

A Feasibility Report on Local Area Short-term Immigration Estimation

November 2008

1. Introduction

1.1 Short-term migration is broadly defined as moves made for more than a few weeks but less than twelve months. ONS have been developing short-term migration estimates since 2006 as part of the Improving Migration and Population Statistics (IMPS) project. This has included:

- An initial feasibility report at the beginning of the project outlining a vision for what might be produced, discussion of how to define short-term migration, and assessment of the available sources and methods.
- A user consultation to the initial feasibility report, specifically on the most appropriate definition to be used.
- Publication of experimental estimates at England & Wales level.
- A reconciliation exercise between IPS based migration estimates (both long and short-term) and counts from administrative sources.

1.2. There are two priorities in the further development of the short-term migration product. Firstly, there is the production of estimates at local area level. Temporary visitors to the UK are unlikely to be spread evenly across the country and they will not necessarily be concentrated in the same areas as long-term migrants. Secondly, ONS is keen to improve the timeliness of the estimates and to further the research work aimed at adapting the timing of their release to be more consistent with other migration products.

1.3 This report sets out progress to date on investigating the feasibility of producing local authority level estimates of short-term migration. Specifically the paper covers (i) the definition of short-term migration to be used at local area level, (ii) local area immigration distributions from administrative sources, and (iii) the potential use of modelling methods. The use of alternative basic distribution approaches is also considered in an appendix to the main report. The paper concludes by outlining the next steps for the work including how ONS plan to engage with users ahead of planned publication in summer 2009.

2. Definitions

2.1 It is proposed that for local area estimation ONS define short-term migration as any visit made for between one and twelve months. The choice of this definition is consistent with local authority user feedback received following the initial feasibility report.

2.2 Consideration was also given to the sample size of the International Passenger Survey (IPS). Without sufficient sample size it would not be possible to establish, at local area level, which variables to use in any model based estimation. This is discussed further in section four. Table 1 demonstrates that the total number of immigrants sampled under the UN definition¹ is only 10% of the total sampled using a definition of all individuals who stay for between one and twelve months. The sample size is relatively small under the UN definition and these estimates are only sufficiently robust to publish at England & Wales level.

Table 1 – IPS Sample Size at England & Wales Level, immigration mid-2006

Short-term migrant definition	Immigrant sample (3 to 12 months)	Immigrant sample (1 to 12 months)
Employment	101	197
Study	111	309
UN Definition (Employment or Study)	212	506
Other Reasons ²	287	1,636
All Reasons	499	2,142

2.3 Local users also expressed interest in being able to separately identify temporary foreign workers. This interest was mainly from London Boroughs and rural authorities with seasonal

¹ The United Nations (UN) defines short-term migration as visits made for work or study for between three and twelve months.

² Other Reasons include visiting friends or family, business, holiday and joining/accompanying

agricultural workers. Given the small sample size of temporary foreign workers shown in Table 1 it will not be possible to produce estimates at local area level using the approaches described in this paper.

2.4 Visits made for less than one month are excluded as previous research has shown that the number of such moves is extremely large, the majority of them being made for very short periods for reasons such as holidays, short business trips or visiting relatives. In the year to mid-2006, moves for less than a month made up 95% of all visits to England & Wales lasting less than 12 months.

2.5 Local area level short-term migration estimates will only be produced for the in-flow of overseas residents to England & Wales. At present, no suitable sources of information on short-term emigration of British residents have been identified that would allow local area estimates to be produced.

2.6 ONS have already published experimental short-term migration estimates at England & Wales level using a range of definitions including the UN definition, all visits for between one and twelve months regardless of reason for visit and separate estimates of short-term migrant workers.

3. Distributions in Administrative Sources

3.1 The comprehensive counts provided by administrative data sources can be produced at any geographic level and so are of great potential use for providing local distributions of those registered. However, there is no administrative source set up expressly for the purpose of international migration measurement. Often when migrants do appear on an administrative source they are included regardless of how long they stay. As there is generally no incentive to de-register with these sources, it is not possible to distinguish between long and short-term migrants. Use of these sources for estimating local distributions is therefore not straightforward, further research into how to use them is required.

3.2 This section reviews local area distributions from administrative sources known to include short-term migrants. Specifically, distributions from the Worker Registration Scheme (WRS) and National Insurance Number (NINo) allocations to overseas nationals are considered. Though the WRS only covers migrant workers³ from the A8 countries that joined the European Union in 2004, it is of particular interest as migrants are asked how long they

³ Self-employed migrant workers are not required to register with the WRS.

intended to stay⁴. As NINo applicants are not asked how long they intend to stay, it is not possible to distinguish between long and short-term migrants in the same way. Comparing the distribution of A8 migrant workers from NINo allocations to the WRS distribution should indicate the extent to which the inclusion of long-term migrants (in the former source) biases the distribution of short-term migrants⁵. The distribution of NINos allocated to all overseas nationals is also provided. However, as this covers migrant workers from all over the world it is difficult to conclude whether the patterns shown are short-term migrant specific or not.

