

Migrant Worker Scan: quality assurance of administrative data used in population statistics, Jan 2017

Report detailing the quality assurance undertaken on administrative data for migrant worker scan data used within Population Statistics Division (PSD) publications.

Contact:
Pete Large
pop.info@ons.gsi.gov.uk

Release date:
20 January 2017

Next release:
To be announced

Table of contents

1. [Introduction](#)
2. [Operational context and administrative data collection](#)
3. [Communication with Data Supply Partners](#)
4. [Quality Assurance principles, standards and checks applied by data suppliers](#)
5. [Producer's Quality Assurance investigations and documentation](#)
6. [Conclusion](#)

1 . Introduction

The Migrant Workers Scan (MWS) contains information on all overseas nationals who have registered for and allocated a National Insurance Number (NINo). The data is held by HM Revenue and Customs (HMRC) and is delivered on an annual basis by the Department for Work and Pensions (DWP) to the Office for National Statistics (ONS).

This quality assurance document refers to the dataset delivered to ONS which is used by the Population and Statistics Division (PSD), in their calculation of the [annual mid-year population estimates for England and Wales](#).

This document covers the processes, from data collection through to how the outputs from the MWS are used in the estimation of international immigrants and emigrants at local authority level in England and Wales. It focuses on the quality assurance of the MWS data and identifies potential risks in data quality and accuracy as well as how those risks are mitigated.

This document does not aim to report on the whole of the mid-year population estimate processing or the quality assurance relating to the processing of the other components used in its production. Further information relating to the quality of the mid-year population estimates can be found in [Population Estimates Quality Tools](#) and in the [Annual Mid-Year Population Estimates QMI](#) on our website.

We have assessed the MWS to determine the appropriate level of assurance, using the [Administrative Data Quality Assurance Toolkit](#) provided by the UK Statistics Authority. The toolkit outlines four areas for assurance; the rest of this document will be split into these areas. The areas for assurance are:

- operational context and administrative data collection
- communication with data supply partners
- QA principles, standards and checks applied by data suppliers
- producer's QA investigations and documentation

The MWS has been assigned an A1 rating, meaning that a basic level of assurance is required for these sources. An explanation of why it has been given this rating is given in Section 6 of this report. If you feel that this document does not adequately provide this assurance then please contact pop.info@ons.gov.uk with your concerns.

2 . Operational context and administrative data collection

2.1 Operational Context

2.1.1 What is a NINo and why would you have one?

National Insurance (NI) in the United Kingdom is a system of contributions paid by workers and employers towards the cost of certain state benefits. Contributions are collected by HM Revenue and Customs (HMRC) through the Pay as You Earn (PAYE) system¹. As well as being needed by an employer and HMRC it is also used by the other government organisations, such as the Department for Work and Pensions (DWP), and the Student Loan Company when sorting out benefits and student loans.

In order to administer the National Insurance system, a unique National Insurance number (NINo) is usually sent to children in the UK shortly before their 16th birthday. Overseas nationals resident in the UK or those intending to come to the UK and who also already have the right to work and/or study in the UK can apply for a NINo.

A NINo is needed by any overseas national looking to work or claim benefits or tax credits in the UK. They must be either:

- working, about to start work, or actively seeking employment, and have the right to work in the UK
- be liable to UK Class 1 National Insurance Contributions²
- making a claim to benefit
- be entitled to and wish to pay voluntary Class 3 National Insurance Contributions³
- be referred by the Student Loans Company

In order to be allocated a NINo, an overseas national needs to prove that he/she has the right to work and/or study in the UK. But it is not a prerequisite to working in the UK. There are circumstances when overseas nationals can work within the UK without paying national insurance or without a national insurance number⁴:

- if you can prove that you have the right to work in the UK and have applied for a NINo it is possible to start a job before a NINo is allocated to you
- if you have a document or certificate proving that you pay NI in another European Economic Area (EEA) Country
- if your home country (for example, USA) has a bilateral agreement⁵ on social security with the UK
- if you come from a non-EEA or non-bilateral agreement country and have been sent by your employer in your home country to work in the UK temporarily, then you will not have to pay national insurance for the first 52 weeks in the UK

2.1.2 Legislation relating to NINo registration

The legislation supporting the necessity of foreign workers registering for a NINo is detailed below:

Regulation 9 of the Social Security (Crediting and Treatment of Contributions, and National Insurance Numbers) Regulations 2001 – (Statutory Instrument 2001 No 769).

