1.261	.325	15.100	1	.000	3.529	1.868	6.667
	[Q30_meet_friends=4]	1.141	.321	12.612	1	.000	3.129	1.667	5.873
	[Q30_meet_friends=5]	1.016	.336	9.116	1	.003	2.762	1.428	5.340
	[Q30_meet_friends=6]	0[b]	.	.	0	.	.	.	.
	[Q8_time_UK=1]	-.918	.208	19.531	1	.000	.399	.266	.600
	[Q8_time_UK=2]	-.442	.191	5.359	1	.021	.643	.442	.934
	[Q8_time_UK=3]	-.471	.223	4.465	1	.035	.625	.404	.966
	[Q8_time_UK=4]	-.306	.189	2.624	1	.105	.736	.508	1.066
	[Q8_time_UK=5]	0[b]	.	.	0	.	.	.	.
	[Q10_religion=1]	-.064	.399	.026	1	.872	.938	.429	2.051
	[Q10_religion=2]	.406	.356	1.301	1	.254	1.500	.747	3.012
	[Q10_religion=3]	1.256	.726	2.991	1	.084	3.510	.846	14.568
	[Q10_religion=4]	-.166	.613	.074	1	.786	.847	.255	2.814
	[Q10_religion=5]	.664	7723.505	.000	1	1.000	1.943	.000	.[c]
	[Q10_religion=6]	.411	.333	1.526	1	.217	1.508	.786	2.896
	[Q10_religion=7]	-.907	1.236	.539	1	.463	.404	.036	4.551
	[Q10_religion=8]	0[b]	.	.	0	.	.	.	.
	[Friends_relatives_ recoded=1]	-1.538	.367	17.543	1	.000	.215	.105	.441
	[Friends_relatives_ recoded=2]	0[b]	.	.	0	.	.	.	.
	[Q4_partner=1]	-1.100	.153	51.765	1	.000	.333	.247	.449
	[Q4_partner=2]	0[b]	.	.	0	.	.	.	.
	[Region=1]	-3.017	.291	107.249	1	.000	.049	.028	.087
	[Region=2]	-1.056	.299	12.507	1	.000	.348	.194	.624
	[Region=3]	.478	.325	2.167	1	.141	1.614	.853	3.051
	[Region=4]	-.019	.334	.003	1	.955	.981	.510	1.888
	[Region=5]	.106	.427	.061	1	.804	1.112	.481	2.569
	[Region=6]	0[b]	.	.	0	.	.	.	.
	[Q14_Emp_recoded=1]	.398	.269	2.199	1	.138	1.489	.880	2.520
	[Q14_Emp_recoded=2]	.449	.272	2.716	1	.099	1.567	.919	2.673
	[Q14_Emp_recoded=3]	.098	.324	.091	1	.762	1.103	.584	2.081
	[Q14_Emp_recoded=4]	-.004	.284	.000	1	.989	.996	.571	1.737
	[Q14_Emp_recoded=5]	.144	.280	.266	1	.606	1.155	.668	1.999
	[Q14_Emp_recoded=7]	.767	.705	1.182	1	.277	2.153	.540	8.580
	[Q14_Emp_recoded=8]	0[b]	.	.	0	.	.	.	.
	[Q31_health=1]	-.874	.355	6.051	1	.014	.417	.208	.837
	[Q31_health=2]	-.481	.349	1.896	1	.169	.618	.312	1.226
	[Q31_health=3]	-.322	.350	.848	1	.357	.725	.365	1.439
	[Q31_health=4]	-.237	.376	.397	1	.529	.789	.378	1.649
	[Q31_health=5]	0[b]	.	.	0	.	.	.	.
	[Q28_meet_relatives=1]	-1.453	.259	31.376	1	.000	.234	.141	.389
	[Q28_meet_relatives=2]	-1.465	.232	39.750	1	.000	.231	.146	.364
	[Q28_meet_relatives=3]	-1.099	.201	29.763	1	.000	.333	.225	.495
	[Q28_meet_relatives=4]	-.803	.178	20.286	1	.000	.448	.316	.635
	[Q28_meet_relatives=5]	-.384	.226	2.887	1	.089	.681	.437	1.061
	[Q28_meet_relatives=6]	0[b]	.	.	0	.	.	.	.
With friends	Intercept	-1.467	.754	3.778	1	.052			
	Q6_Childcount	-.467	.084	30.764	1	.000	.627	.532	.740
	[Q24_need_help=1]	.339	.118	8.241	1	.004	1.404	1.114	1.770
	[Q24_need_help=2]	0[b]	.	.	0	.	.	.	.
	[Q21_contact_ nationalgrps_recoded=1]	.379	.124	9.269	1	.002	1.460	1.144	1.864
	[Q21_contact_ nationalgrps_recoded=2]	0[b]	.	.	0	.	.	.	.
	[Q2_age_group=1]	1.092	.398	7.523	1	.006	2.979	1.366	6.500
	[Q2_age_group=2]	1.484	.388	14.632	1	.000	4.412	2.062	9.438
	[Q2_age_group=3]	1.634	.399	16.773	1	.000	5.122	2.344	11.194
	[Q2_age_group=4]	1.018	.411	6.149	1	.013	2.768	1.238	6.189
	[Q2_age_group=5]	0[b]	.	.	0	.	.	.	.
	[Q1_gender=1]	-.058	.139	.177	1	.674	.943	.719	1.238
	[Q1_gender=2]	0[b]	.	.	0	.	.	.	.
	[Q9_origin=1]	-.245	.383	.409	1	.522	.783	.370	1.657
	[Q9_origin=2]	.602	.360	2.802	1	.094	1.826	.902	3.694
	[Q9_origin=3]	.427	.507	.709	1	.400	1.532	.568	4.135
	[Q9_origin=4]	.748	.477	2.459	1	.117	2.114	.829	5.386
	[Q9_origin=5]	.401	.505	.630	1	.427	1.493	.555	4.018
	[Q9_origin=6]	.500	.850	.347	1	.556	1.649	.312	8.725
	[Q9_origin=7]	-.005	.343	.000	1	.989	.995	.509	1.948
	[Q9_origin=8]	.056	.549	.010	1	.919	1.057	.361	3.100
	[Q9_origin=9]	.008	.351	.001	1	.981	1.008	.507	2.007
	[Q9_origin=10]	.447	.361	1.534	1	.216	1.563	.771	3.169
	[Q9_origin=11]	.522	.370	1.995	1	.158	1.686	.817	3.480

a. The reference category is: With family.
b. This parameter is set to zero because it is redundant.
c. Floating point overflow occurred while computing this statistic. Its value is therefore set to system missing.

Likelihood Ratio Tests

Effect	Model Fitting Criteria: -2 Log Likelihood of Reduced Model	Likelihood Ratio Tests: Chi-Square	df	Sig.
Intercept	7838.67	.000	0	.
Q6_Childcount	7894.99	56.323	3	.000
Q24_need_help	7856.65	17.982	3	.000
Q21_contact_ nationalgrps_recoded	7857.51	18.842	3	.000
Q2_age_group	7889.05	50.380	12	.000
Q1_gender	7846.81	8.146	3	.043
Q9_origin	8198.23	359.563	45	.000
Q30_meet_friends	7890.43	51.765	15	.000
Q8_time_UK	8009.29	170.627	12	.000
Q10_religion	7877.86	39.190	21	.009
Friends_relatives_recoded	7870.32	31.659	3	.000
Q4_partner	7959.61	120.947	3	.000
Region	8799.19	960.522	15	.000
Q14_Emp_recoded	7875.40	36.738	18	.006
Q31_health	7880.75	42.084	12	.000
Q28_meet_relatives	7942.25	103.580	15	.000

The chi-square statistic is the difference in -2 log-likelihoods between the final model and a reduced model. The reduced model is formed by omitting an effect from the final model. The null hypothesis is that all parameters of that effect are 0.

a. This reduced model is equivalent to the final model because omitting the effect does not increase the degrees of freedom.