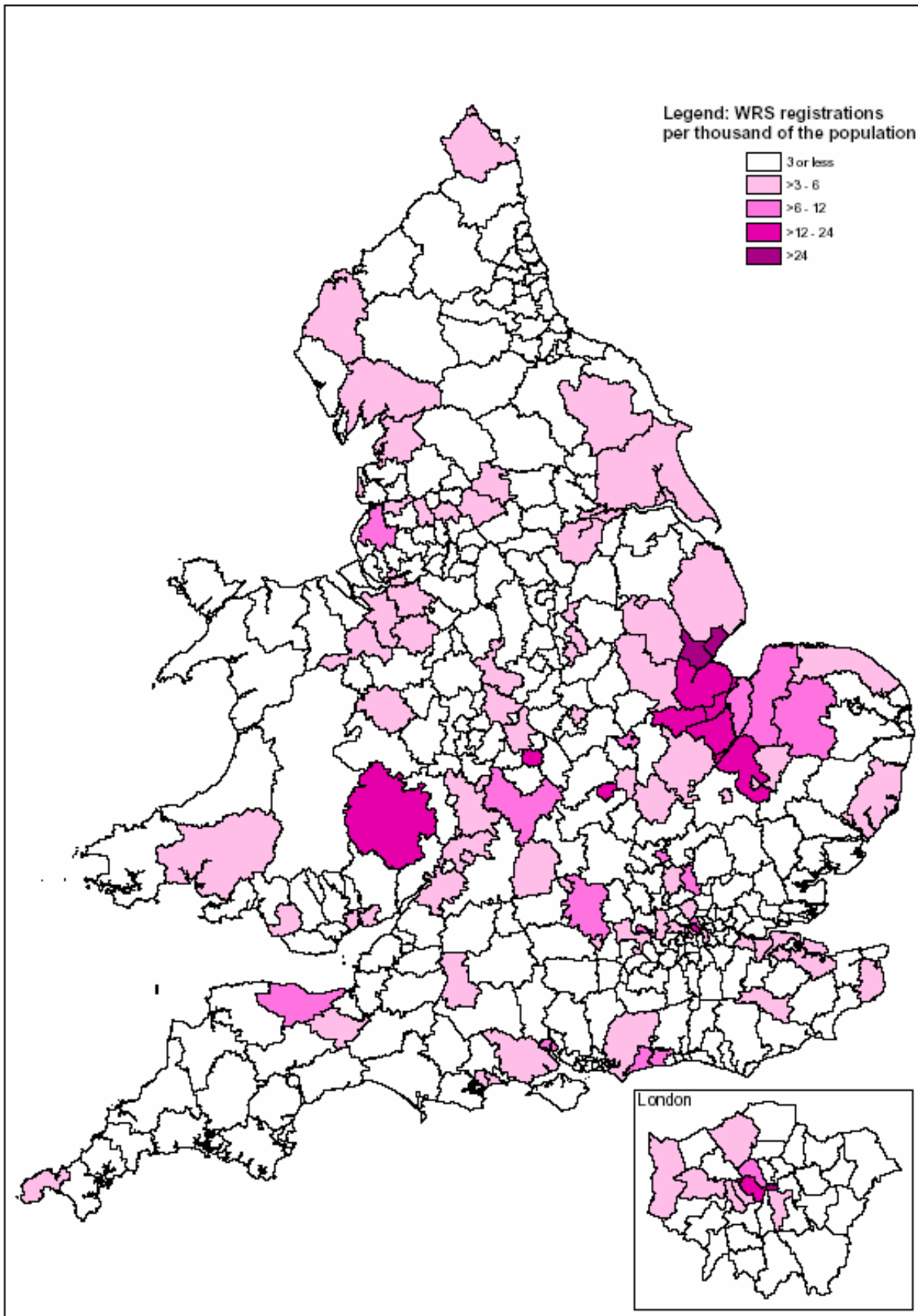
3.3 The local area distribution of immigrants based on patient registration data is not explored in this paper. Though short-term migrants do register with GPs, the coverage of the patient registers is much wider than migrant workers. As a result it is difficult to relate to the WRS, the only source where we are able to separately identify short-term migrants. Patient registration data will however be entered for potential selection in the planned local area modelling of short-term migrants here for one to 12 months.

3.4 Figure 1 below shows the local authority distribution of A8 short-term migrants identified on the WRS between May 2004 and March 2006. It shows that, relative to other local authorities, Boston and City of London had the highest concentration of short-term WRS registrations (over 24 per thousand of the population). Other local authorities with high concentrations (12 to 24 per thousand of the population) include: Coventry, County of Herefordshire, East Cambridgeshire, Fenland, Northampton, Peterborough, South Holland and Westminster. Concentrations in rural areas are of particular interest as generally international long-term migrants are more attracted to cities and towns.

⁴ Intentions data are used as an indicator only. A migrant's intentions are not always realised, they may leave before they intended or stay longer. In addition, some WRS respondents stated that they did not know how long they would stay.

⁵ It is noted that NINos are allocated to those seeking to claim benefits as well as those intending to work.