Someone in the UK who has started work (as an employee or self-employed) must apply for a NINo. People who are entitled to and want to pay Class 3 NI contributions may also apply. This regulation also allows arrangements to be made to allocate NINos to young people in the 12 months before they reach age 16. Since 2006, these regulations also allow that an applicant for a student loan may be directed to apply for a NINo. It also specifies that, in employment inspired applications, the application must be accompanied by right to work documentation.

Regulation 66 of the Social Security (Contributions) Regulations 2001

Every employed earner, in respect of whom any person is liable to pay an earnings-related contribution, shall, on request, supply his NINo to that person.

Regulation 104 of the Social Security (Contributions) Regulations 2001 – Statutory Instrument 2001 No 1004

A self-employed person, who is to be treated as an employee for National Insurance purposes, must tell their employer their NINo so the employer can enter it on the deduction card.

Section 19 of the Social Security Administration (Fraud) Act 1997

This applies to people without a NINo who want to claim benefit from DWP. Section 19 also applies to any other person included in the claim (partner, for example). In all cases the individual must be able to prove that they are who they say they are. If they are unable to do this to DWP's satisfaction, they will not be entitled to claim or receive benefit and so will not need a NINo.

Paragraph 1A of section 13 of the Social Security Administration Act 1992

This applies to people claiming Child Benefit. If someone makes a claim to benefit they must have a NINo or be able to provide sufficient information or evidence for a NINo to be traced, confirmed or allocated to them.

Regulation 5 of Tax Credits (Claims and Notifications) Regulations 2002

This requires a person who is claiming tax credits to provide the Inland Revenue (now HMRC) with a statement of their NINo.

2.2 Administrative Data Collection

2.2.1 Outline and illustrate the administrative data collection process.

Overseas nationals apply for a NINo by making an enquiry to the DWP National Insurance number application line, who will then invite them for an "Evidence of Identity" interview at a specified Jobcentre Plus office. At the interview they must satisfy the criteria for needing a NINo; in particular they must be able to prove their identity and, for employment related applications, that they have a legal right to work in the UK.

The Jobcentre Plus interviewing officer will complete form CA5400 Application for a NINo on behalf of the applicant who will then sign the form. (Applicants may complete the form themselves if they wish.) The form is countersigned then passed to a Jobcentre Plus National Insurance Number Centre (NC) together with copies of the applicant's documents which are required to support the NINo application. A team of specialist trace officers at the NC carries out further checks including whether the applicant already holds a NINo. If a decision is made to allocate a NINo it is allocated on the DWP's departmental Customer Information System (CIS) and then the NC register that National Insurance (NI) account on the HMRC National Insurance and Pay as you Earn System (NPS). Service delivery target is 15 days for a successful applicant to receive their NINo; however it may take 33 days or longer for the process to be completed.

2.3 Implications for accuracy and quality of the data

2.3.1 Main risks to accuracy and quality of data collection

Incorrect recording of the applicant's details, either through typos or inconsistencies between the application form and the evidence of identity, is the largest risk to the quality of the data collection. There is also a risk of an overseas national applying for a NINo when they have been previously allocated one, therefore resulting in 2 separate records for the same person. Fraudulent applications which are not identified as such and are allocated a NINo, also pose a risk to the quality of the MWS dataset: the record may be incomplete or inaccurate on a vital detail used in our processing, such as date of arrival in the UK.

2.3.2 Identified actions taken to minimise risks to quality during data collection

DWP uphold the quality of the data collection process by:

- having a thorough face-to-face interview with the applicant. This enables sensibility checks to be carried out on the applicants' identity and information, for example, is their date of birth plausible for the applicant attending the interview? Do they look like their passport photo?
- having the applicant verbally confirm their details and check the application form before signing it
- having a list of acceptable and robust evidence for proof of identity, visa conditions and employment status

2.3.3 Applying for a National Insurance Number (NINo)

When an overseas national applies for a NINo there are a couple of preconditions that they must meet. These are that they have the right to work and/or study in the UK or they are resident in the UK. The process is started by the applicant contacting the Department of Work and Pensions (DWP) National Insurance number application line. At this stage if it is found that they do not meet the set preconditions they are not given an Evidence of Identity interview and therefore do not receive an NINo.