Case Processing Summary

		N	Marginal Percentage
Where are you currently living?	NASS accommodation	2136.38	47.5%
	With friends	890.07	19.8%
	With family	735.65	16.3%
	Other	738.93	16.4%
In total, how long have you spent living in the UK?	5 years or more	1049.03	23.3%
	At least 2 years but less than 5 years	901.34	20.0%
	At least 1 year but less than 2 years	429.58	9.5%
	At least 6 months but less than 1 year	792.27	17.6%
	Less than 6 months	1328.81	29.5%
What is your religion?	None	210.14	4.7%
	Christian	1836.58	40.8%
	Buddhist	62.78	1.4%
	Hindu	90.04	2.0%
	Jewish	4.92	.1%
	Muslim	2131.16	47.3%
	Sikh	32.12	.7%
	Other	133.29	3.0%
How often do you meet up with friends who are not living with you?	More than twice a week	457.12	10.2%
	Once or twice a week	804.08	17.9%
	Once or twice a month	868.74	19.3%
	Less than once a month	1017.74	22.6%
	Never	494.49	11.0%
	No friends in the UK	858.86	19.1%
Whether have friends or relatives in UK (recoded)	Friends or rels in the UK	3757.78	83.5%
	No friends or rels in UK	743.25	16.5%
Country of origin grouped	Turkey	186.72	4.1%
	Other Europe	248.20	5.5%
	Americas	48.16	1.1%
	DRC/Congo	186.10	4.1%
	Eritrea	780.86	17.3%
	Ethiopia	87.23	1.9%
	Somalia	654.64	14.5%
	Sudan	135.46	3.0%
	Zimbabwe	347.95	7.7%
	Other Africa	306.25	6.8%
	Iran	343.06	7.6%
	Iraq	431.71	9.6%
	Other Middle East	213.09	4.7%
	Afghanistan	137.52	3.1%
	Pakistan	120.31	2.7%
	Other Asia	273.77	6.1%
Are you living with your husband/wife or partner here in the UK?	Yes	950.93	21.1%
	No	3550.10	78.9%
Sex of respondent	Male	2864.32	63.6%
	Female	1636.71	36.4%
Region of residence in the UK (derived from postcode)	London and South East	1666.48	37.0%
	Midlands and East England	892.74	19.8%
	North East, Yorkshire and Humber	921.90	20.5%
	North West	577.84	12.8%
	Scotland and Northern Ireland	122.77	2.7%
	Wales and South West	319.30	7.1%
Age group (derived from raw age data)	18-24	1016.77	22.6%
	25-34	2150.30	47.8%
	35-44	883.36	19.6%
	45-64	365.38	8.1%
	65+	85.22	1.9%
Pre-uk employment recoded to group retired in 'other'	In employment	1295.18	28.8%
	Self-employed	961.95	21.4%
	Unemployed and looking for work	254.46	5.7%
	Student	719.73	16.0%
	Looking after home and family	693.13	15.4%
	Military	344.50	7.7%
	Other	232.08	5.2%
Have contacted national/ethnic group (recoded)	Yes	2281.66	50.7%
	No	2219.37	49.3%
How is your health in general?	Very good	1353.72	30.1%
	Good	1545.96	34.3%
	Fair	1126.12	25.0%
	Bad	349.08	7.8%
	Very bad	126.15	2.8%
Have you ever needed any kind of help or support from any of these groups or organisations?	Yes	2500.24	55.5%
	No	2000.79	44.5%
How often do you meet up with relatives who are not living with you?	More than twice a week	212.34	4.7%
	Once or twice a week	333.30	7.4%
	Once or twice a month	422.98	9.4%
	Less than once a month	610.99	13.6%
	Never	371.95	8.3%
	No relatives in the UK	2549.47	56.6%
Valid		4501.03	100.0%
Missing		1177.57	
Total		5678.60	
Subpopulation		4426[a]	

a. The dependent variable has only one value observed in 4420 (99.9%) subpopulations.

Appendix G – SNR longitudinal models: Stata output

Table 7.18: Model of benefit claims, unweighted SNR.

```
. xtlogit benefits q2_age_group q1_gender q9_origin q3_accom, i(ref) nolog or

Random-effects logistic regression              Number of obs     =      3,651
Group variable: ref                             Number of groups  =      1,754

Random effects u_i ~ Gaussian                   Obs per group:
                                                              min =          1
                                                              avg =        2.1
                                                              max =          3

Integration method: mvaghermite                 Integration pts.  =         12

                                                Wald chi2(4)      =     118.97
Log likelihood  = -1790.9484                    Prob > chi2       =     0.0000

------------------------------------------------------------------------------
    benefits | Odds Ratio   Std. Err.      z    P>|z|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
q2_age_group |   .6139449    .0629727    -4.76   0.000     .5021358    .7506503
   q1_gender |   .1915447     .041099    -7.70   0.000     .1257856    .2916818
   q9_origin |   1.100326    .0267739     3.93   0.000     1.049082    1.154073
    q3_accom |   1.923079    .1690297     7.44   0.000     1.618752    2.284619
       _cons |   .3838024    .1656762    -2.22   0.027     .1646907    .8944299
-------------+----------------------------------------------------------------
    /lnsig2u |   2.103412    .1246446                      1.859113    2.347711
-------------+----------------------------------------------------------------
     sigma_u |    2.86253    .1783995                      2.533385    3.234439
         rho |   .7135246    .0254783                      .6611146    .7607623
------------------------------------------------------------------------------
Note: Estimates are transformed only in the first equation.
Note: _cons estimates baseline odds (conditional on zero random effects).
LR test of rho=0: chibar2(01) = 500.21                 Prob >= chibar2 = 0.000
```

Table 7.19: Model of victim of attack, unweighted SNR.

```
. xtlogit victim_attacknum q2_age_group q1_gender q9_origin q3_accom, i(ref) nolog or

Random-effects logistic regression              Number of obs     =      3,713
Group variable: ref                             Number of groups  =      1,768

Random effects u_i ~ Gaussian                   Obs per group:
                                                              min =          1
                                                              avg =        2.1
                                                              max =          3

Integration method: mvaghermite                 Integration pts.  =         12

                                                Wald chi2(4)      =       6.29
Log likelihood  = -1888.1878                    Prob > chi2       =     0.1786

----------------------------------------------------------------------------------
victim_attacknum | Odds Ratio   Std. Err.      z    P>|z|     [95% Conf. Interval]
-----------------+----------------------------------------------------------------
    q2_age_group |    .931108    .0387362    -1.72   0.086     .8581993    1.010211
       q1_gender |    .933817    .0780254    -0.82   0.412     .7927556    1.099979
       q9_origin |   .9874882    .0099547    -1.25   0.212     .9681688    1.007193
        q3_accom |   .9790315    .0342503    -0.61   0.545     .9141519    1.048516
           _cons |   5.870652    1.091213     9.52   0.000       4.0782    8.450925
-----------------+----------------------------------------------------------------
        /lnsig2u |  -14.88197    19.94409                     -53.97166    24.20773
-----------------+----------------------------------------------------------------
         sigma_u |   .0005867    .0058507                      1.91e-12    180568.4
             rho |   1.05e-07    2.09e-06                      1.10e-24           1
----------------------------------------------------------------------------------
Note: Estimates are transformed only in the first equation.
Note: _cons estimates baseline odds (conditional on zero random effects).
LR test of rho=0: chibar2(01) = 0.00                   Prob >= chibar2 = 1.000
```

Table 7.20: Model of wanting to stay, unweighted SNR.

```
. xtlogit stay_towncitynum q2_age_group q1_gender q9_origin q3_accom, i(ref) nolog or

Random-effects logistic regression              Number of obs     =      3,682
Group variable: ref                             Number of groups  =      1,758

Random effects u_i ~ Gaussian                   Obs per group:
                                                              min =          1
                                                              avg =        2.1
                                                              max =          3

Integration method: mvaghermite                 Integration pts.  =         12

                                                Wald chi2(4)      =      16.24
Log likelihood  = -1309.6182                    Prob > chi2       =     0.0027

------------------------------------------------------------------------------
stay_towncitynum | Odds Ratio   Std. Err.      z    P>|z|     [95% Conf. Interval]
-----------------+------------------------------------------------------------
    q2_age_group |   1.186602   .0994022     2.04   0.041      1.00693    1.398333
       q1_gender |   1.044417   .1696112     0.27   0.789     .7596975    1.435844
       q9_origin |   .9376492   .0186074    -3.24   0.001     .9018795    .9748375
        q3_accom |   1.117769   .0769813     1.62   0.106     .9766284    1.279307
           _cons |    15.6343   5.836825     7.36   0.000     7.521342    32.49836
-----------------+------------------------------------------------------------
        /lnsig2u |    1.12346   .1698517                      .7905566    1.456363
-----------------+------------------------------------------------------------
         sigma_u |   1.753704   .1489348                      1.484797    2.071311
             rho |   .4831595   .0424148                      .4012425    .5659917
------------------------------------------------------------------------------
Note: Estimates are transformed only in the first equation.
Note: _cons estimates baseline odds (conditional on zero random effects).
LR test of rho=0: chibar2(01) = 122.18                 Prob >= chibar2 = 0.000
```

