

The Cathie Marsh Centre for Census and Survey Research

A comparison of internal migration data derived from the Pupil Level Annual School Census with the National Health Service Central Register and 2001 Census data

CCSR Working Paper 2009-04

Stephen Jivraj and Naomi Marquis

Stephen.Jivraj@postgrad.manchester.ac.uk,

Naomi.Marquis@postgrad.manchester.ac.uk

Measuring internal migration is central to understanding demographic trends. However, it is more difficult to measure than other components of population change. In the UK, estimates are usually made using National Health Service Central Register (NHSCR) or decennial Census data. This paper will examine the potential of the Pupil Level Annual School Census (PLASC), a relatively new source of internal migration data that can provide more up to date information than the Census and more detailed socioeconomic and geographical information than the NHSCR. The paper provides an empirical comparison of migration data derived from the PLASC with the NHSCR and 2001 Census during the period 2002 to 2007.

A comparison of internal migration data derived from the Pupil Level Annual School Census with the National Health Service Central Register and 2001 Census data

Stephen Jivraj and Naomi Marquis

Cathie Marsh Centre for Census and Survey Research, School of Social Sciences,
University of Manchester, Manchester M13 9PL, UK

stephen.jivraj@postgrad.manchester.ac.uk

naomi.marquis@postgrad.manchester.ac.uk

Abstract

Measuring internal migration is central to understanding demographic trends. However, it is more difficult to measure than other components of population change. In the UK, estimates are usually made using National Health Service Central Register (NHSCR) or decennial Census data. This paper will examine the potential of the Pupil Level Annual School Census (PLASC), a relatively new source of internal migration data that can provide more up to date information than the Census and more detailed socioeconomic and geographical information than the NHSCR. The paper provides an empirical comparison of migration data derived from the PLASC with the NHSCR and 2001 Census during the period 2002 to 2007.

Keywords

Pupil Level Annual School Census, 2001 Census, National Health Service Central Register, internal migration, comparison.

Acknowledgements

The authors would like to thank Mark Brown, Nissa Finney, Ludi Simpson and Nikos Tzavidis for their useful comments on earlier drafts of this paper.

We are also grateful to the Department for Children, Schools and Families for the supply of PLASC datasets, Migration Statistics Unit at the Office for National Statistics for the supply of NHSCR data, and Centre for Interaction Data Estimation and Research for assistance with the 2001 Census Special Migration Statistics.

Stephen Jivraj's research is funded by an ESRC postgraduate studentship. Naomi Marquis's research is funded by the ESRC and Oldham Rochdale Housing Market Renewal Pathfinder as part of a CASE postgraduate studentship. We acknowledge this financial support.

1. Introduction

Measuring internal migration is central to understanding demographic trends. In the UK, it is a more important element of sub-national population redistribution than natural change (Rees et al., 2002). However, it is much more difficult to measure than other components of population change. The two main datasets which have been used to measure internal migration are the decennial Census and the National Health Service Central Register (NHSCR) of all patients registered with a general practitioner (GP). This paper will examine the potential of the Pupil Level Annual School Census (PLASC), a relatively new source of internal migration data that can provide more up to date information than the Census and more detailed socioeconomic and geographical information than the NHSCR.

The paper is structured as follows. Section 2 will give a brief introduction to the PLASC. Section 3 compares the characteristics of migration data which can be derived from the PLASC, NHSCR and decennial Census. Section 4 compares migration data from the PLASC with the NHSCR. Section 5 compares migration data from the PLASC with the 2001 Census. All comparison is based on people of compulsory school age (aged 5 to 15) in each dataset. Section 6 provides conclusions.

2. Pupil Level Annual School Census

The Pupil Level Annual School Census (PLASC) records details of all state-school pupils in England, and is updated annually. The PLASC data source is derived from

an electronic administrative form completed by each school to cover all enrolled pupils in January of each year (Machin et al., 2006). It is collated nationally by the Department for Children School and Families (DCSF) through Local Education Authorities. Completion of the PLASC has been a statutory requirement for all state maintained primary, secondary and special schools since 2002 under section 537A of the Education Act 1996 (Harland and Stillwell, 2007).