If they meet the set preconditions they are then invited to an Evidence of Identity interview at their local JobCentre Plus. At this face to face interview the DWP staff complete the application form (CAS5400). The checks include using document examination tools and corroborative checks with third parties (including other government departments) to verify information given. This is then signed by the applicant and any evidence is then photocopied. If at this point the evidence is incomplete, or they do not fully meet the preconditions for applying, the application is halted. If all the relevant evidence is in place then the application is countersigned by another DWP member of staff.

Further checks are then carried out by the Jobcentre Plus National Insurance Number Centre (NC). These checks include checking whether the applicant already has a NINo in place alongside checking the application and evidence that the applicant has provided. If at this point the NC considers the evidence of identity is incomplete or suspect then the application is halted. Equally if it is found that the person already has a NINo in place, again the process is then halted.

If at this stage the application passes the agreed checks carried out by the NC then a NINo is allocated on the DWP's departmental Customer Information System (CIS). The NC then register that National Insurance (NI) account on the Her Majesty's Revenue and Customs National Insurance and Pay as you Earn System (NPS). Thereafter the applicant is sent their NINo.

2.3.4 Statistical implications of the accuracy and quality of the data collection process

The MWS is a dataset from a long standing administrative source that is likely to continue being used by the HMRC and DWP. Like any administrative dataset, the MWS is domain-specific and this is its biggest strength as well as being a potential weakness.

The MWS dataset includes everyone who has been allocated a National Insurance number through the adult allocation process since January 2002. It is a complete dataset of overseas nationals who have arrived in the UK, or intend to come to the UK, and successfully applied for a NINo.

By definition, the dataset does not include overseas nationals resident in the UK who have not registered for a National Insurance number, either because they do not intend to work, cannot register as they have no right to work or study in the UK or are working without being registered with the HMRC. It is therefore a proxy for counts of immigrant workers, but should not be seen or used as a proxy for immigrants generally. Also it cannot be used as a proxy for the net flow of migrant workers into the UK, as it does not contain any information about emigration from the UK. By using this administrative dataset alongside other administrative datasets and survey data, the methodology used by PSD to estimate immigration takes account of this weakness.

As MWS is one of several administrative sources used in the estimation of international immigration at local authority (LA) level, record linkage is used both within and between sources to minimise definitional differences and duplication. PSD match records using a combination of identifiers, such as sex, date of birth and postcode. In order to do this they need the MWS dataset to be an accurate description of the overseas nationals' particulars. Missing and inaccurate data makes it much harder to match records together, in turn reducing the robustness of the population estimates.

Any future changes by DWP to the process of allocating NINOs has the potential to introduce bias in the outputs of the data analysis of the MWS, either because the new procedure would make it easier or harder for an applicant to apply, by reducing the level of quality of the details recorded or by changing the time taken for an application to be processed. If DWP ceased to collect specific pieces of information during the application process this could also be problematic for ONS, especially if this information was vital to the calculation of population estimates.

Notes for Data Collection:

1. <https://www.gov.uk/national-insurance>
2. Information on the different types of National Insurance known as classes: <https://www.gov.uk/national-insurance/national-insurance-classes>
3. Information on the different types of National Insurance known as classes: <https://www.gov.uk/national-insurance/national-insurance-classes>
4. <https://www.gov.uk/tax-come-to-uk>
5. The countries with a bilateral agreement are: Barbados, Bermuda, Bosnia and Herzegovina, Canada, Isle of Man, Israel, Jamaica, Japan, Guernsey, Republic of Korea, Macedonia, Mauritius, Montenegro, Philippines, Serbia, Turkey and the USA.
6. Information on the different types of National Insurance known as classes: <https://www.gov.uk/national-insurance/national-insurance-classes>.