Table 7.21: Model of economic activity, unweighted SNR.

```
. xtlogit econ_activity_bin q2_age_group q1_gender q9_origin q3_accom, i(ref) nolog o

Random-effects logistic regression              Number of obs     =      2,194
Group variable: ref                             Number of groups  =      1,197

Random effects u_i ~ Gaussian                   Obs per group:
                                                              min =          1
                                                              avg =        1.8
                                                              max =          3

Integration method: mvaghermite                 Integration pts.  =         12

                                                Wald chi2(4)      =      47.38
Log likelihood  = -1168.9137                    Prob > chi2       =     0.0000

------------------------------------------------------------------------------
econ_activity_bin | Odds Ratio   Std. Err.      z    P>|z|     [95% Conf. Interval]
------------------+-----------------------------------------------------------
     q2_age_group |   .8078447   .1054666    -1.63   0.102     .6254621     1.04341
        q1_gender |   .5289493   .1404606    -2.40   0.016     .3143261    .8901182
        q9_origin |   1.045575   .0298891     1.56   0.119      .988604    1.105828
         q3_accom |   2.013659    .217693     6.47   0.000     1.629161    2.488901
            _cons |   2.621611   1.414149     1.79   0.074     .9107795    7.546114
------------------+-----------------------------------------------------------
         /lnsig2u |   2.095111   .1772225                      1.747761    2.442461
------------------+-----------------------------------------------------------
          sigma_u |   2.850674   .2526018                      2.396192    3.391358
              rho |   .7118249   .0363537                      .6357381     .777579
------------------------------------------------------------------------------
Note: Estimates are transformed only in the first equation.
Note: _cons estimates baseline odds (conditional on zero random effects).
LR test of rho=0: chibar2(01) = 225.01                 Prob >= chibar2 = 0.000
```

Table 7.22: Model of health, unweighted SNR.

```
. xtlogit health_grouped_bin q2_age_group q1_gender q9_origin q3_accom, i(ref) nolog or

Random-effects logistic regression              Number of obs     =      8,940
Group variable: ref                             Number of groups  =      5,234

Random effects u_i ~ Gaussian                   Obs per group:
                                                              min =          1
                                                              avg =        1.7
                                                              max =          4

Integration method: mvaghermite                 Integration pts.  =         12

                                                Wald chi2(4)      =     259.91
Log likelihood  = -4906.9204                    Prob > chi2       =     0.0000

------------------------------------------------------------------------------------
health_grouped_bin | Odds Ratio   Std. Err.      z    P>|z|     [95% Conf. Interval]
-------------------+----------------------------------------------------------------
      q2_age_group |   .3944915   .0279518   -13.13   0.000      .343341    .4532624
         q1_gender |   .2362249   .0312759   -10.90   0.000     .1822332    .3062132
         q9_origin |   .9744547   .0148924    -1.69   0.090     .9456989    1.004085
          q3_accom |   1.260317   .0686076     4.25   0.000     1.132774    1.402221
             _cons |   174.4746   54.36045    16.57   0.000     94.73847    321.3203
-------------------+----------------------------------------------------------------
          /lnsig2u |   2.274774   .0889867                      2.100363    2.449185
-------------------+----------------------------------------------------------------
           sigma_u |   3.118609   .1387573                       2.85817    3.402779
               rho |   .7472363   .0168073                      .7129011    .7787397
------------------------------------------------------------------------------------
Note: Estimates are transformed only in the first equation.
Note: _cons estimates baseline odds (conditional on zero random effects).
LR test of rho=0: chibar2(01) = 1293.92                Prob >= chibar2 = 0.000
```

Table 7.23: Model of perception of job being below skill level, unweighted SNR.

```
. xtlogit work_skills_bin q2_age_group q1_gender q9_origin q3_accom, i(ref) nolog or

Random-effects logistic regression              Number of obs     =      1,493
Group variable: ref                             Number of groups  =        842

Random effects u_i ~ Gaussian                   Obs per group:
                                                              min =          1
                                                              avg =        1.8
                                                              max =          3

Integration method: mvaghermite                 Integration pts.  =         12

                                                Wald chi2(4)      =      16.20
Log likelihood  = -875.02341                    Prob > chi2       =     0.0028

---------------------------------------------------------------------------------
work_skills_bin |         OR   Std. Err.      z    P>|z|     [95% Conf. Interval]
----------------+----------------------------------------------------------------
   q2_age_group |    1.44943   .2701554     1.99   0.046     1.005875    2.088577
      q1_gender |   .9043831   .3364578    -0.27   0.787     .4361941    1.875103
      q9_origin |   .9301306   .0354886    -1.90   0.058     .8631113    1.002354
       q3_accom |   .6314208   .0889071    -3.27   0.001     .4791435    .8320937
          _cons |   6.568094    4.87848     2.53   0.011     1.531813    28.16262
----------------+----------------------------------------------------------------
       /lnsig2u |   2.457373   .2000608                      2.065262    2.849485
----------------+----------------------------------------------------------------
        sigma_u |   3.416739   .3417778                      2.808444    4.156788
            rho |   .7801475   .0343139                      .7056633    .8400551
---------------------------------------------------------------------------------
LR test of rho=0: chibar2(01) = 256.79                 Prob >= chibar2 = 0.000
```

Table 7.24: Model of benefit claims, weighted SNR.

```
. xtlogit benefits q2_age_group q1_gender q9_origin q3_accom wtl_b123, i(ref) nolog or

Random-effects logistic regression              Number of obs     =      2,389
Group variable: ref                             Number of groups  =        816

Random effects u_i ~ Gaussian                   Obs per group:
                                                              min =          1
                                                              avg =        2.9
                                                              max =          3

Integration method: mvaghermite                 Integration pts.  =         12

                                                Wald chi2(5)      =      70.42
Log likelihood  = -1118.4258                    Prob > chi2       =     0.0000

------------------------------------------------------------------------------
    benefits | Odds Ratio   Std. Err.      z    P>|z|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
q2_age_group |   .4694981   .0732722    -4.84   0.000      .345773    .6374947
   q1_gender |   .1701052   .0488684    -6.17   0.000     .0968683    .2987127
   q9_origin |   1.073042   .0366656     2.06   0.039     1.003532    1.147366
    q3_accom |   1.613025   .1904074     4.05   0.000     1.279859    2.032919
    wtl_b123 |   .7042579   .1685559    -1.46   0.143     .4405615    1.125789
       _cons |    2.24156   1.885091     0.96   0.337     .4312339    11.65166
-------------+----------------------------------------------------------------
    /lnsig2u |    2.12328    .140134                      1.848623    2.397938
-------------+----------------------------------------------------------------
     sigma_u |   2.891109   .2025713                      2.520132    3.316695
         rho |   .7175686   .0284001                      .6587604    .7697837
------------------------------------------------------------------------------
Note: Estimates are transformed only in the first equation.
Note: _cons estimates baseline odds (conditional on zero random effects).
LR test of rho=0: chibar2(01) = 427.62                 Prob >= chibar2 = 0.000
```