The data forms part of the National Pupil Database, which is a data warehouse of education data for key stage performance and information relating to schools and their staff. The PLASC, which provides a link to other data sources, consists of entries for every pupil on roll including information such as home postcode, ethnicity, free school meal eligibility, gender and first language. Through the inclusion of a unique pupil matching reference the data can be matched between years to form a longitudinal source.

Marquis and Jivraj (2009) have assessed the quality of data as supplied by the DCSF and have used appropriate interpolation techniques developed by Harland and Stillwell (2007) to clean errors and omissions. The data is found to be of high quality, though changes over time in the way some data is recorded means that certain variables are more prone to inconsistency than others. The main conclusion regarding the use of the PLASC data for any research purposes is that cleaning is considered essential before analysis (Ewens, 2005, Harland and Stillwell, 2007, Marquis and Jivraj, 2009).

3. Comparison of migration dataset's characteristics

Table 1 provides a comparison of the measure, coverage (population, spatial and temporal) and attributes recorded for the migration data which can be derived from the PLASC, NHSCR and decennial Census.

Table 1 - Characteristics of migration data from PLASC, Census, and National Health Service Central Register

Feature	PLASC	Census	NHSCR
Migration measure	Transitions	Transitions	Moves
Population coverage	Compulsory state school-aged pupils	All people	NHS GP registered patients
Spatial coverage	Postcode to higher	Output areas to higher	Districts to higher
Temporal coverage	Annually from 2002	1980-81, 1990-91, 2000-01	Quarterly from 1975 for Health Authorities; Annually from 1999 for districts
Attributes recorded	Age, sex, language, ethnicity, free school meal status	Age, sex, ethnicity, family status, limiting long term illness, economic activity, socioeconomic class, tenure, household composition	Age and sex

Note: Adapted from Boden, Stillwell and Rees (1992).

In the PLASC, a change of a postcode for a pupil in the matched data can be used to measure migration. This could be between two consecutive years over a longer period for which the data is available providing the pupil is present at both time points. This type of measurement of migration is referred to as a transition because not all movements made will be identified. For example, if a pupil moves more than once during the period between two PLASC collections only one movement will be recorded and if a pupil moves away and then returns to their original location between PLASC collections they will not be recorded as a migrant.

The Census also records migration in this way through a retrospective question which asks whether a respondent had a different usual address one year prior to enumeration. The NHSCR, on the other hand, records all moves an individual makes when they re-register with a GP in a different Health Authority Area¹. An update of the NHSCR is supplied to the Office for National Statistics (ONS) each week and it is combined with data from individual Health Authorities to produce annual migration estimates.

The main limitation of PLASC migration data is that it only includes school-aged children attending state schools. Therefore, the data may not reflect the movement of

¹ Health Authority areas were created by the government in 2002. There were originally 28 across England

all people as families with school-aged children are less likely to migrate than most other households (Bailey and Livingston, 2005, Meen et al., 2005). The data may not even reflect the movements of all families with school-aged children, as there is no information for those in private education. In England, approximately 92% of school-aged children attend a state school, however, this varies between different parts of the country (ONS, 2004). The Census and the NHSCR are not limited to small subsets of the population and, in theory, should include almost all people.

The Census, however, does not provide migration information on people aged less than 1 at the time of enumeration because they would not have had an address one year previously. A greater concern with the Census is data quality. For example, not everyone completes the Census form and, of those that do, the nature of the migration question has led to issues with recall error. In 2001, the problem with non-response was remedied using the One Number Census Methodology (ONS et al., 2001). This process imputed data for both item and unit non-response. The issue with recall error is much more difficult to rectify. The 1991 Census Validation Survey found that about 10% of people who said they did not have a different address one year previously did actually move in the year before enumeration (Rees et al., 2002).

The NHSCR does provide information for people of all ages, however, it only includes people registered with a NHS GP and migration data derived from the data source is dependent on people re-registering with another GP when they move.