3 . Communication with Data Supply Partners

3.1 Data Provision Arrangements

A Memorandum of Understanding (MOU), for the supply of the Migrant Workers Scan (MWS), was finalised in 2010 between HM Revenue and Customs (HMRC), the Department for Work and Pensions (DWP) and ONS. This was updated in June 2016 and is valid for the next 5 years. The MOU sets out the intentions of the three organisations in relation to data sharing, identifying what, how, why and when the MWS dataset is to be shared.

3.1.1 Legal Access to the MWS

The permissive legal gateway for sharing the relevant MWS data is covered in Section 122AA of the Social Security Administration Act 1992. All three organisations are Data Controllers under the Data Protection Act 1998, with the DWP assigned the role of Data Processor. In accordance with Section 39(4)(c) of the Statistics and Registration Service Act 2007, the MOU covers arrangements for equivalent government departments in Scotland, Wales and Northern Ireland to have access to the MWS data passed to ONS. HMRC considers that the disclosure of information to DWP and ONS is necessary and proportionate because of the need for HMRC, other Government Departments and Public Bodies need to understand the level of migrant workers present in the UK.

3.1.2 Statement of Needs

ONS requires an extract of MWS, consisting of individual records, from the HMRC National Insurance and Pay as you Earn System (NPS). Our on-going requirement is for updated quarterly extracts, covering the UK, delivered annually (usually in December). The transfer consists of 4 separate SAS files, labelled to show the day and month the extract was taken.

The MWS datasets that DWP receive from HMRC only include records of overseas nationals who have applied for a NINo, that is migrants. These are selected in accordance with standard criteria – all cases where the “Nationality” field does not have a country code relating to Great Britain or a British protectorate and has an end of period of “Abroad Liability” after 6 April 1975¹.

The dataset transferred to ONS contains the DWP encrypted NINos and provides cumulative data of NINo allocations back to 2002. This means that in each extract individual records are updated to take account of new information received by HMRC (for example, details on change of address or date of death).

The MWS dataset that the Population Statistics Division (PSD) uses contains the following variables:

- address and postcode
- date of birth
- age at extract date (derived variable)
- sex
- nationality
- date of arrival in the UK
- date of registration (HMRC date of entry variable)
- date of death
- address indicator
- encrypted NINo (DWP derived from the HMRC NINo)
- geographical referencing variables (derived during DWP processing)
- date of extract (derived at load)

ONS engages with DWP at least 8 weeks prior to the next data delivery to ensure there are no changes to the next data delivery and to discuss any updates or changes to the metadata.

3.1.3 Transfer of data between HMRC, DWP and ONS

The prepared MWS datasets are physically transferred from HMRC to DWP. DWP sign off the transfer of data, by email, once the data has been successfully loaded into their secure computer system. The whole process is underpinned by DWP/HMRC Operational Control procedures to ensure a secure process of transfer and receipt.

After carrying out cleaning, geocoding the records and carrying out quality assurance, DWP prepares the encrypted datasets as SAS files. They physically transfer the MWS data to ONS using a secure procedure. ONS confirms to DWP that the data has arrived.

3.1.4 Procedures for managing issues related to data provision

The MOU sets out the named contacts for each of the 3 organisations (HMRC, DWP and ONS). Communication on issues, problems and reviews of the exchange of data is handled by these personnel by email and telephone conversations as needed.

3.2 Engagement with users

PSD continually engages with users to understand how well outputs meet their requirements. PSD's user engagement activities include formal consultations on proposed changes to outputs, regular communication on plans through a quarterly newsletter and external events open to all users. In addition, where evaluating changes to methods or sources has required specialist knowledge of local areas, PSD has organised Local Insight Reference Panels to elicit the views of relevant local authorities. From these activities, any issues relating to the sources and their fitness for the proposed use, will naturally come out. Issues restricted to one output will generally be addressed by the team responsible for that output while the Stakeholder Engagement team in PSD takes an overview of any issues with more general implications and ensures that this is considered in development of outputs across the division. It should be noted that users are more likely to comment on the overall methodology and the effect that it has on the final statistics than on a contributory data source.

Any issues around the quality of the statistics are described in the Quality and Methodology Information report accompanying each output. Issues around specific administrative data sources used in producing the statistics are considered in Quality Assurance of Administrative Data reports such as this.