Table 7.25: Model of victim of attack, weighted SNR.

```
. xtlogit victim_attacknum q2_age_group q1_gender q9_origin q3_accom wtl_b123, i(ref) nolog
> or

Random-effects logistic regression              Number of obs     =      2,427
Group variable: ref                             Number of groups  =        817

Random effects u_i ~ Gaussian                   Obs per group:
                                                              min =          1
                                                              avg =        3.0
                                                              max =          3

Integration method: mvaghermite                 Integration pts.  =         12

                                                Wald chi2(5)      =       3.44
Log likelihood  = -1463.4398                    Prob > chi2       =     0.6328

----------------------------------------------------------------------------------
victim_attacknum | Odds Ratio   Std. Err.      z    P>|z|     [95% Conf. Interval]
-----------------+----------------------------------------------------------------
    q2_age_group |   .9595385   .0489913    -0.81   0.419     .8681655    1.060528
       q1_gender |   .9564597   .0882466    -0.48   0.629     .7982365    1.146045
       q9_origin |   1.000088   .0115052     0.01   0.994     .9777908    1.022894
        q3_accom |   .9634487   .0381121    -0.94   0.347     .8915726    1.041119
        wtl_b123 |   .8785019   .0679148    -1.68   0.094     .7549853    1.022226
           _cons |   3.530512   1.000718     4.45   0.000     2.025657    6.153318
-----------------+----------------------------------------------------------------
        /lnsig2u |  -17.94672   36.13768                     -88.77526    52.88183
-----------------+----------------------------------------------------------------
         sigma_u |   .0001267   .0022901                      5.28e-20    3.04e+11
             rho |   4.88e-09   1.76e-07                      8.48e-40           1
----------------------------------------------------------------------------------
Note: Estimates are transformed only in the first equation.
Note: _cons estimates baseline odds (conditional on zero random effects).
LR test of rho=0: chibar2(01) = 0.00                   Prob >= chibar2 = 1.000
```

Table 7.26: Model of want to stay, weighted SNR.

```
. xtlogit stay_towncitynum q2_age_group q1_gender q9_origin q3_accom wtl_b123, i(ref) nolog
> or

Random-effects logistic regression              Number of obs     =      2,414
Group variable: ref                             Number of groups  =        816

Random effects u_i ~ Gaussian                   Obs per group:
                                                              min =          1
                                                              avg =        3.0
                                                              max =          3

Integration method: mvaghermite                 Integration pts.  =         12

                                                Wald chi2(5)      =      10.81
Log likelihood  = -866.93371                    Prob > chi2       =     0.0552

------------------------------------------------------------------------------
stay_towncitynum | Odds Ratio   Std. Err.      z    P>|z|     [95% Conf. Interval]
-----------------+------------------------------------------------------------
    q2_age_group |   1.174793    .1377033     1.37   0.169     .9336576    1.478205
       q1_gender |   1.002826    .2086406     0.01   0.989     .6670081    1.507719
       q9_origin |   .9607646     .025521    -1.51   0.132     .9120242     1.01211
        q3_accom |   1.122401    .1006161     1.29   0.198     .9415501     1.33799
        wtl_b123 |   1.596949    .3276604     2.28   0.023     1.068175     2.38748
           _cons |   7.908779    5.282429     3.10   0.002     2.135872     29.2849
-----------------+------------------------------------------------------------
        /lnsig2u |   1.116942    .1860925                       .7522073    1.481677
-----------------+------------------------------------------------------------
         sigma_u |   1.747998    .1626447                       1.456598    2.097693
             rho |    .481532    .0464597                        .392065    .5721991
------------------------------------------------------------------------------
Note: Estimates are transformed only in the first equation.
Note: _cons estimates baseline odds (conditional on zero random effects).
LR test of rho=0: chibar2(01) = 105.58                 Prob >= chibar2 = 0.000
```

Table 7.27: Model of economic activity, weighted SNR.

```
. xtlogit econ_activity_bin q2_age_group q1_gender q9_origin q3_accom wtl_b123, i(ref) nolog
>  or

Random-effects logistic regression              Number of obs     =      1,430
Group variable: ref                             Number of groups  =        597

Random effects u_i ~ Gaussian                   Obs per group:
                                                              min =          1
                                                              avg =        2.4
                                                              max =          3

Integration method: mvaghermite                 Integration pts.  =         12

                                                Wald chi2(5)      =      25.22
Log likelihood  = -716.14482                    Prob > chi2       =     0.0001

------------------------------------------------------------------------------
econ_activity_bin | Odds Ratio   Std. Err.      z    P>|z|     [95% Conf. Interval]
------------------+-----------------------------------------------------------
     q2_age_group |   .6248165    .1193949    -2.46   0.014     .4296336    .9086713
        q1_gender |   .5489142    .1852715    -1.78   0.076     .2832722    1.063665
        q9_origin |   1.026723    .0382849     0.71   0.479     .9543622     1.10457
         q3_accom |   1.624069    .2169847     3.63   0.000     1.249911     2.11023
         wtl_b123 |   .6105992    .1581016    -1.91   0.057     .3675834    1.014277
            _cons |   17.54339    17.56708     2.86   0.004     2.464691    124.8719
------------------+-----------------------------------------------------------
         /lnsig2u |   1.968841    .1965427                       1.583624    2.354058
------------------+-----------------------------------------------------------
          sigma_u |   2.676261    .2629998                       2.207393    3.244719
              rho |   .6852475     .042391                        .596951    .7619155
------------------------------------------------------------------------------
Note: Estimates are transformed only in the first equation.
Note: _cons estimates baseline odds (conditional on zero random effects).
LR test of rho=0: chibar2(01) = 178.42                 Prob >= chibar2 = 0.000
```

Appendix G

Table 7.28: Model of health, weighted SNR.

```
. xtlogit health_grouped_bin q2_age_group q1_gender q9_origin q3_accom wtl_b123, i(ref) nolo
> g or

Random-effects logistic regression              Number of obs     =      3,235
Group variable: ref                             Number of groups  =        817

Random effects u_i ~ Gaussian                   Obs per group:
                                                              min =          2
                                                              avg =        4.0
                                                              max =          4

Integration method: mvaghermite                 Integration pts.  =         12

                                                Wald chi2(5)      =      81.74
Log likelihood  =  -1536.576                    Prob > chi2       =     0.0000
```

health_grouped_bin	Odds Ratio	Std. Err.	z	P>|z|	[95% Conf.	Interval]
q2_age_group	.3310162	.0512137	-7.15	0.000	.2444298	.4482748
q1_gender	.2274225	.0624537	-5.39	0.000	.132764	.3895707
q9_origin	.9657584	.0327854	-1.03	0.305	.9035914	1.032202
q3_accom	1.241632	.1440075	1.87	0.062	.9891647	1.558538
wtl_b123	.8699681	.2029744	-0.60	0.550	.5506903	1.374356
_cons	435.6687	374.1476	7.08	0.000	80.93821	2345.088
/lnsig2u	2.323489	.1170552			2.094065	2.552913
sigma_u	3.195503	.1870252			2.849184	3.583918
rho	.7563261	.0215729			.7116103	.7960952

```
Note: Estimates are transformed only in the first equation.
Note: _cons estimates baseline odds (conditional on zero random effects).
LR test of rho=0: chibar2(01) = 917.30                 Prob >= chibar2 = 0.000
```

Table 7.29: Model of perception of job being below skill level, weighted SNR.

```
. xtlogit work_skills_bin q2_age_group q1_gender q9_origin q3_accom wtl_b123, i(ref) nolog or

Random-effects logistic regression              Number of obs     =      1,012
Group variable: ref                             Number of groups  =        464

Random effects u_i ~ Gaussian                   Obs per group:
                                                              min =          1
                                                              avg =        2.2
                                                              max =          3

Integration method: mvaghermite                 Integration pts.  =         12

                                                Wald chi2(5)      =      17.23
Log likelihood  = -572.00743                    Prob > chi2       =     0.0041
```

work_skills_bin	OR	Std. Err.	z	P>|z|	[95% Conf.	Interval]
q2_age_group	.864627	.2088995	-0.60	0.547	.5384843	1.388304
q1_gender	.7432665	.3308747	-0.67	0.505	.3106141	1.778558
q9_origin	.8923849	.0421745	-2.41	0.016	.8134373	.9789948
q3_accom	.5776383	.1010433	-3.14	0.002	.4099771	.8138648
wtl_b123	.3286453	.1164488	-3.14	0.002	.1641048	.6581631
_cons	149.3446	202.5411	3.69	0.000	10.46622	2131.027
/lnsig2u	2.281856	.218522			1.853561	2.710152
sigma_u	3.129672	.3419511			2.526363	3.877055
rho	.7485717	.0411285			.6598697	.820436

```
LR test of rho=0: chibar2(01) = 198.52                 Prob >= chibar2 = 0.000
```

Appendix H – Variable harmonisation

Table 7.30: Variable categories in the SNR and APS

Variable	SNR	APS
Age	**Q2_age_group** 18-24, 25-34, 35-44, 45-64, 65+	**aage** 0-15, 16-17, **|**18-19, 20-24, **|**25-29, 30-34, **|**35-39, 40-44, **|**45-49, 50-54, 55-59, 60-64, **|**65-99
Sex	**Q1_gender** Male, female	**sex** Male, female
Country of origin	**Q9_origin** 16 countries (or regions)	**cryo7** All individual countries
Region of UK (usual residence)	**Region** London and South East, Midlands and East England, North East, Yorkshire and Humber, North West, Scotland and Northern Ireland, Wales and South West	**GOVTOF** London, South East **|** East Midlands, West Midlands, Eastern**|** North East, Yorkshire & Humberside**|** North West, Merseyside**|** Scotland, Northern Ireland**|** Wales, South West

Source: Author's creation using SNR and APS data (Home Office, 2010e; ONS, 2018).