The release of the PLASC with a postcode identifier for individual pupils means that data can be aggregated to any higher level geography. This is useful when one is interested in arbitrary geographies or movement between small areas. Small area analysis is also possible with the 2001 Census migration data which was released at output area² at the lowest geographical level. To protect confidentiality, however, ONS have used a procedure called Small Cell Adjustment Mechanism (SCAM) which can lead to data between the same released tables at different geographies not adding to a consistent figure (Duke-Williams and Stillwell, 2007). This can create noise in small area analyses. The NHSCR can only provide migration data between Health

² Output Areas were created using the result of the 2001 Census and contain, on average, 300 people

Authorities at the lowest geography. Nonetheless, through the combination of individual registers, ONS have been able to produce annual migration estimates at local authority district level since 1999 (ONS, 2007).

PLASC data has only been collected since 2002. However, its annual release means that it can provide more up-to-date information than the decennial Census. The release of information for each academic year between 2002 and 2007 provides the necessary information to evaluate the short-term effects of policy intervention, for example. The PLASC has been collected tri-annually since 2006 for secondary schools and since 2007 for primary schools, and as a result future data releases will be able to provide even more detail of the residential movement of school-aged pupils (Harland and Stillwell, 2007). The temporal coverage of the NHSCR is much more complete than both the PLASC and the Census, with migration data between former Health Authority areas being made available since 1975.

The recorded characteristics of migrant pupils in the PLASC include age, gender, ethnicity, first language and free school meal status (an indicator of family socio-economic status). These breakdowns are limited in comparison with the variables available in the 2001 Census for migrants including age, sex, ethnicity, family status, limiting long term illness, economic activity, socioeconomic class, tenure, household composition. All Census variables, however, are recorded at the time of enumeration which makes it difficult to determine, for example, how many migrants were living in rented accommodation one year before the Census. For certain variables, including age and sex it is possible to allocate migrants back to their previous status one year previously. The additional information recorded in the NHSCR is even more limited than the PLASC. Migrants can only be disaggregated by age and sex.

4. Comparison with the National Health Service Central Register

Table 2 compares levels of migration between districts within England in the PLASC and the NHSCR for each consecutive year during the period 2002 to 2007. The time points to which the data relate are not the same for the NHSCR as the PLASC. The NHSCR includes all moves between July of each year whereas the PLASC is collected in January of each year. In Table 2, the difference column shows that the

NHSCR always records a higher number of moves between districts than the PLASC. The ratio between the two datasets indicates that the PLASC accounts for, on average, 75% of the moves in the NHSCR between 2002 and 2007. This is not surprising because the NHSCR records all moves when a patient re-registers with a GP in a different Health Authority. Comparison between the 2001 Census and the NHSCR for the period 2000 to 2001 (not shown here) suggests that the Census accounts for a similar proportion of moves in the NHSCR, 77%.