When changes are proposed to methods (including changes in data sources being used in producing statistics) the ONS Population Methodology and Statistical Infrastructure Division will assess the resultant methods prior to implementation to assure that they are of sufficient statistical quality to meet user needs and are an improvement on the previous method. An independent evaluation by academic experts may also be undertaken, should methodological changes be extensive. The methods are also subject to scrutiny by the UK Statistics Authority as part of the National Statistics accreditation programme under Principle 4 of the Code of Practice for Official Statistics (sound methods and assured quality).

The Responsible Statistician is named for each release and contact details for them are provided, so should someone have concerns over the statistics they are able to communicate them with us. Methodology documents are published to enable users to provide scrutiny.

Notes for Communication with Data Supply Partners:

1. Abroad liability relates to a non-UK national. Source data from the National Insurance and Pay as You Earn System (from which the Migrant Worker Scan is extracted) is available from 1975.

4 . Quality Assurance principles, standards and checks applied by data suppliers

This section details the checks and standards applied to the data prior to receipt by our Population Statistics Division (PSD). The checks carried out by PSD upon receipt of the data are detailed in Section 5.

HM Revenue and Customs (HMRC) carry out checks on the Migrant Workers Scan (MWS) extracts before sending them to the Department for Work and Pensions (DWP). These include quality checks on the file size and format, the number of variables loaded and the reference date of the extract.

DWP clean the data, adhering to a series of data cleansing rules applied to control for known quality issues. This includes “cleaning” the nationality variable for nationality codes that are now redundant. As a result, quality is improved as the statistics represent current world areas and are consistent over time.

The data are subject to routine quality assurance in order to produce the statistical outputs. There are checks at the file loading, file extraction and processing stages as well as checks on the time-series and checks against internal management information. Currently DWP have noted only one error within processing, where a single NINo record in the raw input file (which had a corrupted format) caused the loading of a scan to fail. This issue was manually corrected by DWP and did not impact on the subsequent quality of the datasets sent to PSD. The outgoing dataset to PSD is checked for the appropriate number of records and that it contains the correct variables.

DWP are committed to producing accurate, timely and high-quality official statistics publications that take into account user needs and which are produced and disseminated in accordance with the UK Statistics Authority Code of Practice. There is a stakeholder group within DWP who deals with any issues regarding data, operations or policy which may impact the National Insurance number (NINo) statistics. Users are informed of issues relating to interpretation of the NINO statistics (especially as a time-series).

Code for producing official statistics may at times be peer-reviewed, for example when it is especially complex. Some statistical processes are also subject to formal independent peer review, internal audit assessment, peer review by other Departments and/or UK Statistics Authority assessment. In all cases the DWP conduct reasonableness checks when statistics are produced following changes to the processing regime which includes looking for discontinuities in time series of the data.

5 . Producer’s Quality Assurance investigations and documentation

5.1. How does our Population Statistics Division (PSD) use the Migrant Workers Scan (MWS) in its calculation of international migrants?

We produce annual estimates of the resident population of England and Wales as at 30 June every year, by local authority, sex and age. The most authoritative population estimates come from the Population and Household Census which takes place every 10 years in the UK. Population estimates from a census are updated every year until the next census, to produce mid-year population estimates.

We use a cohort component method to estimate the population of England and Wales in a year when a census has not occurred: the population, by age, sex and local authority, is aged on from the previous year, births are added and deaths are removed and adjustments are made for internal and international migration and special populations such as prisoners. More information on the cohort component method is available in the [Methodology Guide for mid-2015 UK Population Estimates \(England and Wales\) June 2016](#) on our website.

Our national estimates of international migration are based on the International Passenger Survey (IPS), which is a long-running ONS survey that operates at UK ports of arrival and departure. This is the best source of information currently available on the numbers and characteristics of international migrant flows. But sample sizes from the IPS are too small to provide robust estimates of migration at local authority level, and a migrant's initial intentions about where they will settle may not be realised. Local authority estimates of international immigration and emigration are derived using a variety of data sources and methods. The MWS is one of the administrative sources used in these methods.

For immigration into England and Wales, the long-term international migration estimate for first-time migrants and returning non-UK born migrants is streamed (into students, workers and others), and relevant administrative sources are used to distribute immigrants to each local authority. Record linkage is used both within and between sources to minimise definitional differences and duplication.