Figure 1 – WRS registrations of short-term migrants, per thousand of population, 2004-2006



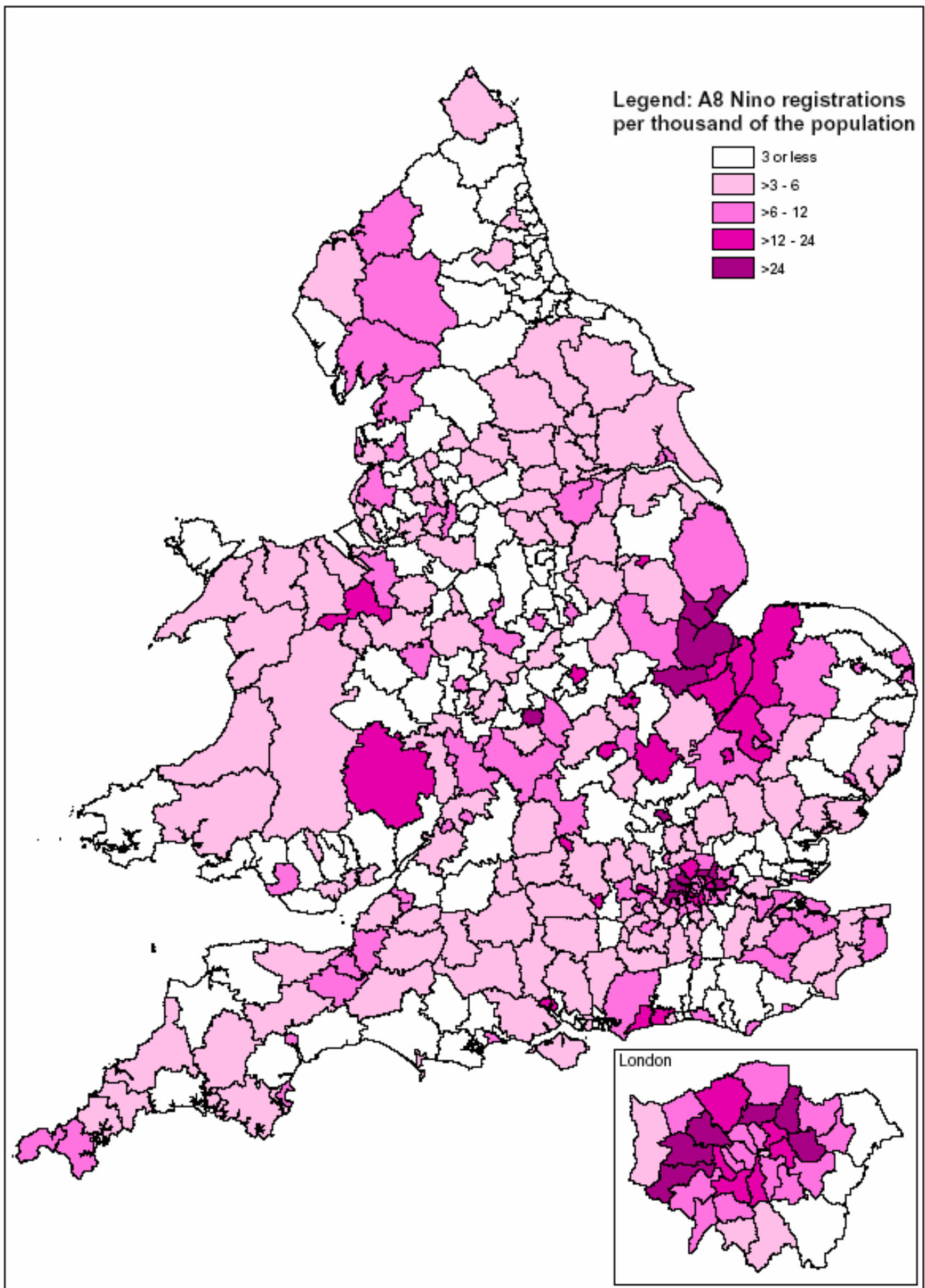
3.5 The distribution of NINOs allocated to A8 nationals in figure 2 shows many of the same concentrations as in figure 1 (again there are concentrations in rural areas in the East and East Midlands) but there are important differences. Some LAs identified by the WRS are not present in the NINOs such as North Norfolk and City of London. This is likely to be because the WRS is based on the address of the employer whereas NINO allocations are based on the migrant's residential address. In Westminster and City of London this is likely to be a key issue as many migrant workers, like the resident population, will work in the borough but live elsewhere.

3.6 There are also many areas with concentrations identified by NINOs that are not in the WRS, including the London Boroughs of Brent, Ealing, Hounslow, Waltham Forest, Newham and Haringey. This is likely to be due to two reasons:

- NINOs include both short and long-term migrants and as it is not possible to identify each group separately it is assumed that areas with higher NINOs contain concentrations of long-term A8 migrants.
- Self-employed migrant workers are not required to register with the WRS, whereas they do require a NINO.