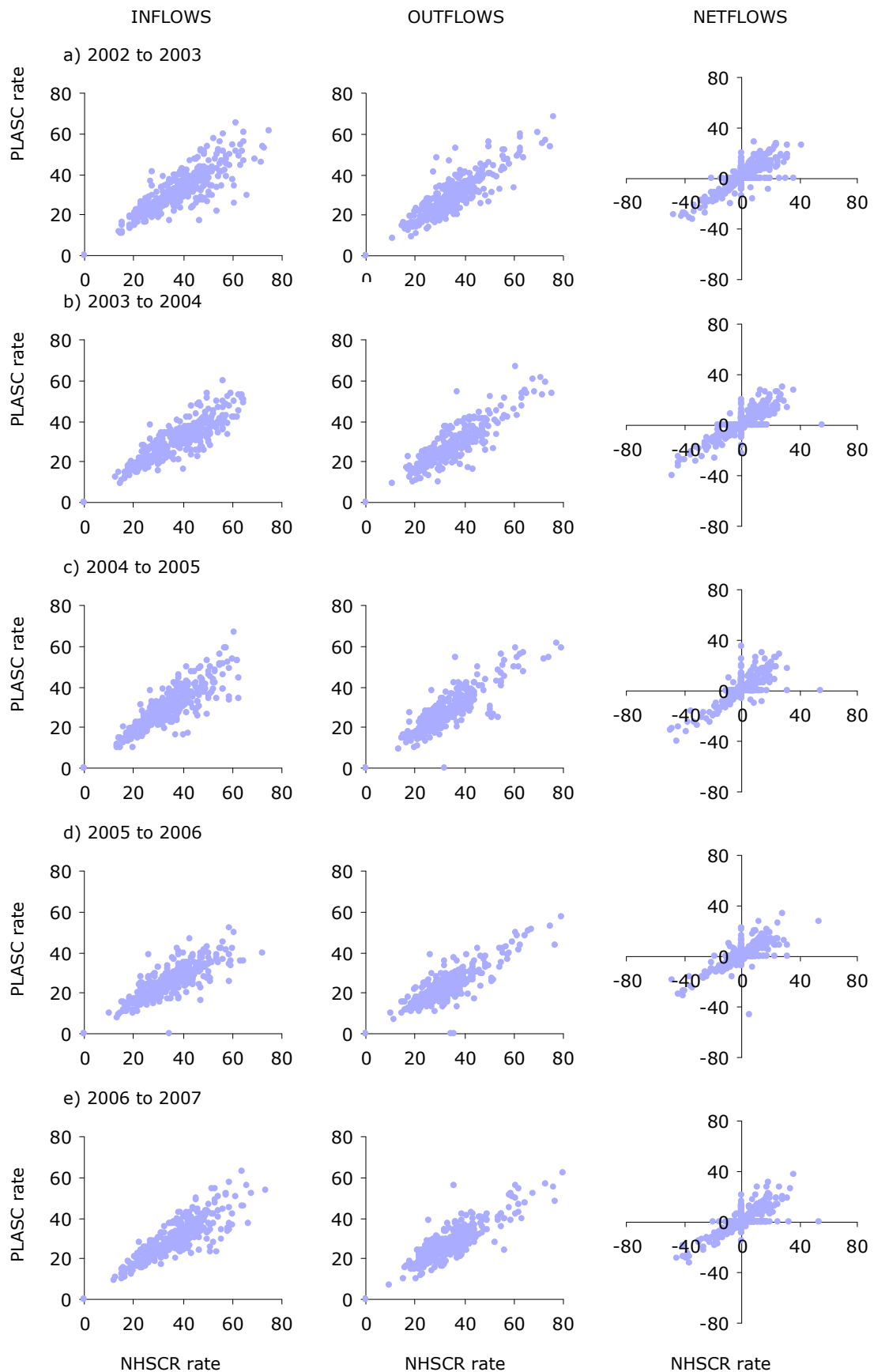
Table 2 - Comparison between level of migration between districts for people aged 5 to 15 in PLASC and NHSCR, 2002-07

Period	PLASC	NHSCR	Difference	Ratio
2002-03	179,498	236,240	-56,742	0.760
2003-04	173,464	233,390	-59,926	0.744
2004-05	173,189	214,800	-41,611	0.807
2005-06	145,605	212,100	-66,495	0.687
2006-07	166,143	218,120	-51,977	0.762

Source: National Health Service Central Register / Patient Register Data System, 2003-07

Figure 1 shows the degree of association between in, out and net migration rates for each individual district in England in the PLASC and NHSCR for each consecutive year during the period 2002 to 2007. The rates are calculated for the NHSCR using mid-year population estimates for people aged 5 to 15 for the corresponding end of year and for the PLASC using compulsory school-aged pupil population at the end of year. The scatterplots indicate a strong positive relationship between the migration rates in the PLASC and NHSCR. The outliers in each scatterplot tend to include selected rural areas, including Richmondshire, North Dorset and Rutland, and London Boroughs. These areas tend to include the lowest proportion of pupils in state schools based on comparison between the PLASC and ONS mid-year estimates for people aged 5 to 15. The outlying rural areas also appear to be areas which contain armed forces bases whose population will be included in the mid-year estimates but not in the NHSCR migration estimates.

Figure 1 – Relationship between inflow, outflow and net flow migration rates for each district in England for 5 to 15 year olds in PLASC and NHSCR, 2002-07



The migration rates are expressed as per 1,000 population aged 5 to 15 in ONS mid-year population estimates for NHSCR and pupils aged 5 to 15 in PLASC
 Source: National Health Service Central Register, 2007; ONS mid-year population estimates, 2007 8

5. Comparison with the 2001 Census

Comparison between migration data available from the 2001 Census and the PLASC is constrained by the different time points for which each dataset is available.