MWS data is used in the calculation to distribute international higher education student immigrants (HE students), immigrant workers (Workers) and other international immigrants (who are not working or are not students or who are children, known as Others). Specific details identified on the MWS records, namely the difference between date of arrival in the UK and date of registration (date of allocation of a NINo), are used to identify whether migrant workers are long-term or short term migrants. Distributions of long-term migrants by Local Area (LA) are used in the calculations of long-term immigrants in the mid-year population estimates, whereas distributions of short-term migrants by LA are given to our Migration Statistics Unit to use in their calculation of short-term international migration (STIM).

MWS data is also used in the calculation of international emigration at local authority level. We calculate these using a statistical model, which produces a more robust estimate of emigration at local authority level than the IPS alone can provide. The model estimates the numbers of international emigrants over the year using relationships established between the estimate of emigration from the IPS and estimates from other data sources (covariates). MWS counts of EU8¹ nationals by local authority are one of the covariates used in the model. Further information on the emigration model is given in the [Population Estimates Methodology Guide](#).

5.2 Regular quality assurance checks carried out on the received MWS datasets

These checks are scheduled to be carried out by PSD once the MWS datasets have been loaded into the secure working environment at ONS. The checks are designed to identify any obvious errors that have occurred in the size and content of the dataset as well as confirm that the data are plausible. They also check that the datasets are fit for purpose, that they can be processed in line with the agreed methodology to produce distributions of immigrants by LA and that they can be easily merged with the cumulative dataset of previous MWS extracts.

For each check, parameters are given based on our knowledge of the quality of these data. As long as the data lies within these parameters, the results are recorded but no further action is taken. Otherwise concerns are reported to Data as a Service Division (DSD), within ONS who communicate with DWP as needed. Immediate action on the checks is taken as needed and a copy of the final report is saved by PSD and sent to DSD.

Table 1: Checks carried out by PSD on the MWS datasets

Category of Check	Checks to be carried out	Action to be taken
Check the correct metadata has been delivered into the secure environment	<p>Check that the latest metadata is available and that it reflects the variables in the dataset.</p> <p>Check that the metadata includes changes to the composition of the dataset since the last extract and why these changes have occurred, or states that no changes have taken place.</p>	<p>If metadata missing, incomplete, outdated or ambiguous, report concerns to DSD, who will contact DWP as needed.</p> <p>Do not carry on processing until this has been resolved.</p>
Check the correct data has been delivered into the secure environment	<p>Check that the number of records in the dataset is as recorded by DSD.</p> <p>Check number of records against the numbers in previous extracts. The number of records should increase over time, usually by around 10% over a 12 month period. e.g. Oct11 to Oct12 extract.</p> <p>Check the dates of registration covered by the data span from January 2002 to a date slightly before, or on the extract date</p> <p>Using the metadata, check that dataset only contains expected variables.</p>	<p>If numbers differ, report concerns to DSD, who will investigate further.</p> <p>If number of records is less than previous extracts, or has increased by over 15%, or if earlier records have been included, report concerns to DSD, who will contact DWP as needed.</p> <p>Check information on unexpected or missing variables in the metadata. Report concerns to DSD, who will contact DWP as needed.</p> <p>Do not carry on processing until these issues have been resolved.</p>
Check format of responses for consistency	Check that the length and format (character/numeric /date) of the variables agrees with those in the cumulative file of MWS created previously.	If the format of the data is not as expected (as in previous extracts) it will be cleaned or, via DSD, referred back to DWP as appropriate.