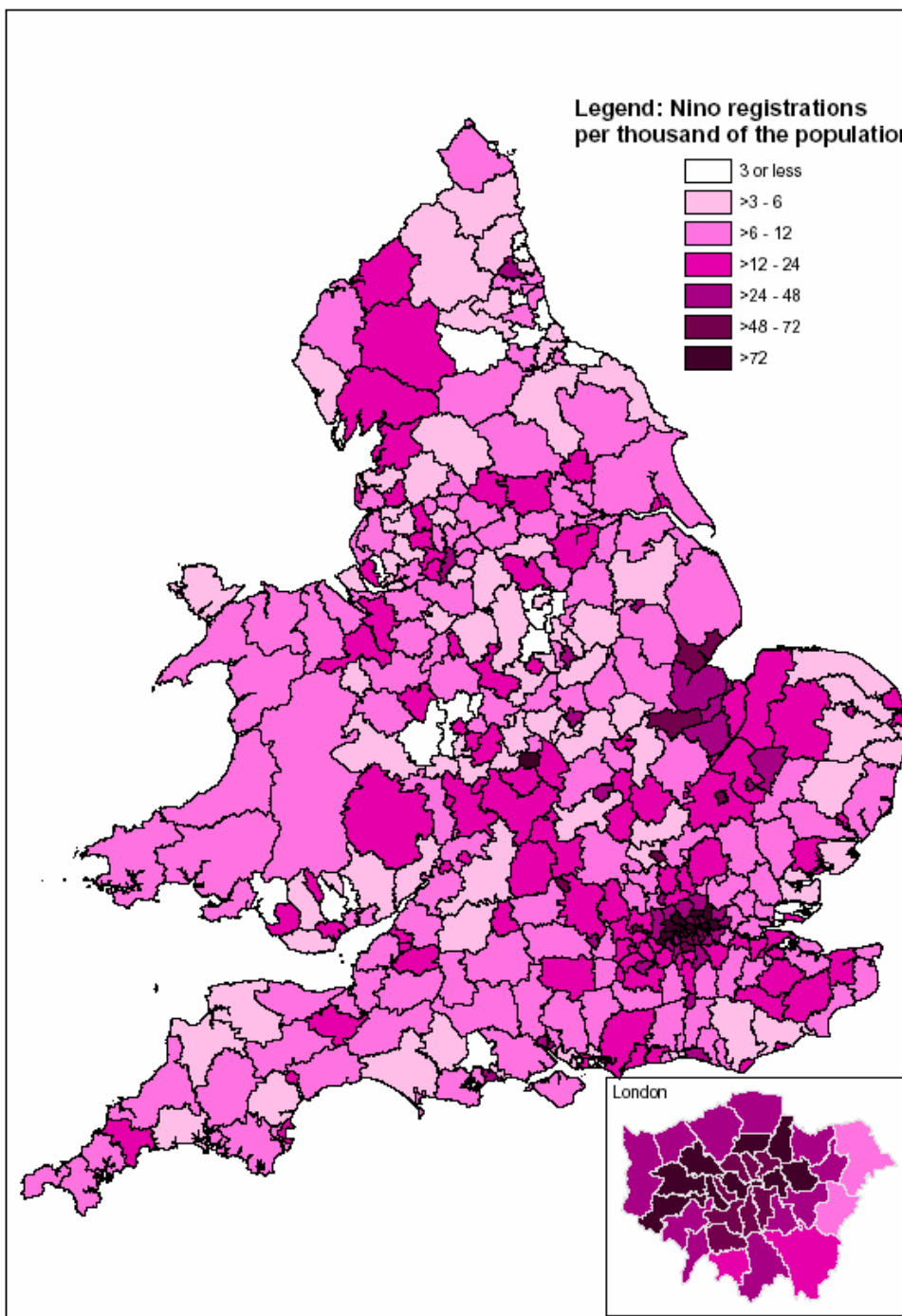
These definitional differences between the sources make it difficult to draw firm conclusions.

Figure 2 – NINo allocations to A8 nationals, per thousand of population, 2004-2006



3.7 A third distribution is presented in figure 3 using NINo allocations to all overseas nationals. When all migrants are included (regardless of where they are from) it is apparent that there is an even greater concentration of migrants to London Boroughs. Although the concentration identified in the other two figures is still apparent in some parts of the East of England, it is difficult to conclude whether concentrations are the result of large numbers of short-term migrants or not.

Figure 3 – NINo allocations to all overseas nationals, per thousand of population, 2004-2006



3.8 This section has highlighted how difficult it is to distinguish between distributions of long and short-term migrants from administrative sources. While it has been possible to identify a distribution of some short-term migrant workers from A8 countries (from the WRS) it is a partial picture and there is no evidence to confirm whether it is representative of all short-term migrants. It becomes increasingly difficult to relate this to a distribution based on NINOs as the latter includes both short and long-term migrants from across the world. Again there are definitional differences which further complicate comparisons.

3.9 An alternative way of assessing administrative data for distributing short-term migrants is through comparisons at a regional level between IPS based estimates and distributions based on administrative sources. This assessment is presented in Annex A. In addition to using WRS, NINo and Patient Registration data, regional estimates are also produced using estimates of long-term migration and population size. It is concluded that although none of these sources give the same distribution as the IPS, distributions based on three of the sources (long-term immigration, NINOs and patient registers) are relatively consistent. Estimates based on these three basic distribution approaches will be used to validate local area estimates based on more complex methods.

4. Modelling Approach

Why use a modelling approach?

4.1 Statistical modelling has already been used to improve the distribution of international migration at local area level. As part of the package of improvements made to migration statistics in 2007, modelling was used to improve the local area distribution of where emigrants lived prior to leaving the UK. ONS is building on this experience and are researching and developing immigration models to improve population and migration statistics. This work is being undertaken on both long and short-term migration.