Nonetheless, comparison can be made at a lower spatial scale than between districts, which only account for a quarter of total migration flows in the PLASC. Comparison can also be made between the PLASC and the 2001 Census by age and ethnicity.

NHSCR data is available by age, however, rounding of the data (to the nearest 100 for total outflows and inflows) by ONS has a distorting effect for flows between districts for single year of age.

Table 3 provides a comparison of the distribution of movers recorded in the PLASC and the 2001 Census by type of flow. Data from the PLASC for the period 2002 to 2003 was used as it is the year closest to the 2001 Census. The count of all flows shows that the PLASC recorded a higher number of moves than the Census. This was a result of a much higher level of within-district moves recorded in the PLASC than the Census. As a result, the proportion of within district moves was higher in the PLASC than the Census, however, the proportion of between district moves was higher in the Census than the PLASC.

The higher level of within district moves in the PLASC than the 2001 Census is not unexpected as it is thought that not all people who move will state that they have done so in the Census as a result of recall error (Rees et al., 2002). These moves are most likely to be short distance. It is more likely that short distance moves will be reported in the PLASC, as parents are likely to inform their child's school of a change of address soon after they move house. On the other hand, the difference between migration flows recorded in the PLASC and the Census could in part be due to the different time points over which each dataset was collected and therefore when the migration took place.

Table 3 – Comparison of the distribution of migrants aged 5 to 15 by type of flow from PLASC and 2001 Census

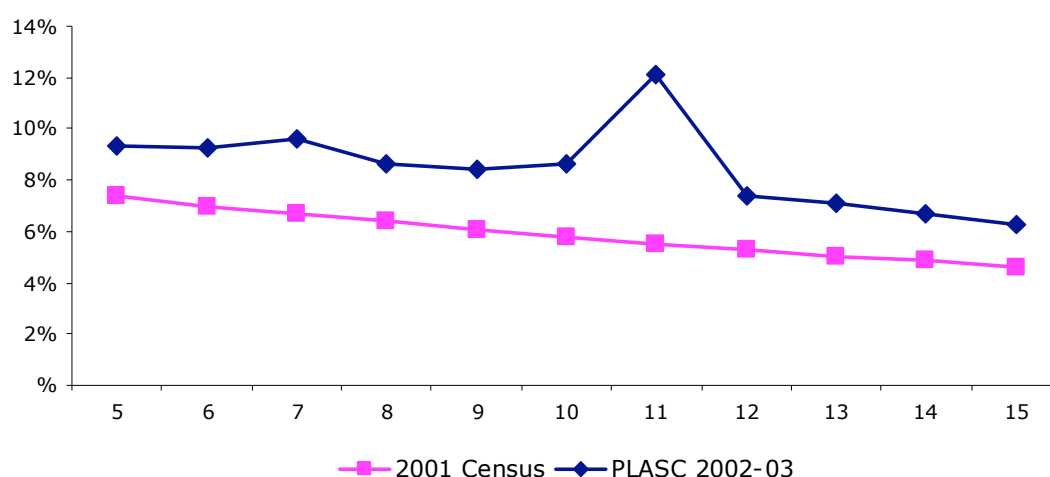
Type of flow	Census (2000-01)		PLASC (2002-03)	
	Count	%	Count	%
Within district	408,881	69.8	544,117	75.2
Between district	177,218	30.2	179,498	24.8
All flows	586,099	100	723,615	100

Note: PLASC data for period 2002 to 2003; Census data for period 2000 to 2001

Source: 2001 Census SMS Table 1

Figure 2 compares the migration rate within districts from the PLASC between 2002 and 2003 and the 2001 Census between 2000 and 2001 by year of age. The rate is, on average, 2.5 percentage points higher in the PLASC than the Census for each single year of age. The trend in the migration rate within districts is similar for the PLASC and the Census for each year of age as it tends to be lower later in a pupil's compulsory school career. In the PLASC, there is a one year rise in the move rate within-districts at age 11 when pupils moves from primary to secondary school. The rise is thought to be a result of movement of pupils in order to obtain a place in a preferred school by parents (Harland and Stillwell, 2007). It is unclear why there is no rise in the rate of within district migration in the 2001 Census migration data.