Appendix I – Borrowing information on individuals: models of dispersal

Table 7.31: Model 1 of dispersal showing coefficients based on SNR data.

Variable		Coef.	Std. Err.	z	P>z	[95% Conf. Interval]	
Age group	18-24	0.44	0.31	1.45	0.146	-0.15	1.04
	25-34	0.57	0.30	1.91	0.056	-0.02	1.17
	35-44	0.59	0.31	1.91	0.056	-0.01	1.19
	45-64	0.55	0.32	1.69	0.09	-0.09	1.18
	65+	0.00	(omitted)				
Sex	Female	0.12	0.08	1.35	0.176	-0.05	0.28
	Male	0.00	(omitted)				
Period of arrival	2014	1.61	0.13	12.48	0	1.36	1.87
	2013	1.38	0.13	10.39	0	1.12	1.64
	2012	1.08	0.15	7.19	0	0.78	1.37
	2008-11	0.92	0.12	7.72	0	0.69	1.16
	2000-07	0.00	(omitted)				
Region	London and South East	-2.70	0.17	-16.38	0	-3.03	-2.38
	Midlands and E.England	-0.81	0.16	-5.04	0	-1.13	-0.50
	NE, Yorks. and Humber	0.39	0.16	2.35	0.019	0.06	0.71
	North West	-0.20	0.17	-1.19	0.236	-0.54	0.13
	Scotland and NI	0.36	0.25	1.47	0.142	-0.12	0.85
	South West and Wales	0.00	(omitted)				
Country of origin	Turkey	-0.09	0.29	-0.31	0.76	-0.67	0.49
	Other Europe	0.60	0.25	2.43	0.015	0.12	1.09
	Americas	-0.64	0.60	-1.07	0.285	-1.81	0.53
	DRC/Congo	1.32	0.25	5.34	0	0.84	1.81
	Eritrea	2.66	0.22	12.29	0	2.24	3.09
	Ethiopia	2.45	0.35	7	0	1.77	3.14
	Somalia	0.35	0.20	1.73	0.083	-0.05	0.74
	Sudan	1.81	0.28	6.35	0	1.25	2.37
	Zimbabwe	-0.35	0.22	-1.64	0.101	-0.77	0.07
	Other Africa	1.05	0.23	4.67	0	0.61	1.49
	Iran	0.92	0.21	4.37	0	0.51	1.34
	Iraq	0.41	0.21	1.95	0.051	0.00	0.82
	Other Middle East	0.35	0.24	1.49	0.137	-0.11	0.82
	Afghanistan	0.25	0.30	0.82	0.414	-0.34	0.83
	Pakistan	0.46	0.28	1.64	0.1	-0.09	1.00
	Other Asia	0.00	(omitted)				
Constant		*					

*constant removed due to disclosure control requirements

Source: Author's analysis of SNR data (Home Office, 2010e).

Table 7.32: Model 2 of dispersal showing coefficients based on SNR data.

	Variable	Coef.	Std. Err.	z	P>z	[95% Conf. Interval]	
Age group	18-24	0.47	0.30	1.56	0.120	-0.12	1.06
	25-34	0.46	0.30	1.55	0.120	-0.12	1.04
	35-44	0.42	0.30	1.38	0.166	-0.17	1.02
	45-64	0.37	0.32	1.15	0.248	-0.26	0.99
	65+	0.00	(omitted)				
Sex	Female	0.16	0.08	1.89	0.059	-0.01	0.32
	Male	0.00	(omitted)				
Region	London and South East	-2.83	0.16	-17.64	0.000	-3.15	-2.52
	Midlands and E.England	-0.94	0.16	-6.04	0.000	-1.24	-0.63
	NE, Yorks. and Humber	0.35	0.16	2.23	0.026	0.04	0.66
	North West	-0.19	0.17	-1.14	0.253	-0.52	0.14
	Scotland and NI	0.18	0.24	0.75	0.451	-0.29	0.65
	South West and Wales	0.00	(omitted)				
Country of origin	Turkey	-0.34	0.29	-1.19	0.236	-0.90	0.22
	Other Europe	0.33	0.24	1.37	0.171	-0.14	0.80
	Americas	-1.13	0.59	-1.91	0.056	-2.29	0.03
	DRC/Congo	1.40	0.24	5.9	0.000	0.94	1.87
	Eritrea	3.02	0.21	14.57	0.000	2.61	3.42
	Ethiopia	2.49	0.33	7.52	0.000	1.84	3.14
	Somalia	0.69	0.19	3.59	0.000	0.31	1.07
	Sudan	2.03	0.28	7.33	0.000	1.49	2.58
	Zimbabwe	-0.40	0.21	-1.95	0.051	-0.81	0.00
	Other Africa	0.84	0.21	3.9	0.000	0.42	1.26
	Iran	1.06	0.20	5.16	0.000	0.66	1.46
	Iraq	-0.02	0.20	-0.12	0.906	-0.42	0.37
	Other Middle East	0.59	0.23	2.56	0.011	0.14	1.04
	Afghanistan	0.24	0.29	0.83	0.409	-0.33	0.81
	Pakistan	0.39	0.27	1.45	0.148	-0.14	0.93
	Other Asia	0.00	(omitted)				
Constant		*					

*constant removed due to disclosure control

Source: Author's analysis of SNR data (Home Office, 2010e).

Table 7.33: Model 3 of dispersal showing coefficients based on SNR data.

	Variable	Coef.	Std. Err.	z	P>z	[95% Conf. Interval]	
Region	London and South East	-2.81	0.16	-17.74	0.000	-3.12	-2.50
	Midlands and E.England	-0.92	0.15	-6.02	0.000	-1.23	-0.62
	NE, Yorks. and Humber	0.39	0.16	2.48	0.013	0.08	0.69
	North West	-0.17	0.16	-1.04	0.297	-0.49	0.15
	Scotland and NI	0.22	0.24	0.94	0.349	-0.24	0.69
	South West and Wales	0.00	(omitted)				
Country of origin	Turkey	-0.42	0.28	-1.48	0.139	-0.97	0.14
	Other Europe	0.33	0.24	1.39	0.165	-0.14	0.80
	Americas	-1.16	0.59	-1.96	0.050	-2.31	0.00
	DRC/Congo	1.38	0.24	5.89	0.000	0.92	1.84
	Eritrea	3.02	0.20	14.81	0.000	2.62	3.41
	Ethiopia	2.50	0.33	7.58	0.000	1.86	3.15
	Somalia	0.65	0.19	3.43	0.001	0.28	1.02
	Sudan	2.01	0.27	7.33	0.000	1.47	2.55
	Zimbabwe	-0.41	0.20	-2.01	0.045	-0.81	-0.01
	Other Africa	0.81	0.21	3.84	0.000	0.40	1.23
	Iran	1.01	0.20	4.99	0.000	0.61	1.40
	Iraq	-0.09	0.20	-0.46	0.643	-0.48	0.30
	Other Middle East	0.54	0.23	2.37	0.018	0.09	0.99
	Afghanistan	0.21	0.29	0.71	0.475	-0.36	0.77
	Pakistan	0.37	0.27	1.37	0.171	-0.16	0.89
	Other Asia	0.00	(omitted)				
Constant		*					

*constant removed due to disclosure control

Source: Author’s analysis of SNR data (Home Office, 2010e).

Appendix J – Borrowing information on aggregates: contingency tables

Table 7.34: Predicted number dispersed and not dispersed by economic activity.