4.2 The basis of the models, which are regression estimation models, is to identify explanatory variables which have a significant association with a dependent variable, the local area count of immigrants. In addition to identifying the significant variables, the model also identifies the importance of each, controlling for the other variables in the model.

4.3 A key strength of the regression modelling approach is that data from a range of sources can be entered for potential selection. For migration distribution this is particularly useful as although administrative sources provide annual figures of the numbers of migrants at local

area level, no single source is able to identify the number of short-term migrants. Using a single administrative source as the basis for a distribution may be potentially biased.

4.4 Data sharing is central to the programme of improving migration statistics. As access to administrative sources is established, there is scope to refine the migration models being developed by entering new data for potential selection. At present all administrative data entered into these migration models is provided at aggregate level such as NINo and WRS data. Improved data sharing provides an opportunity to develop more sophisticated variables by using the extra information on the characteristics of migrants that can only be obtained from individual record level data. In the future, the data most likely to improve the models is from the Department for Work and Pensions (DWP).

4.5 Previous research has identified a basic model structure for distributing international immigration to local level. ONS now plan to test this structure using explanatory variables specifically relevant to short-term migration. Key characteristics of the model are that it is:

- based on a Poisson distribution
- based on counts of migrants (as opposed to weighted estimates)
- applied directly at local authority level

What data will be modelled?

4.6 International Passenger Survey (IPS) data will be modelled for short-term migration. The IPS data used will be 'completed flow' as is currently used to produce England & Wales level short-term migration estimates. 'Completed flow' refers to data based on individuals sampled by the IPS at the end of their visit. For short-term visitors, sampling individuals at the end of their visit is preferable to sampling at the beginning as the latter is based on intentions which will change for some visitors. Intentions can change both in terms of how long and where people stay. As some long-term migrants may never return to their country of previous residence, it is not possible to use completed flow data for long-term migration modelling.

4.7 An assessment of the sample size of IPS short-term immigration suggests that a model can be developed using a single year's data. The number of short-term immigrants sampled in the year to mid-2006 is shown in table 2. It is noted that the number of short-term immigrants sampled reflects the broad definition used. All visits lasting between one and twelve months are included in this definition, regardless of why people visited.

Table 2 – IPS sample size at different geographic levels, short-term immigration mid-2006

Geographic level	Short-term immigrant sample size
Regional	-
North-East	29
North-West	125
Yorkshire & The Humber	73
East Midlands	89
West Midlands	80
Eastern	150
London	781
South-East	348
South-West	160
Wales	54
(Missing Information)	253
Mean	189
England & Wales	2,142

4.8 Table 2 shows how the number of short-term migrants contacted by the IPS varies across Wales and the regions of England. This reflects the spread of short-term migrants across the country and indicates that, when using this broad definition, London is by far the most common destination for short-term visits with more than a third of all visitors sampled. It is difficult to draw firm conclusions from this however as the sample weights attached to each contact will differ.

4.9 In addition to the geographic distribution provided by table 2, it is noted that a number of records (approximately 12 per cent) are labelled 'missing information'. Location of visit information is missing for two reasons. Firstly, the location of stay question was only introduced for visitors with a definite job to go to in 2008. Short-term migration estimates are the only ONS product using IPS data which include temporary foreign workers hence the recent introduction of the question. The second reason for missing information relates to those who have not fully responded to the survey. Before IPS data can be modelled, a methodology will need to be developed to impute geographic information for these records. This method will make use of geographic information from other IPS contacts where location of stay was collected and IPS data on where short-term migrants intended to stay when they were sampled at the beginning of their visit.

What information will be used in the model?

4.10 The regression model is designed to select the variables that are most important for estimating the number of short-term migrants at local area level. However, it is still necessary to decide which variables to input to the model for possible selection. A number of the variables will reflect factors generally associated with migration. For instance, an employment type variable is likely to be important and other migration indicators such as long-term migration or migration volume rates are also likely to be.

4.11 Analysis of IPS data at England & Wales level has identified the reasons why short-term migration visits are made in greater detail than previously published. It is proposed that the set of potential explanatory variables in the model will be designed to ensure that all of the main reasons for visit are represented. It is these 'other' reasons for visit, such as tourism or visiting friends and family that account for the majority of short-term visits to England & Wales whilst being of little importance when considering long-term migration. The diagram below illustrates the explanatory variables that are proposed for the model.

Figure 4 – Proposed explanatory variables for modelling short-term immigration to local area level

IPS Reason for Visit	Employment	Study	Visiting friends or family	Join or Accompany	Business	Holiday	Other	
Potential Sources:	WRS (for A8 workers)							
	Unemployment estimates							
	Job Centre Plus Vacancies							
	NINo Allocations Data for overseas nationals							
		HESA data						
		Census students data						
	Flag 4 GP Registrations							
	Houses of Multiple Occupation							
			Long-term Emigration					
					Counts from IDBR			
						IPS data		
	Long-term Immigration Estimates							
	Ethnic Group Proportions (Annual Ethnic Population Estimates)							
	Country of Birth Proportions (Census)							

4.12 Figure 4 demonstrates that a range of variables have already been identified for potential selection. As noted previously, a key limitation with administrative data is the inability to distinguish between long and short-term migrants. Without this differentiation there is a risk that long-term migrants could bias the model. However, only those variables which are significant in terms of the short-term migrant count being modelled should be included. Key points to note about the different variables are:

