Office for National Statistics

National Insurance number (NINo) data: quality assurance of administrative data used in population statistics, Jan 2017

The quality assurance undertaken on administrative data for National Insurance Number (NINo) data used within Population Statistics Division (PSD) publications.

Table of contents

1 . Introduction

National Insurance number (NINo) data for adult overseas nationals for the UK are produced and published by the Department for Work and Pensions (DWP), with collaboration from Her Majesty's Revenue and Customs (HMRC). These data are also used by the Office for National Statistics (ONS) for their migration outputs and research. However, ONS do not publish the full range of published NINo statistics, so these are not covered by this document. The use of NINo data within ONS outputs is as a comparative tool within our publications and not to produce ONS statistics.

Within the Population Statistics Division (PSD) of ONS, NINo data are used by the Migration Statistics Unit and the Migration Analysis Group:

- within the publication of the Migration Statistics Quarterly Report

- within briefings shared amongst other government departments

- in ad hoc research comparing data sources (see the examples below)

This report covers the processes, from data collection through to publications produced by PSD and focuses on quality assurance. It identifies potential risks in data quality and accuracy as well as details of how those risks are mitigated.

This report does not cover the publication of NINo statistics; its purpose is to examine the quality assurance processes that NINo data go through when they are used as a data source in migration estimates. This document is intended to supplement existing documentation:

- the ONS Long-Term International Migration Quality and Methodology Information document and Migration Statistics Quarterly Report User Information

- National Insurance number allocations to adult overseas nationals entering the UK

- A Reconciliation of ONS International Migration Estimates, which compares 3 administrative sources (NINo allocations, patient registrations and the worker registration scheme) with IPS estimates of both Long-Term and Short-Term International Migration

- a summary note called Sources of Migration Statistics, which explored the differences between IPS estimates and NINo data

- a Note on the difference between National Insurance number registrations and the estimate of Long-Term International Migration, which was a more in-depth comparison of IPS estimates and NINo data

This report has been published to help understand data processing and provide reassurance that the subsequent statistics produced by PSD using NINo data are suitably robust.

Within PSD, NINo data were assessed using the UK Statistics Authority's Quality Assurance of Administrative Data Toolkit. The results of those assessments are that NINo data received an A2 rating. The scores provided by each team and the rationale behind those scores will be provided, in terms of both the risk and profile components, in Section 6 of this report.

The A2 rating means that an enhanced level of assurance is required for these sources and this document will provide information to meet this level of assurance. If you feel that this report does not adequately provide this assurance then please contact pop.info@ons.gsi.gov.uk with your concerns.

The toolkit outlines 4 areas for assurance; the rest of this document will be split into these areas, with further subdivisions by country. The areas for assurance are:

- operational context and administrative data collection

- communication with data supply partners

- quality assurance (QA) principles, standards and checks applied by data suppliers

- producer's QA investigations and documentation

2 . Operational context and administrative data collection

2.1 Background information

In the case of an individual already living in the UK when they turn 16 years of age, a NINo is automatically issued before their 16th birthday by HM Revenue and Customs (HMRC). This is based on their Child Reference Number. Individuals who move to the UK after the age of 16 (adult overseas nationals), or were not allocated a Child Reference Number originally, must apply for a NINo through the Department for Work and Pensions (DWP).

The purpose of a NINo is to make sure that the National Insurance contributions and tax a person pays are properly recorded against their name. These contributions count towards certain benefits, such as Maternity Allowance and the State Pension. A NINo also acts as a reference number when communicating with DWP and HMRC.

The NINo figures encompass adult overseas nationals allocated a NINo for whatever reason, that is, the figures cover benefit or tax credit recipients as well as workers (including self-employed). This extract is known as the Migrant Worker Scan (MWS). All adult overseas nationals allocated a NINo are included, regardless of their length of stay in the UK. They are based on the recorded registration date onto HMRC's National Insurance and PAYE Service (NPS), that is, after the NINo application and allocation process has been completed. This may be a number of weeks or months after arriving in the UK, so they do not measure the date of arrival in the UK, nor do they measure the job start date.

There is inherent bias on late registrations (particularly ex-students) and early registrations (before they arrive in the UK). An individual can start work before their NINo arrives, if they can prove they can work in the UK and tell their employer they have applied for one. DWP acknowledge this and provide substantial information as to the limitations of the data.

2.2 Process from collection to analysis

In order to apply for a National Insurance number (NINo), in the first instance, a migrant worker makes an enquiry to the DWP National Insurance number application line. The majority of applicants must attend an "Evidence of Identity interview" (EOI) at a specified Jobcentre Plus office where 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.

The service delivery target for allocation of a NINo is 15 days. If a decision is made to allocate a NINo and registration is confirmed, the NINo is recorded on DWP's departmental Customer Information System (CIS) and the NINo Centre create a full National Insurance (NI) account on the HMRC National Insurance Pay As You Earn System (NPS) from which the Migrant Worker Scan (MWS) is sourced.

DWP receive a quarterly cumulative scan from the HMRC's NPS, which contains details of all adult overseas nationals allocated a NINo up to the date the scan was run. This is used to produce these National Statistics. ONS are granted pre-release access before the publication of NINo data for QA purposes, which means any issues would be flagged up before the NINo data were used.

2.3 Revisions

DWP's general policy is to be transparent with users about the need for revisions, how and when to expect scheduled revisions and how these will be communicated to users. These are outlined in the Policy on Revisions of Official Statistics. Scheduled revisions may cause a delay in receiving the finalised data but ensure greater accuracy. Any revisions or errors would not affect the Long-Term International Migration estimates as these data do not feed into the estimates.