<p>Check for missing data in key variables; assess the importance of missing data and record how many items are missing for which variables and whether variables are vital to processing.</p>	<p>Check the following variables for missing data:</p> <p>CCNINO (Encrypted NINo)</p> <p>CDDOB (Date of Birth)</p> <p>CCSEX (Sex)</p> <p>CDUKARRV (Date of arrival)</p> <p>CDREGIST (Date of registration)</p> <p>CCADDRST (Address status)</p> <p>GEOFLAG (Geographic area)</p> <p>COAIND (Quality of geo-referencing indicator)</p> <p>CCNATCODNEW (Nationality)</p> <p>Age (derived from date of birth)</p> <p>Check number of missing records in variable CCPCODE (Post code)</p> <p>Check number of records with no address information in any of the four address line fields.</p>	<p>It is not anticipated that any of these variables will have missing data. However, if less than 0.1% of records do have missing data for any of these variables it will be recorded but no action will be taken. If more than 0.1% of records have values missing from a particular variable this needs to be investigated further, in order to establish whether these are new records or old records. Report concerns to DSD, who will contact DWP as needed. Do not carry on processing until this has been resolved. Previous checks have indicated that the Nationality variable (CCNATCODNEW) will contain records which have been allocated to 'unspecified', 'abroad – not known', 'not yet recorded or 'stateless'. Unless the number of records allocated to these categories exceeds 0.5% no action will be taken. Previous checks have indicated postcodes are often missing from records. If more than 15% of postcodes are missing, report concerns to DSD who will contact DWP about the declining quality of address data. Previous checks have indicated address information is sometimes missing from records. Unless more than 0.5% of addresses are missing it will be recorded but no action will be taken.</p>
<p>Check for repeated records</p>	<p>Check for duplicate encrypted NINO's to ensure no record appears more than once in the data set</p>	<p>Record the number of duplicate encrypted NINOs and de-duplicate the dataset.</p>
<p>Check all dates are plausible, i. e., no actions before birth /after death</p>	<p>Check no-one was born after their death date</p> <p>Check no-one was registered for a NINo before they were born</p> <p>Check no-one is under 15 yrs at the date of extract</p>	<p>Report concerns to DSD, if more than 0.1% of records fail any of these checks.</p>

Check no-one is over 120 years of age without a date of death at the date of extract

Check for people who have an age of over 100 years at registration

Check no-one was registered 33 days after date of death (as it should not take more than 33 days between application and registration)

Check no-one has an arrival date before they were born

Check non-date responses are valid (e.g. sex)	Check the variables which do not contain dates only have coded responses which are listed in the metadata.	If any coded responses are found which the metadata does not explain, refer the discrepancy back to DWP via DSD.
---	--	--

Look at time /area trends to highlight any areas where data might be missing	Check whether records 'disappear' from one extract to the next. The data is cumulative so this should not happen Check sex variable is approximately 50:50 overall Check age distribution for new records shows a peak for those of working age at date of registration Check distribution of new records by LA is similar to distribution seen in previous extracts.	As it is known there is a problem with data 'disappearing' between extracts only report this to DSD if more than 0.1% of records disappear. As long as the sex ratio is found to be within a 45:55 boundary this is acceptable. Report any concerns to DSD, who will contact DWP as needed. If there is no peak for those of working age undertake further checks on the age distribution pattern. If there appears to be an inconsistency with past extracts, report concerns to DSD. Report concerns to DSD if distribution for an LA differs by more than 5% from the previous extract.
--	---	--

Source: Office for National Statistics

Notes on Quality Assurance by ONS:

1. Estonia, Latvia, Lithuania, Poland, Hungary, Slovenia, Slovakia and Czech Republic who joined the European Union in 2004.

6 . Conclusion

6.1 Strengths and limitations of the MWS datasets, in relation to how we use our datasets.

Strengths

- Large cumulative dataset of all adult overseas nationals (non-UK citizens) who have registered for and been allocated a NINo.
- Established system of quality checks carried out by Department for Work and Pensions (DWP) at data collection and on the MWS dataset before being delivered to Office for National Statistics (ONS).
- Robust proxy for immigrant workers; analysis by ONS carried out in 2013 showed that the data was of high quality.
- Long-term continuation of data collection by DWP.
- Migrant Workers Scan (MWS) dataset contains records since Jan 2002 onwards.
- Timely delivery of extracts. Extracts are taken quarterly and delivered annually to ONS – so extracts taken in January, April, July and October in a year are delivered to ONS in the December.
- Ability to link new activity on a record (such as change of address) to a previous entry using the encrypted NINo as a unique identifier.
- Large amount of details (variables) available, including sex, date of birth, postcode, date of registration and date of arrival in the UK. This is particularly useful in categorising whether an individual is a recent migrant and to link to records in other administrative sources.
- Nationality of individuals is recorded, allowing analysis between non-EU and EU migrant registrations.
- Date of birth is recorded, allowing age to be derived which in turn allows the dataset to be filtered by age group as required.
- Address variables enable geo-referencing of a record even if the postcode is missing or incorrect.