- There are many more sources of data available that cover people who migrate for either employment or study purposes than for other reasons
- People visiting friends or family fall into a number of categories; UK born returning home, overseas residents visiting British nationals or overseas residents visiting foreign nationals resident in the UK. The first of these can be represented using long-term emigration data, the second two via long-term immigration data and ethnicity/country of birth information.
- Persons stating joining or accompanying as their main reason for visit are represented by those they are meeting/travelling with.
- IPS data can be used to identify areas that attract a high number holiday makers. Including tourist visits for any length of stay ensure sufficient sample size at local area level.
- The Inter-Departmental Business Register (IDBR) can be used to identify areas where large businesses/multi-national corporations are located. These are the types of business that short-term business migrants are likely to work for.

5. Next Steps

5.1 ONS will continue to develop the short-term migration models outlined in this paper. The first step in this work will be to finalise the local area level IPS count data which are to be modelled. Following this, the model development stage will draw heavily on the work already carried out on modelling long-term immigration and emigration. Estimates produced as a result of this research are planned for publication in summer 2009. They will not be included in the mid-year population estimates as short-term migrants are not usually resident.

5.2 Input from interested users is clearly important to the development work that will take place between now and the planned publication of these local area statistics. To facilitate this, a small working advisory group will be convened early in 2009 to review the modelling work ahead of any results being published. This group will be similar to the one convened to review previous improvements published in 2007 but will be consulted earlier in the process and able to advise ONS on engaging with the wider user community. A small number of

representatives from local authorities, central government and academia who have a particular interest in this work will be invited to participate.

5.3 Users will also be kept informed of this work through the quarterly Improving Migration and Population Statistics (IMPS) updates, the CLIP population sub-group, and the joint Local Government Association/ONS workshops. In addition a series of presentations are planned for 2009 covering the wider package of improvements the IMPS project plans to introduce.

5.4 In addition to local area estimation, the introduction to this paper also identified timing and timeliness as priorities for the England & Wales level estimates. Research on improving timeliness has indicated that by using provisional (rather than final) IPS data England & Wales level estimates could be published more quickly. The difference between provisional and final IPS data has been shown to be small. England & Wales estimates for the year to mid-2007, based on final IPS data, would have been published in August 2009. ONS now plan to publish these estimates, based on provisional IPS data, in February 2009. If local area estimates can be produced they will also refer to the year to mid-2007.

6. References

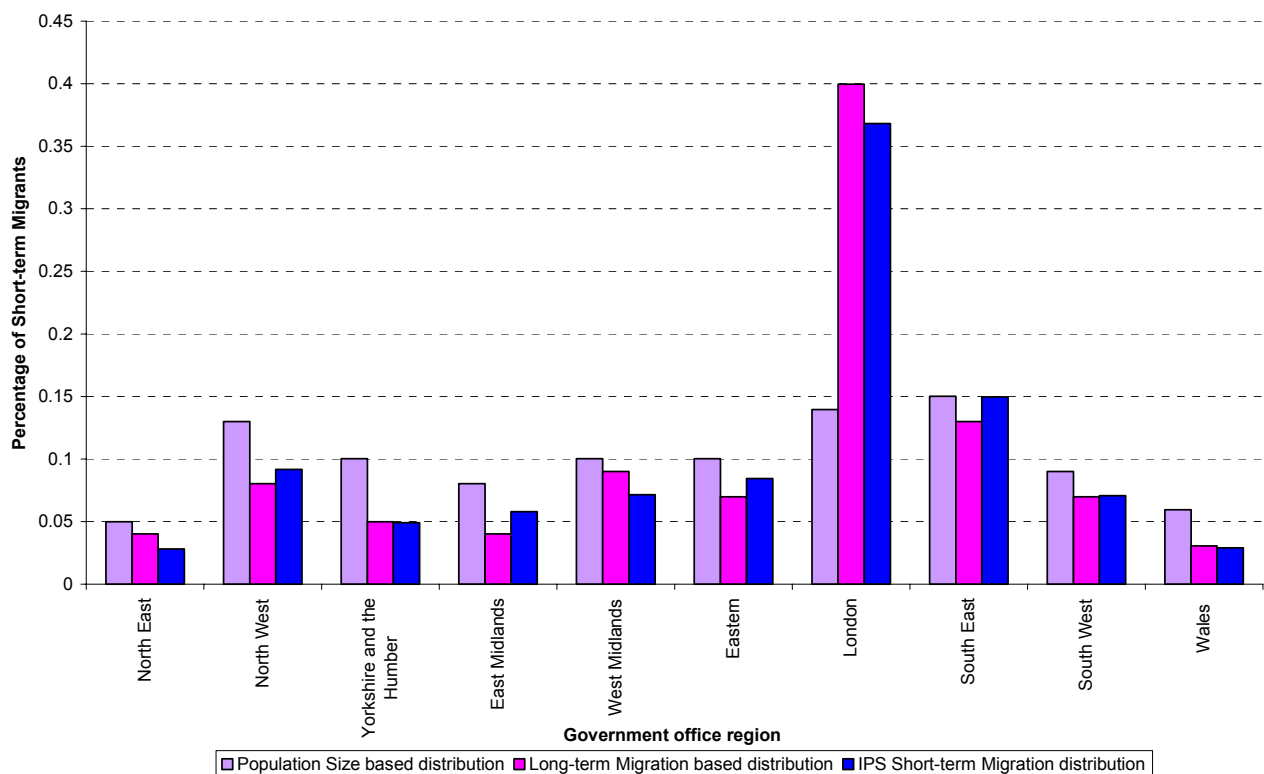
1. Office for National Statistics (2007) *Short-Term Migration Feasibility Report*. Available at: <http://www.statistics.gov.uk/about/data/methodology/specific/population/future/imps/updates/downloads/Short-termMigrationFeasibilityReport.pdf>
2. Office for National Statistics (2007) *Summary of Feedback Received on Short-Term Migration Feasibility Report*. Available at: http://www.statistics.gov.uk/about/data/methodology/specific/population/future/imps/updates/downloads/STM_feedback.pdf
3. Office for National Statistics (2008) Mid-2006 Short-term Migration Estimates for England & Wales. Available at: http://www.statistics.gov.uk/about/data/methodology/specific/population/future/imps/updates/downloads/STM_mid06.pdf
4. Office for National Statistics (2008) A Reconciliation Exercise between IPS Based Estimates of International Immigration and Counts from Administrative Sources, available at: http://www.statistics.gov.uk/about/data/methodology/specific/population/future/imps/updates/downloads/Reconciliation_Exercise.pdf
5. Office for National Statistics (2007) *A Review of the Potential Use of Administrative Sources in the Estimation of Population Statistics*. Available at: <http://www.statistics.gov.uk/about/data/methodology/specific/population/future/imps/updates/downloads/admin.pdf>