COUNTS	Employed	Unemployed	Inactive	**Total**
Dispersed	136	38	115	**288**
Not dispersed	239	48	183	**471**
Total	**375**	**86**	**298**	**759**
%				
Dispersed	47.1	13.1	39.8	**100**
Not dispersed	50.8	10.2	39.0	**100**
Total	**98**	**23**	**79**	**200**

Pearson chi2(2) = 1.9158 Pr = 0.384
Likelihood-ratio chi2(2) = 1.8952 Pr = 0.388
Fisher's exact = 0.377

Source: Author's analysis of SNR and APS data (Home Office, 2010e; ONS, 2018).

Table 7.35: Predicted number dispersed and not dispersed by economic activity, with Poisson model.

COUNTS	Employed	Unemployed	Inactive	**Total**
Dispersed	141	38	118	**297**
Not dispersed	234	48	180	**462**
Total	**375**	**86**	**298**	**759**
%	Employed	Unemployed	Inactive	**Total**
Dispersed	47.4	12.8	39.7	**100**
Not dispersed	50.7	10.4	39.0	**100**
Total	**98**	**23**	**79**	**200**

Pearson chi2(2) = 1.3189 Pr = 0.517
Likelihood-ratio chi2(2) = 1.3083 Pr = 0.520
Fisher's exact = 0.515

Source: Author's analysis of SNR and APS data (Home Office, 2010e; ONS, 2018).

Table 7.36: Predicted number dispersed and not dispersed by benefits claims.

COUNTS	Yes	No	**Total**
Dispersed	194	91	**285**
Not dispersed	298	164	**462**
Total	**492**	**255**	**747**
%	Yes	No	**Total**
Dispersed	68.1	31.9	**100**
Not dispersed	64.5	35.5	**100**
Total	**133**	**67**	**200**

Pearson chi2(1) = 0.9981 Pr = 0.318

Likelihood-ratio chi2(1) = 1.0024 Pr = 0.317

Fisher's exact = 0.341

1-sided Fisher's exact = 0.179

Source: Author's analysis of SNR and APS data (Home Office, 2010e; ONS, 2018).

Table 7.37: Predicted number dispersed and not dispersed by benefits claims, with Poisson model.

COUNTS	Yes	No	**Total**
Dispersed	199	94	**293**
Not dispersed	293	161	**454**
Total	**492**	**255**	**747**
%	Employed	Unemployed	**Total**
Dispersed	67.8	32.2	**100**
Not dispersed	64.6	35.4	**100**
Total	**132**	**68**	**200**

Pearson chi2(1) = 0.9052 Pr = 0.341

Likelihood-ratio chi2(1) = 0.9086 Pr = 0.340

Fisher's exact = 0.385

1-sided Fisher's exact = 0.192

Source: Author's analysis of SNR and APS data (Home Office, 2010e; ONS, 2018).

Table 7.38: Predicted number dispersed and not dispersed by health status.

COUNTS	Very good	Good	Fair	Bad or Very Bad	**Total**
Dispersed Not dispersed	100 151	111 189	43 62	28* 48*	282 450
Total	251	300	105	76*	732
%	Very good	Good	Fair	Bad or Very Bad	**Total**
Dispersed Not dispersed	35.6 33.5	39.2 42.1	15.1 13.8	10.1* 10.6*	100 100
Total	69.1	81.3	29.0	20.6*	200

*Categories ‘Bad’ and ‘Very Bad’ combined for presentation due to disclosure control

Pearson chi2(4) = 0.8719 Pr = 0.929
Likelihood-ratio chi2(4) = 0.8718 Pr = 0.929
Fisher's exact = 0.928

Source: Author’s analysis of SNR and APS data (Home Office, 2010e; ONS, 2018).

Table 7.39: Predicted number dispersed and not dispersed by health status, with Poisson model.

COUNTS	Very good	Good	Fair	Bad or Very Bad	**Total**
Dispersed	103	115	44	30*	291
Not dispersed	148	185	61	46*	441
Total	251	300	105	76*	**732**
%	Very good	Good	Fair	Bad	Total
Dispersed	35.4	39.4	15.0	10.2*	100
Not dispersed	33.6	42.0	13.9	10.5*	100
Total	68.9	81.4	28.9	20.7*	**200**

*Categories ‘Bad’ and ‘Very Bad’ combined for presentation due to disclosure control

Pearson chi2(4) = 0.7139 Pr = 0.950
Likelihood-ratio chi2(4) = 0.7137 Pr = 0.950
Fisher's exact = 0.950

Source: Author’s analysis of SNR and APS data (Home Office, 2010e; ONS, 2018).

Table 7.40: Predicted number dispersed and not dispersed who are renting by housing landlord.

COUNTS	LA	Housing association	Another organisation	Individual employer	Other individual private landlord	Total
Dispersed	110	67	*	*	76	*
Not dispersed	158	103	*	*	165	*
Total	**268**	**170**	*	*	**241**	**705**
%	**LA**	**Housing association**	**Another organisation**	**Individual employer**	**Other individual private landlord**	**Total**
Dispersed	40.7	24.7	*	*	28.0	**100**
Not dispersed	36.4	23.8	*	*	38.1	**100**
Total	**77.0**	**48.4**	*	*	**66.0**	**200**

*Categories ‘Another organisation’ and ‘Individual employer’ removed for presentation due to disclosure control
Pearson chi2(4) = 18.2149 Pr = 0.001
Likelihood-ratio chi2(4) = 18.3497 Pr = 0.001
Fisher's exact = 0.001

Source: Author’s analysis of SNR and APS data (Home Office, 2010e; ONS, 2018).

Table 7.41: Predicted number dispersed and not dispersed who are renting by housing landlord, with Poisson model.

COUNTS	LA	Housing association	Another organisation	Individual employer	Other individual private landlord	Total
Dispersed	113	68	*	*	80	*
Not dispersed	155	102	*	*	161	*
Total	268	170	*	*	241	705
%	**LA**	**Housing association**	**Another organisation**	**Individual employer**	**Other individual private landlord**	
Dispersed	40.4	24.5	*	*	28.8	100
Not dispersed	36.5	23.9	*	*	37.7	100
Total	76.8	48.4	*	*	66.5	200

*Categories ‘Another organisation’ and ‘Individual employer’ removed for presentation due to disclosure control
Pearson chi2(4) = 13.9798 Pr = 0.007
Likelihood-ratio chi2(4) = 14.1873 Pr = 0.007
Fisher's exact = 0.005

Source: Author’s analysis of SNR and APS data (Home Office, 2010e; ONS, 2018).

List of References

1993 Asylum and Immigration Appeals Act c23. [Available at: http://www.legislation.gov.uk/ukpga/1993/23/enacted].

ADRN (2016) Legal issues for ADRN users [Available at: https://adrn.ac.uk/media/174205/legal_guide_final.pdf].

Ager, A. and Strang, A. (2004) Indicators of integration: final report. Home Office Development and Practice Report 28 [Available at: http://webarchive.nationalarchives.gov.uk/20110218135832/http:/rds.homeoffice.gov.uk/rds/pdfs04/dpr28.pdf].

Anie, A., Daniel, N., Tah, C. and Petruckevitch, A. (2005) 'An exploration of factors affecting the successful dispersal of asylum seekers', Home Office Online Report 50/05, London: Research, Development and Statistics Directorate.

Asylum and Immigration (Treatment of Claimants, etc.) Act 2004 c. 19. [Available at: http://www.legislation.gov.uk/ukpga/2004/19/contents].

Bijak, J., Disney, G., Lubman, S., & Wisniowski, A. (2013). Towards Reliable Migration Statistics for the United Kingdom: Response to the House of Commons Public Administration Select Committee Call for Evidence on Migration Statistics. Southampton, GB: Centre for Population Change.

Bloch, A. and Schuster, L. (2005) 'At the extremes of exclusion: Deportation, detention and dispersal', Ethnic and Racial Studies, 28:3, 491-512, [Available at: http://www.tandfonline.com/doi/pdf/10.1080/0141987042000337858].

Burnett, J. (2011) Public spending cuts savage dispersal system Institute of Race Relations, London [Available at: www.irr.org.uk/news/public-spending-cuts-savage-dispersal-system].