Figure 2 - Within district migration rates for 5 to 15 year olds from PLASC and 2001 Census by year of age

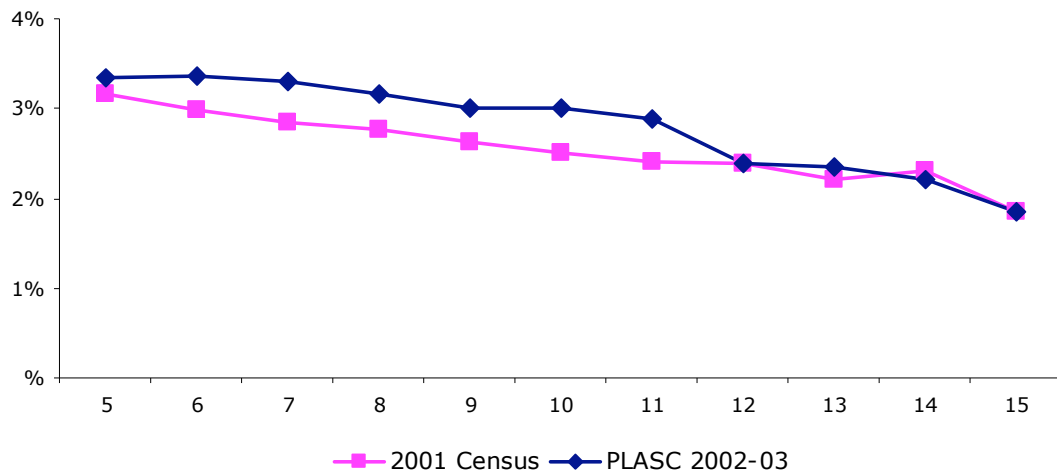


Note: PLASC data for period 2002 to 2003; Census data for period 2000 to 2001

Source: 2001 Census Commissioned Table C0527

Figure 3 compares the single year of age migration rates between districts from the PLASC between 2002 and 2003 and the Census between 2000 and 2001. The rate is broadly the same for each single year of age. There is also a decline in the migration rate between districts, as there is for the within district rate, for older compulsory school aged pupils in both the PLASC and the Census.

Figure 3 - Between district migration rates for 5 to 15 year olds from the PLASC and the 2001 Census by year of age