Annex A – Basic Approaches to Short-term Migration Distribution

A.1 The aim of this section is to consider whether, at regional level, sources other than the IPS can be used to produce short-term migration estimates sufficiently close to IPS short-term migration estimates⁶. Regional estimates are based on distributing the published England & Wales short-term migration estimate.

A.2 An obvious first source to assess is the distribution of long-term migrants. As with the proposed short-term migration definition, long-term migration estimates are not limited to certain types of migrants such as workers. A second alternative is to distribute on the basis of population size. Regional estimates based on long-term migration and population size are compared against IPS based short-term migration estimates in figure 5.

Figure 5 – Mid-2005 England & Wales short-term immigration distributed to regional level using (i) population size, (ii) long-term immigration and (iii) IPS short-term immigration.



A.3 The long-term immigration based distribution is clearly a closer match to the IPS based distribution of short-term migration than that based on population size. However, there are still clear differences, with higher long-term immigration in London, the North-East and West

⁶ IPS based estimates of short-term at regional level are shown for indicative purposes only. There are known issues with the quality of these estimates. Specifically there are IPS records without location of stay information which have been distributed across the regions.

Midlands and higher short-term immigration in the other regions. These differences are likely to be even greater within individual local authorities, although it is not possible to confirm this by making a comparison at that level.

A.4 Regional estimates were also produced using three key administrative sources (WRS, NINOs and patient registers). These are presented in figures 6 and 7. NINOs and patient registers appear to perform as well as the long-term immigration distribution shown in figure 5. The WRS distribution produces regional estimates which are very different to those based on the IPS.

Figure 6 – Mid-2004-06 England & Wales short-term immigration distributed to regional level using (i) WRS, (ii) NINOs and (iii) IPS short-term immigration