Calderwood, L. and Lessof, C. 'Enhancing Longitudinal Surveys by Linking to Administrative Data' in Lynn, P. (2009) Methodology of Longitudinal Surveys. Chichester: John Wiley and Sons.

Castles S. and Miller, M.J. (2009). The Age of Migration: International Population Movements in the Modern World (4th edition). Basingstoke: Palgrave MacMillan.

Cheung, S. Y. and Phillimore, J. (2013) Social networks, social capital and refugee integration. Report for Nuffield Foundation. Available at: http://www.birmingham.ac.uk/Documents/college-social-sciences/social-policy/iris/2013/nuffield-refugees-integration-research-report.pdf].

Commission of the European Communities, (2007) Green Paper on the future Common European Asylum System in Europe, Brussels.

Commission on Inequality in Education (2016) Educational inequalities in England and Wales, Social Market Foundation [Available at: http://www.smf.co.uk/wp-content/uploads/2016/01/Publication-Commission-on-Inequality-in-Education-Initial-Findings-Slide-Pack-120116.pdf]

Czaika, M. and Hobolth, M. (2016) 'Do restrictive asylum and visa policies increase irregular migration into Europe?' in European Union Politics, 17(3): 345-365.

Da Lomba, S., (2010) 'Legal Status and Refugee Integration: a UK Perspective', Journal of Refugee Studies, 23 (4): 415-436. [Available at: http://jrs.oxfordjournals.org/content/23/4/415.abstract].

Darling, J. (2013) 'Dispersal in the UK' part of Producing Urban Asylum [Available at: http://www.producingurbanasylum.com/four-cities/dispersal-in-the-uk/]

Department for Business, Innovation and Skills (BIS) (2013) Improving access for research and policy: The Government response to the Report of the Administrative Data Taskforce [Available at: https://www.gov.uk/government/publications/administrative-data-taskforce-report-government-response].

Department for Communities and Local Government (DCLG) (2008) English Indices of Deprivation 2007 [Available at: http://webarchive.nationalarchives.gov.uk/20100410180038/http://communities.gov.uk/documents/communities/pdf/733520.pdf].

Department for Communities and Local Government (DCLG) (2011) English Indices of Deprivation 2010 [Available at: https://www.gov.uk/government/uploads/system/uploads/attachment_data/file/6320/1870718.pdf].

D'Orazio, M., Di Zio, M. and Scanu, M. (2006) 'Statistical Matching for Categorical Data: Displaying Uncertainty and Using Logical Constraints.' Journal of Official Statistics 22: 137–157.

EMN (2014) EMN Inform: The Organisation of Reception Facilities for Asylum Seekers in different Member States. [Available at: https://emnbelgium.be/sites/default/files/publications/emn_inform_reception_facilities_3_2.pdf].

Finney, N. and Robinson, V. (2008) 'Local press, dispersal and community in the construction of asylum debates', Social & Cultural Geography, Volume 9, Issue 4 [Available at: http://www.tandfonline.com/doi/pdf/10.1080/14649360802077038].

Friedman, E. and Klein, R. (2008) Reluctant Refuge: The story of asylum in Britain. London: The British Library.

Gomatam, S. et al. (2001) 'An empirical comparison of record linkage procedures', Statistics in medicine 21(10): 1485-1496.

Goodson, L. and Phillimore, J. with Black, J., Jones, P., Lutz, J., Tice, A., Williams, J. and Decanntan, C. (2005) New migrant communities: education, training, employment and integration matters. Report for Learning and Skills Council, Black Country [Available at: http://www.download.bham.ac.uk/curs/pdf/new_migrant_communities.pdf].

Greenslade, R. (2005) 'Seeking Scapegoats: the coverage of asylum in the UK Press', Asylum and Migration Working Paper 5, London: Institute for Public Policy Research.

Harron, K., Dibben, C., Boyd, J., Hjern, A., Azimaee, M., Barreto, M. L. and Goldstein, H. (2017) 'Challenges in administrative data linkage for research' in Big Data & Society, 4(2): 1-12.

Herzog, T. N., Scheuren, F. J. and Winkler. W. E. (2007) Data Quality and Record Linkage Techniques. Washington DC: Springer.

HM Government (2012) Open Data White Paper: Unleashing the Potential [Available at: https://data.gov.uk/sites/default/files/Open_data_White_Paper.pdf].

Home Office (1998), Fairer, Faster and Firmer – A Modern Approach to Immigration and Asylum (white paper), London: The Stationary Office [Available at: https://www.gov.uk/government/uploads/system/uploads/attachment_data/file/264150/4018.pdf]

Home Office (2002) Secure Borders, Safe Haven: Integration with Diversity in Modern Britain (white paper), London: The Stationary Office.

Home Office (2005) Integration Matters: A National Strategy for Refugee Integration. IND Corporate Communications [Available at: http://www.employabilityforum.co.uk/documents/Integration%20Matters%20Strategy.pdf].

Home Office (2010) Spotlight on Refugee Integration: Findings from the Survey of New Refugees in the United Kingdom, Research Report 37.

Home Office (2010a) Summary of the Survey of New Refugees December 2005-March 2009, Research Report 35.

Home Office (2010b) Helping New Refugees Integrate into the UK: Baseline Data Analysis from the Survey of New Refugees, Research Report 36.

Home Office (2010c) Spotlight on Refugee Integration: Findings from the Survey of New Refugees in the United Kingdom, Research Report 37.

Home Office (2010d) Technical notes: Survey of New Refugees [Available at: http://doc.ukdataservice.ac.uk/doc/6556/mrdoc/pdf/6556_technical_and_data_notes.pdf].

Home Office, UK Border Agency, Analysis, Research and Knowledge Management. (2010e) Survey of New Refugees, 2005-2009. [data collection]. UK Data Service. SN: 6556, http://doi.org/10.5255/UKDA-SN-6556-1.

Home Office (2011) Asylum Data Tables Immigration Statistics July to September 2011 volume 3 [Available at: https://www.gov.uk/government/statistics/tables-immigration-statistics-july-to-september-2011].

Home Office (2014) The reason for migration and labour market characteristics of UK residents born abroad, Occasional Paper 110 [Available at: https://www.gov.uk/government/uploads/system/uploads/attachment_data/file/350927/occ110.pdf].

Home Office (2014b) Asylum Data Tables Immigration Statistics July to September 2014 volume 1 [Available at: https://www.gov.uk/government/statistics/immigration-statistics-july-to-september-2014-data-tables].

Home Office (2015) User Guide to Home Office Immigration Statistics [Available at: https://www.gov.uk/government/uploads/system/uploads/attachment_data/file/455513/user-guide-immigration-statistics.pdf].

Home Office (2015b) Asylum Data Tables Immigration Statistics April to June 2015 volume 4 [Available at: https://www.gov.uk/government/statistics/immigration-statistics-april-to-june-2015].

Home Office (2018) Asylum Data Tables Immigration Statistics Year ending September 2018 volume 1 [Available at: https://www.gov.uk/government/statistics/immigration-statistics-year-ending-september-2018-data-tables].

House of Commons Committee of Public Accounts (2014) COMPASS: Provision of Asylum Accommodation, Fifty-fourth Report of Session 2013-14, London: The Stationary Office [Available at: http://www.publications.parliament.uk/pa/cm201314/cmselect/cmpubacc/1000/1000.pdf].

House of Commons (HC) Home Affairs Committee (2017) Asylum accommodation: Twelfth Report of Session 2016–17. [Available at: https://publications.parliament.uk/pa/cm201617/cmselect/cmhaff/637/637.pdf].

House of Commons (HC) Home Affairs Committee (2016) The work of the Immigration Directorates (Q3 2015): Sixth Report of Session 2015–16. London: The Stationery Office Limited [Available at: https://publications.parliament.uk/pa/cm201516/cmselect/cmhaff/772/772.pdf].

Hynes, P. (2011) The dispersal and social exclusion of asylum seekers: Between liminality and belonging. Bristol: The Policy Press.

Immigration and Asylum Act 1999 c. 33. [Available at: http://www.legislation.gov.uk/ukpga/1999/33/contents].

Johnson, M. (2003) Asylum seekers in dispersal - healthcare issue Home Office [Available at: https://core.ac.uk/download/files/25/10037.pdf]

Knauth, B. (2011) Migration Statistics Mainstreaming. Paper presented at the 58th ISI Congress, Dublin (Session STS018) [Available at: http://2011.isiproceedings.org/papers/650162.pdf].