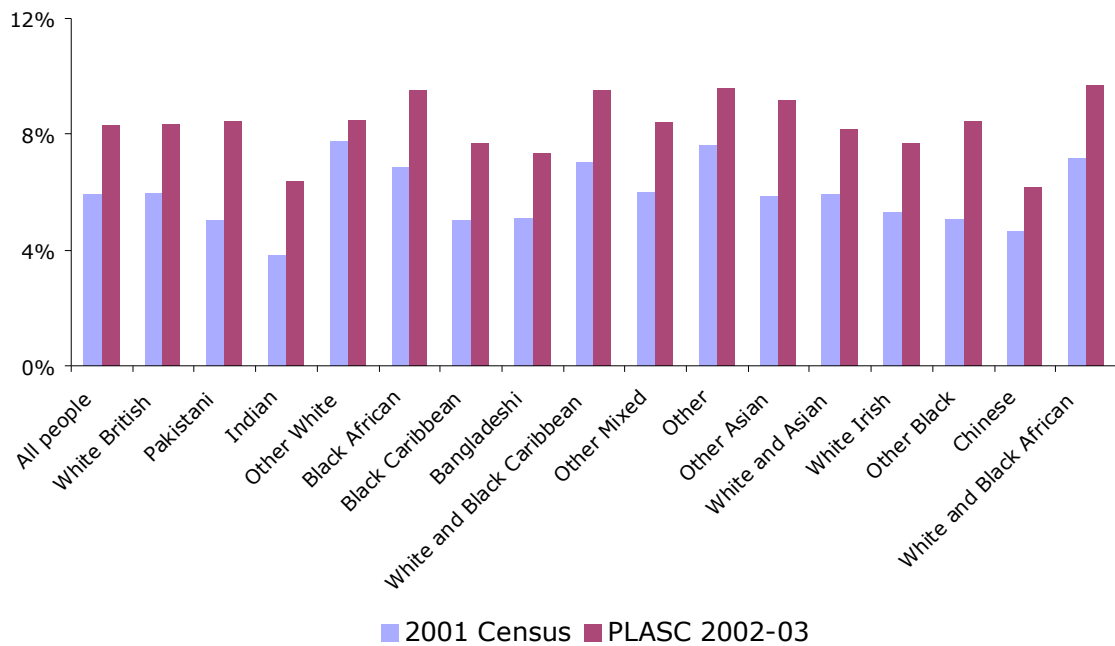


Note: PLASC data for period 2002 to 2003; Census data for period 2000 to 2001

Source: 2001 Census Commissioned Table C0527

Figure 4 provides a comparison of the percentage of people aged 5 to 15 moving within districts in the PLASC between 2002 and 2003 and the Census between 2000 and 2001 by ethnic group. For all ethnic groups the proportion of movers within districts is higher in the PLASC than the Census. This is not surprising as the PLASC tends to record a greater number of within district moves for all 5 to 15 year olds. The distribution of the rate of moves within districts between ethnic groups is similar in both the PLASC and the Census. For example, ethnic groups which have the highest within district move rates in the 2001 Census also have the highest within district move rates in the PLASC. The only group which does not tend to follow the same distribution in both datasets is the Other White group. This is possibly a result of the residual nature of this group rather than any real change in the migration propensity of people aged 5 to 15 classified as Other White. Nonetheless, the difference in the periods for which each dataset was collected cannot be ruled out as a cause.

Figure 4 – Within district migration rates for 5 to 15 year olds from the PLASC and the 2001 Census by ethnic group

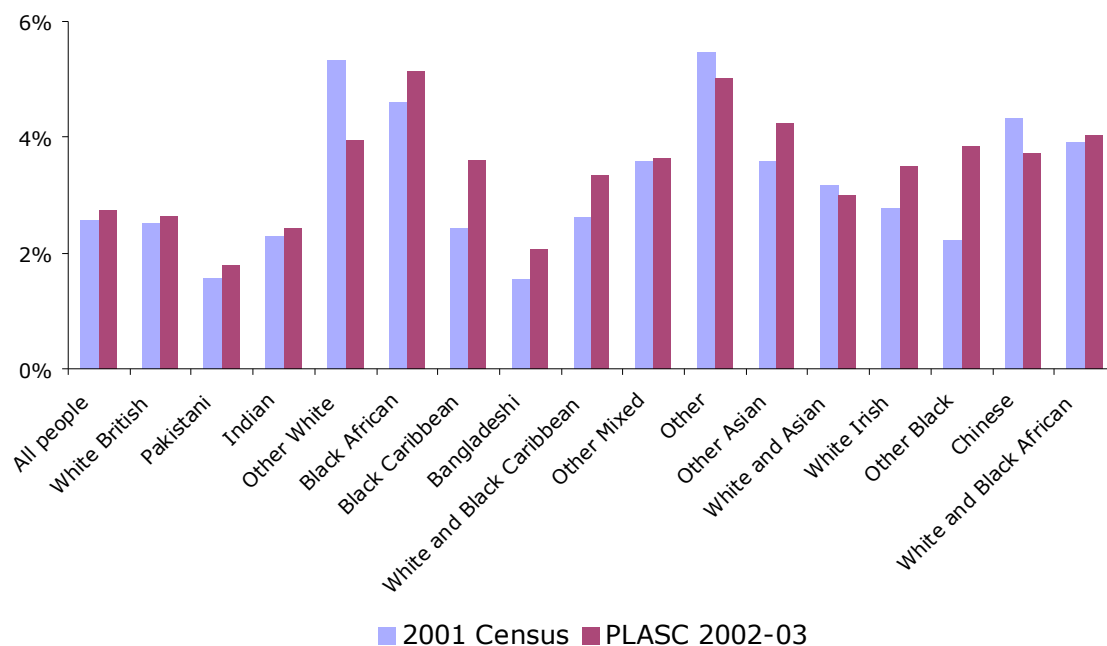


Note: PLASC data for period 2002 to 2003; Census data for period 2000 to 2001; ethnic groups sorted by number of pupils recorded in the PLASC in 2003.