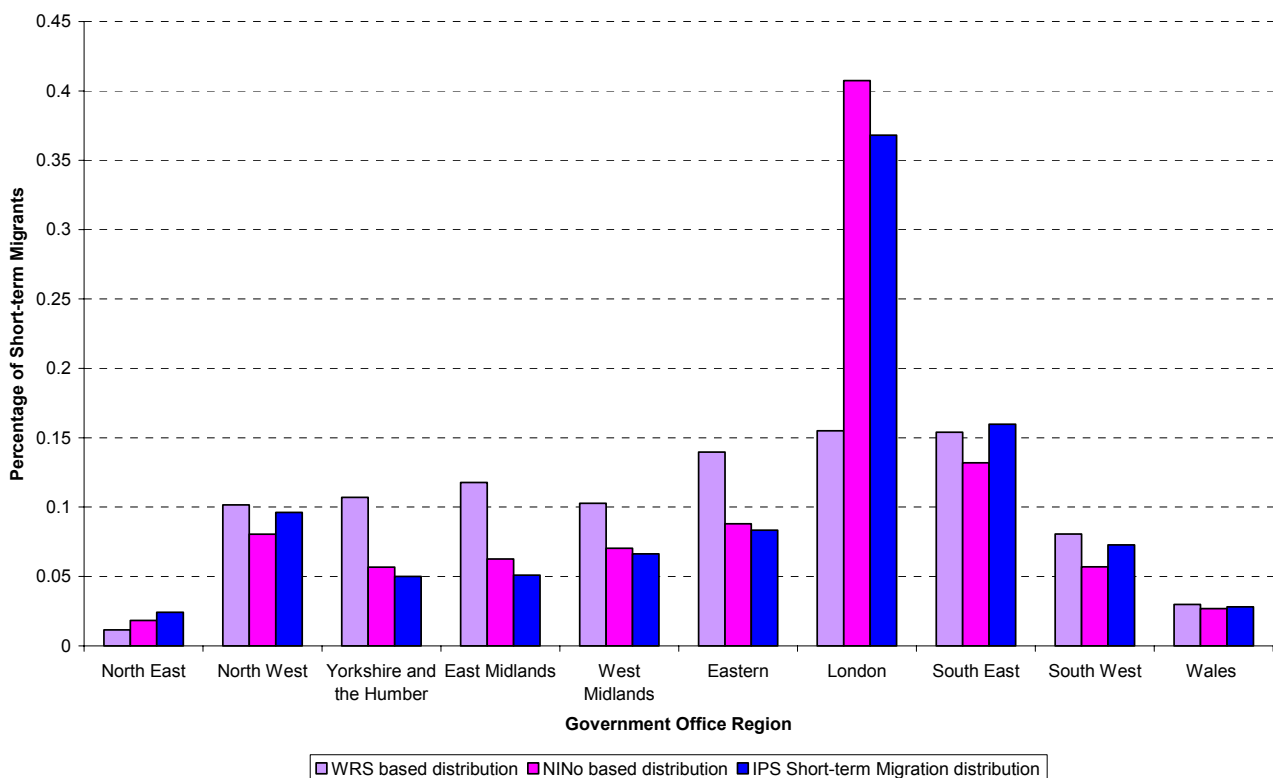
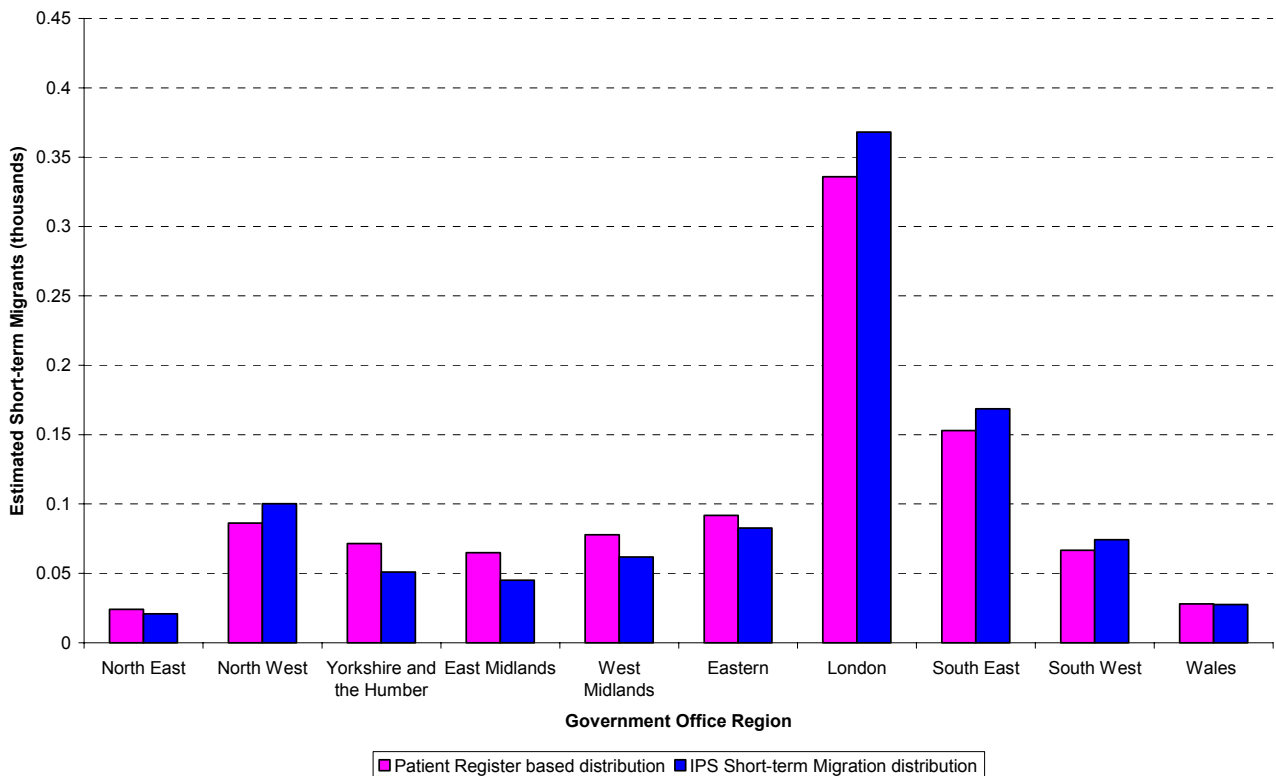


Figure 7 – Mid-2006 England & Wales short-term immigration distributed to regional level using (i) patient registers and (ii) IPS short-term immigration



A.5 Three of the basic distribution approaches (long-term immigration, NINOs and patient registers) considered in this section appear to be of some potential use for distributing short-term migration estimates to regional level. However, it is not possible to repeat these comparisons at local area level as IPS based estimates of short-term migration are not sufficiently robust. Local area estimates based on a number of these sources will be developed for comparison and validation with local level estimates produced using the modelling approach. In addition it is also concluded that these sources should be included in the explanatory variables entered for potential selection by the short-term migration model.