Kushner, T. and Knox K. (1999) Refugees in an Age of Genocide, London: Frank Cass.

Lynn, P. (2009) Methodology of Longitudinal Surveys. Chichester: John Wiley and Sons.

Manchester City Council Communities Scrutiny Committee (March 2014) Manchester City Council Report for Information, Item 7 [Available at: http://www.manchester.gov.uk/meetings/committee/81/communities_scrutiny_committee].

Manly, B. and Navarro Alberto, J. A (2017) Multivariate Statistical Methods: A Primer, Fourth Edition. Florida: CRC Press.

Molenberghs, G. and Verbeke, G. (2005) Marginal versus random-effect models, In Models for Discrete Longitudinal Data, New York: Springer.

National Audit Office (NAO) (2014) COMPASS contracts for the provision of accommodation for asylum seekers London: The Stationery Office [https://www.nao.org.uk/wp-content/uploads/2014/01/10287-001-accommodation-for-asylum-seekers-Book.pdf]

Nationality, Immigration and Asylum Act 2002 c. 41. [Available at: http://www.legislation.gov.uk/ukpga/2002/41/contents].

Norusis, Marija. (2011) IBM SPSS Statistics 19 Statistical Procedures Companion, Pearson Education: Upper Saddle River, N.J.

Nowok, B., and Willekens, F. (2011) 'A Probabilistic Framework for Harmonisation of Migration Statistics', Population Space and Place, 17(5), 521-533.

OECD (2016), Making Integration Work: Refugees and others in need of protection, OECD Publishing, Paris. [Available at: http://dx.doi.org/10.1787/9789264251236-en].

ONS (2010), Mid-1981 to 2010 Population Estimates: Population density in England and Wales; estimated resident population. [Available at: http://www.ons.gov.uk/ons/search/index.html?newquery=71currmye&newoffset=50&pageSize=50&sortBy=&sortDirection=DESCENDING&applyFilters=true.

ONS (2010a) Mid-2005 Population Estimates: Single year of age and sex for local authorities in the United Kingdom; estimated resident population.

ONS (2010b) Mid-2008 Population Estimates: Single year of age and sex for local authorities in the United Kingdom; estimated resident population.

ONS (2016) The 2021 Census: Assessment of initial user requirements on content for England and Wales. Migration and citizenship topic report for the ONS Census Transformation Programme.

ONS (2016b) Labour Force Survey User Guide Volume 1 – LFS Background and Methodology 2016 [Available at: http://doc.ukdataservice.ac.uk/doc/6727/mrdoc/pdf/lfs_vol1_background2016.pdf].

ONS (2018) Social Survey Division, Northern Ireland Statistics and Research Agency. Central Survey Unit. Quarterly Labour Force Survey, 1992-2017: Secure Access. [data collection]. 12th Edition. UK Data Service. SN: 6727, http://doi.org/10.5255/UKDA-SN-6727-13.

Phillimore, J. and Goodson, L. (2006) 'Problem or Opportunity? Asylum Seekers, Refugees, Employment and Social Exclusion in Deprived Urban Areas', Urban Studies 43.10: 1-22.

Phillimore, J. and Goodson, L. (2008), 'Making a Place in the Global City - The relevance of indicators of integration', Journal of Refugee Studies, 21: 305-325.

Phillimore, J., Goodson, L. and Oosthuizen, R. (2003) Asylum seekers and refugees: education, training, employment, skills and services in Coventry and Warwickshire. Report for Learning and Skills Council Coventry and Warwickshire.

Phillimore, J., Goodson, L., Beebejaun, Y. and Ferrari, E. (2004) The access, learning and employment needs of newcomers from abroad and the capacity of existing provision to meet those needs. Report for Learning and Skills Council, Birmingham.

Politowski, B. and McGuinness, T. (2016) Research Briefing: Policy on the dispersal of asylum seekers. House of Commons Library [Available at: http://researchbriefings.parliament.uk/ResearchBriefing/Summary/CDP-2016-0095#_ftnref9].

Raab, G., Nowok, B. and Dibben, C. (2016) 'Practical data synthesis for large samples' in Journal of Privacy and Confidentiality, 7(3).

Rabe- Hesketh, S. and Everitt, B. (2004) A handbook of Statistical Analyses using Stata, third edition.

Rassler, S. (2004). Data Fusion: Identification Problems, Validity, and Multiple Imputation. Austrian Journal of Statistics, 33:153–171.

Raymer, J., Abel, G. and Smith, P. W. F. (2007) 'Combining census and registration data to estimate detailed elderly migration flows in England and Wales', Journal of the Royal Statistical Society A 170: 891-908.

Right Care (2015) The NHS Atlas of Variation in Healthcare: Reducing unwarranted variation to increase value and improve quality, Public Health England [Available at: http://www.rightcare.nhs.uk/atlas/RC_nhsAtlas3_HIGH_150915.pdf]

Robinson,V., Andersson, R. and Musterd, S. (2003) Spreading the 'Burden'? A Review of Policies to DisperseAasylum Seekers and Refugees, Bristol: The Policy Press.

Schuster, L. (2003) The Use and Abuse of Political Asylum in Britain and Germany, London: Frank Cass Publishers.

Smith, P., Cleary, A., Jones, M., Johnston, S., and Bremner, P. (Ipsos-Mori), Brown, J. and Wiggins, R. (Institute of Education) (2011) A feasibility study for a survey of migrants. Home Office Occasional Paper 92 [Available at: https://www.gov.uk/government/uploads/system/uploads/attachment_data/file/115908/occ92.pdf].

Somerville, W. (2007) Immigration under New Labour, Bristol: The Policy Press.

Stewart, E. (2004) 'Deficiencies in UK Asylum Data: Practical and Theoretical Challenges', Journal of Refugee Studies 17 (1): 29-49.

Stewart, E. (2011) 'UK Dispersal Policy and Onward Migration: Mapping the Current State of Knowledge', Journal of Refugee Studies.

Stuart, E. A. (2010) Matching Methods for Causal Inference: A Review and a Look Forward. Statist. Sci. 25(1): 1-21.

Thomsen, I., and Zhang, L-C. (2001) 'The effects of using administrative registers in economic short term statistics: the Norwegian labour force survey as a case study', Journal of Official Statistics, 17(2), 285-294.

Townsend, P. (1987) 'Deprivation'. Journal of Social Policy, 16(2): 125-146.

Twisk, J. W. R. (2003) Applied Longitudinal Data Analysis for Epidemiology: A Practical Guide. Cambridge: Cambridge University Press.

UNHCR (2010) Convention and Protocol relating to the Status of Refugees, Geneva [Available at http://www.unhcr.org/3b66c2aa10.html].

Watson, N. and Wooden, M. 'Identifying factors affecting longitudinal survey response' in Lynn, P. (2009) Methodology of Longitudinal Surveys. Chichester: John Wiley and Sons.

Yildiz, D. and Smith, P. W. F. (2015) 'Models for combining aggregate-level administrative data in the absence of a traditional census', Journal of Official Statistics, 31(3): 431-451.

Yorkshire and Humberside Consortium for asylum seekers and Refugees (2003) Integration Strategy [Available at: https://www.migrationyorkshire.org.uk/userfiles/file/AboutUs/Integration/RegIntegrationStrat2003-06.pdf].

Zetter, R., Griffiths, D. and Sigona, N. (2005) 'Social capital or social exclusion? The impact of asylum-seeker dispersal on UK refugee community organizations.' Community Development Journal, 40.2: 169-181. [Available at: https://www.researchgate.net/profile/Roger_Zetter/publication/31263741_Social_capital_or_social_exclusion_The_impact_of_asylum-seeker_dispersal_on_UK_refugee_community_organizations/links/55f15ad608ae0af8ee1d5b6f.pdf].

Zetter, R., Griffiths, D., Sigona, N. and Hauser M. (2002) Survey on Policy and Practice related to Refugee Integration. Oxford Brookes University/European Commission [Available at: https://www.academia.edu/368863/A_Survey_of_Policy_and_Practice_Related_to_Refugee_Integration].

Zhang, L-C. (2015) 'On Proxy Variables and Categorical Data Fusion', Journal of Official Statistics, 31(4): 783-807.