3 . Communication with data supply partners

There are quarterly Migration Reporting Working Group (MRWG) meetings, which act as a means to discuss:

- the progress toward publication of NINo data

- any proposed changes for the future

- any risks and issues (such as processing issues) that exist or may emerge and the mitigation for those risks and issues

The MRWG meetings involve representatives from some of the devolved administrations and certain government departments of relevance to migration statistics.

The process of data transfer between the Department for Work and Pensions (DWP) and Population Statistics Division (PSD) is via a template. PSD send a template comprising text and tables to DWP, who update the template with the latest data. The reference period for the data is 3 months ahead of the Long-Term International Migration data and is published the same day as the Migration Statistics Quarterly Report (MSQR) publication by ONS. Basic checks are made by PSD, for example, on percentage calculations, or change over time. Any errors found within the template completed by DWP for PSD, are addressed via email.

DWP hold consultations for a minimum of 4 weeks and a maximum of 13 weeks depending on the nature of the consultation. They will respond to every consultation providing a summary of all responses, the way forward that will be taken and the rationale for the decisions on the way forward.

The views of statistics users have been sought and as a result, 2 reports have been published in regards to reconciliation of NINo data with IPS data. Differences between the 2 sets of data and their definitions were explained using other data such as Long-Term International Migration estimates and Short-Term International Migration estimates.

3.1 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.

4 . QA principles, standards and checks applied by data suppliers

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

The Department for Work and Pensions (DWP) is 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 Code of Practice for Official Statistics.

The data have shown no known data quality issues. The data are 100% counts, which do accurately reflect the number of NINos registered over time. Data are not subsequently amended and are not affected by retrospection. These processes are the same throughout the devolved administrations and so there is no difference in quality throughout England, Wales, Scotland and Northern Ireland.

All statistical sources need to be subjected to quality checks when producing official statistics from them. All applicants for a UK NINo are subjected to a rigorous identity checking process. For the majority of applicants, this process includes thorough face-to-face interviews and the use of document examination tools to ensure the authenticity of any documentary evidence provided in support of the application. Also corroborative checks are conducted with third parties (including other government departments) to verify information supplied during the interview.

At the point of data loading within HM Revenue and Customs (HMRC), the extracts they receive are checked for file format, file size, reference date and variables loaded to ensure the integrity of the input data. The only noted error within processing, which does not impact on the subsequent quality of the data feed returned to PSD, is regarding a single NINo record in the raw input file that had a corrupted format. This caused the loading of the scan to fail, which meant it had to be manually corrected until a fix can be put in place by HMRC.

The prepared data are physically transferred from HMRC to DWP using a secure DWP/HMRC Operational Control procedure.

Further quality assurance is undertaken by DWP in order to produce the statistical outputs. This includes cleaning the data, adhering to a series of data cleansing rules. For example, the nationality variable is "cleaned" for nationality codes that are now redundant. There are further checks at the file loading, file extraction and processing stages as well as checks on the time series and checks against internal management information.

The outgoing dataset to PSD is checked for the appropriate number of records and contains the correct variables. 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 DWP conduct reasonableness checks when new statistics are produced (for example, are they in line with previous figures, are they what would be expected given what has happened since, and are there plausible explanations for changes).

5 . Producer's QA investigations and documentation

5.1. Checks applied within PSD

Some data are sent to Population Statistics Division (PSD) in the form of embedded text where PSD checks the narrative for sense to make sure that the message being given matches what the figures show. Completed table templates are also sent from the Department for Work and Pensions (DWP) to PSD, with the data from these tables also used within corresponding downloadable Excel spreadsheets created within PSD. Percentage changes or numerical differences over time are checked by several different individuals within PSD for accuracy before they are signed off.

6 . Conclusion

6.1 Strengths and limitations: comparison with ONS's Long-Term International Migration estimates of immigration

For a summary of the differences between NINo and IPS data, please see [Sources of Migration Statistics](). ONS's overall data provide a better indication of long-term trends of international immigration than NINo data for adult overseas nationals, because ONS data include children and migrants arriving who do not sign onto the HM Revenue and Customs (HMRC) National Insurance Pay As You Earn System (NPS) system as they are not in the UK to work (for example, many full-time students).

Furthermore, NINo allocations for adult overseas nationals include visitors and short-term migrants. This definition of migrants (adult overseas nationals entering the UK and allocated a NINo) differs from other published migration statistics, such as IPS estimates and statistics on foreign workers derived from the Labour Force Survey (LFS). The IPS statistics define a migrant as someone who stays in the UK for at least a year following arrival (the UN definition). The LFS defines a foreign worker as someone who works but has foreign citizenship, and a foreign-born worker as anyone born outside of the UK, including British citizens. These various definitions mean that different numbers will be produced if these data sources are used as proxies or measures of international migration

As the NINo data are used as a comparator to the Long-Term International Migration (LTIM) data and do not feed into it, there is no risk to the quality of the LTIM estimates.

6.2 Justification of rating

NINo data used within ONS outputs have been given an "A2" rating based on risk profile scoring. This is an enhanced level of assurance. NINo data are considered a low concern in terms of quality ("risk") and of high public interest ("profile").

The NINo data are considered lower concern in terms of quality as:

- they are used as a comparator rather than to feed into ONS data

- the data are published elsewhere by the Department for Work and Pensions (DWP)

- DWP conduct validation checks on the data throughout the process of collection and analysis

- the data are used for the same purpose they were designed for

The NINo data are considered to be of high public interest as:

- they have recently received substantial media attention regarding comparisons between NINos and Long Term International Migration

- migration levels are of great interest to the general public