Limitations

Coverage

- undercoverage of all migrants. The MWS specifically excludes certain types of migrants who have not yet applied or who are not eligible to apply for a National Insurance number (NINo); UK or dual-nationality, children of migrants aged under 15 years 9 months, migrants who work illegally, adult migrants who neither work nor claim benefits (such as spouses of working migrants) and asylum seekers who have not been granted the right to reside in the UK
- overcoverage of migrants. The MWS is an ongoing list, not a register, and there is no de-regulation policy. Once allocated a NINo an individual stays on the list even if they should subsequently emigrate and after they have died. The list may also include people who do not stay in the UK long enough to meet the definition of a usual resident (that is, short-term migrants) and a very small number of people who successfully apply for a NINo but end up not coming to live in the UK
- we mitigate the under- and over-coverage of migrants in the MWS by using the MWS alongside other administrative datasets and survey data in its methodology to calculate the number of international immigrants to England and Wales
- The unique identifiers (encrypted NINos) used in the MWS are different to those used on other administrative sources, such as the Patient Register, making it harder to link individual records together.
- MWS does not provide any information on how long an individual stays in the UK. There may also be a lag between an overseas national entering the country and registering for a NINo. Consequently, the data are not directly comparable with our estimates of long-term migrants, that is, those who stay in the UK for a period of 12 months or more or estimates of migration from the Annual Population Survey.
- Information on nationality is based on passport data at allocation and does not reflect later changes in nationality.
- Previous country of residence information is available on the dataset, but according to DWP it is not robust enough for ONS to use in its statistics.
- Address information may be out of date. Changes to a migrant's residential address will only be recorded if the migrant or DWP inform HMRC of the change, through interaction with the benefits system or on an annual tax return. Quality assurance work (carried out by ONS in 2009) has indicated that a very small proportion of migrants may be providing a non-residential address when they initially register for a NINo, for example, contact details for a bank, employer or employment agency.
- Date of arrival in the UK. This is subjective and relies on applicant's recollections. The date will apply to their most recent entry into the UK of which DWP/HMRC are aware. In most cases authorities are not told of re-entry to the UK – they only know if the migrant has forgotten their NINo or applies for benefit.
- A tiny proportion of records disappear from one cumulative dataset to another.
- Date on the extract, for example, Oct14 relates to a date at the beginning of October 2014. Extract only contains NINo registrations that have been allocated up to that date and not for the whole of that quarter (October to December 2014). Allocations for the rest of that month and following two months will be included in the following extract (Jan15).

6.2 Justification of rating

The risk of quality concerns over the MWS is judged to be low, for the following reasons:

There is an established process for allocation of a NINo to overseas nationals, which is transparent and understood by data processors. There is completeness of the data in the datasets which feed into the MWS. There is clear agreement about what data will be provided to ONS, when, how, and by whom, as set out in the Memorandum of Understanding, mitigating the higher risk of the data been transferred to ONS from HMRC via DWP.

Quality Assurance processes are followed at each stage of data collection, data supply and processing. There is clear understanding at ONS on the limitations of using this data source as a proxy for migrant counts. The risk of misusing the MWS data to count migrants is mitigated by the design of the international immigration methodology, which uses a combination of administrative data to provide estimates of the distribution of immigrants by LA. International immigration is but one of several components contributing to the mid-year population estimate, including natural change, internal net migration and international emigration. Finally, the [methodology](#) used is concisely documented and published on our website.

We have also considered the public profile of the statistics. The distributions of immigration we produce using the MWS are only used in the production of the mid-year population estimates and are not separately available. However final estimates for international immigration by LA are published by us in the [Components of Change tables](#) which accompany the [annual mid-year population estimates](#) release. The public interest profile of these tables is deemed to be high, for the following reasons:

- they are a component of the mid-year population estimates, which are National Statistics used to make decisions on resource allocation by central and local government and are used as the basis of many other official statistics
- international immigration is a highly politically sensitive topic, which also receives substantial and sustained media coverage

Based on the above, we have therefore judged that the appropriate level of assurance for this data source is A1, meaning that a basic level of assurance is required for these sources.