Source: 2001 Census Commissioned Table C0527

Figure 5 provides a comparison of the percentage of people aged 5 to 15 moving between districts in the PLASC between 2002 and 2003 and the Census between 2000 and 2001 by ethnic group. The rate of movement between districts is similar in both the PLASC and the Census for most ethnic groups. The Black Caribbean and Other Black groups are exceptions as these groups tend to be more mobile between districts in the PLASC than the Census. The Other White group also appears to be an exception as this group tends to be more mobile between districts in the Census than the PLASC.

Figure 5 - Between district migration rates for 5 to 15 year olds from the PLASC, 2002-03 and the Census, 2000-01 by ethnic group



Note: PLASC data for period 2002 to 2003; Census data for period 2000 to 2001; ethnic groups sorted by number of pupils recorded in the PLASC in 2003.

Source: 2001 Census Commissioned Table C0527

6. Conclusions

The PLASC certainly has potential to provide new insight into the levels and patterns of internal migration in England. The data can provide a more up-to-date measure of migration than the decennial Census and more detailed information about the socioeconomic characteristics and geographical location of migrants than the NHSCR. The data are, however, limited to the measure of movement of school aged pupils attending state schools. This means that it is difficult to make inferences about the internal migration patterns of the population as a whole because families with children of compulsory school age are less likely to move than any other family group.

Comparison with migration data from the NHSCR and the 2001 Census shows that the migration data derived from the PLASC records similar patterns and trends. Levels of migration between districts are similar in each dataset when accounting for the different way in which migration is measured in the NHSCR. Differences between

the PLASC and NHSCR appear to be due to the population included in each dataset. For example, migration rates are lower than expected in some rural areas and London Boroughs where there are lower proportions of compulsory school aged children in state schools. Furthermore, compared with the 2001 Census, the PLASC appears to record differing levels of within district moves, moves at age 11, and moves for some of the residual ethnic groups. These differences could be due to different time period each dataset was collected.

References

- BAILEY, N. & LIVINGSTON, M. (2005) Determinant of individual migration: an analysis of SARs data. *Scottish Centre for Research on Social Justice Working Paper No. 3.*
- BODEN, P., STILLWELL, J. & REES, P. (1992) How good are the NHSCR data? IN STILLWELL, J., REES, P. & BODEN, P. (Eds.) *Migration Processes and Patterns. Vol. 2 Population Redistribution in the UK.* London, Belhaven.
- DUKE-WILLIAMS, O. & STILLWELL, J. (2007) Investigating the potential effects of small cell adjustment on interaction data from the 2001 Census. *Environment and Planning A*, 39, 1079-1100.
- EWENS, D. (2005) The National and London Pupil Datasets: An introductory briefing for researchers and research users. *Greater London Authority, DMAG Briefing 2005/8.*
- HARLAND, K. & STILLWELL, J. (2007) Using the PLASC data to identify patterns of commuting to school, residential migration and movement between schools in Leeds. *School of Geography, University of Leeds, Working Paper 07/03.*
- MACHIN, S., TELHAJ, S. & WILSON, J. (2006) *The Mobility of English School Children.*

- MARQUIS, N. & JIVRAJ, S. (2009) Preparation of Pupil Level Annual School Census data for the analysis of Internal Migration. *Centre for Census and Survey Research, The University of Manchester, Working Paper 2009-03.*
- MEEN, G., GIBB, K., GOODY, J., MCGRATH, T. & MACKINNON, J. (2005) *Economic segregation in England: Causes, consequences and policy*, Bristol, Policy Press.
- ONS (2004) Pupils and teachers: by type of school, 2002/03. *Regional Trends*, 38.
- ONS (2007) Estimating Internal Migration: Customer Guidance Notes.
- ONS, GRO & NISRA (2001) Census 2001: A Guide to the One Number Census.
- REES, P., THOMAS, F. & DUKEWILLIAMS, O. (2002) Migration data from the census. IN REES, P., MARTIN, D. & WILLIAMSON, P. (Eds.) *The Census Data System*. Chichester, Wiley.