

Quality assurance of administrative data report for real estate

Quality assurance report investigating the administrative data sources used in the production of short-term economic indicators.

Contact:
Karen Carter
stoi.development@ons.gsi.gov.
uk

Release date:
27 October 2017

Next release:
To be announced

Table of contents

1. [Introduction](#)
2. [Quality assurance of administrative data assessment](#)
3. [Areas of quality assurance of administrative data \(QAAD\)](#)
4. [Summary](#)

1 . Introduction

Background

National Accounts and Economic Statistics (NAES) within Office for National Statistics (ONS) collect data from the Investment Property Databank (IPD) for the activities of real estate. These data form one source in the calculation of short-term economic output indicators namely gross domestic product output approach (GDP (O)) and Index of Services (IoS) for the UK.

This report outlines the process data take from initial collection through to the output of the release. It identifies potential risks in data quality and accuracy as well as details of how those risks are mitigated.

This report forms the latest in a series of quality assurance of administrative data (QAAD) reports produced by NAES to investigate the administrative data sources we use in the production of short-term economic output indicators as set out by the [UK Statistics Authority](#). As such, this report specifically focuses on our administrative data use for real estate, that is, UK Standard Industrial Classification 2007 (SIC 2007) 68.2, and does not aim to cover other aspects of SIC 68 such as 68.1 and 68.3 as these are collected by ONS using Household Final Consumption Expenditure (a non-administrative data source). Separate industries where we use administrative data will be considered in other QAAD reports in the series.

Further information relating to quality and methodology for the short-term economic output indicators can be found in our [Gross domestic product, preliminary estimate](#) report and our [Index of Services](#) bulletin.

Standard Industrial Classification (SIC) overview

The activities of real estate cover all activities under UK SIC 2007 division 68. This includes activities of acting as lessors, agents or brokers. Activities in this industry may be carried out on own or leased property and may be done on a fee or contract basis.

Based on the UK [SIC \(2007\)](#) the industry is classified to three groups:

- 68.1 – Buying and selling of own real estate
- 68.2 – Renting and operating of own or leased real estate
- 68.3 – Real estate activities on a fee or contract basis

According to the Inter-Department Business Register (IDBR) there were 90,990 enterprises classified under division 68 in March 2016. This is an increase of approximately 2,905 enterprises (3.3%) from the previous year (March 2015).

Within division 68, there were 75,030 enterprises with fewer than five employees (82.5% of the division), compared with just 1,930 enterprises with 20 or more employees (2.1% of the division). Of the 90,990 enterprises in division 68, the annual turnover of 41,245 (45.3%) enterprises was below £250,000.

The majority of enterprises within division 68 were allocated to 68.2 – Renting and operating of own or leased real estate – which equates to 52,375 enterprises (57.6 % of the total division).

2 . Quality assurance of administrative data assessment

UK Statistics Authority QAAD toolkit

The assessment of our administrative data sources has been carried out in accordance with the [UK Statistics Authority quality assurance of administrative data toolkit](#).

Each administrative data source investigated has been evaluated according to the toolkit's risk and profile matrix (Table 1) reflecting the level of risk to data quality and the public interest profile of the statistics.

Table 1: UK Statistics Authority quality assurance of administrative data (QAAD) risk and profile matrix

Level of risk of Quality concerns	Public interest profile		
	Lower	Medium	Higher
Low	Statistics of lower quality concern and lower public interest [A1]	Statistics of low quality concern and medium public interest [A1/A2]	Statistics of a low quality concern and higher public interest [A1/A2]
Medium	Statistics of medium quality concern and lower public interest [A1/A2]	Statistics of medium quality concern and medium public interest [A2]	Statistics of medium quality concern and higher public interest [A2/A3]
High	Statistics of higher quality concern and lower public interest [A1/A2/A3]	Statistics of higher quality concern and medium public interest [A3]	Statistics of higher quality concern and higher public interest [A3]

Source: Office for National Statistics

The toolkit outlines four specific areas for assurance and the rest of this report will focus on these areas in turn. These are:

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

In the assurance of our data source we have chosen to give a separate risk and profile matrix score (Table 1) for each of the four areas of assurance. This will allow us to focus our investigatory efforts on areas of particular risk or interest to our users (Table 2).

Table 2: QAAD risk and profile matrix assessment of administrative data used to measure activities of real estate

	Low Medium High		
	A1	A2	A3
Operational context and administrative data collection	A1		
Communication with data supply partners	A1		
Quality assurance principles, standards and checks by data supplier	A1		
Producers quality assurance investigations and documentation	A1		

The risk of quality concern and public interest has been set as “low” due to the small contribution that the real estate statistics feed into the Index of Services (1.8%) and gross domestic product (1.4%). As such, a score of A1 (basic assurance) is deemed appropriate for this data source.

All scoring was carried out by National Accounts and Economic Statistics (NAES) based on the level of risk of the data and interest of our users. Results for each area of assurance for real estate are shown in Table 2. If you feel that this report does not adequately provide this level of assurance or you have any other feedback, please contact stoi.development@ons.gsi.gov.uk with your concerns.

Notes:

1. The [Inter-Departmental Business Register \(IDBR\)](#) is a comprehensive list of UK businesses that is used by government for statistical purposes. It provides the main sampling frame for business surveys carried out by both ONS and other government departments. It is also a main data source for analyses of business activity.

3 . Areas of quality assurance of administrative data (QAAD)

3.1 Operational context and administrative data collection (QAAD matrix score: A1)

This relates to the need for statistical producers to gain an understanding of the environment and processes in which the administrative data are being compiled and the factors that might increase the risks to the quality of the administrative data.

Data collection process

Morgan Stanley Capital Investment (MSCI) formally acquired the Investment Property Databank (IPD) in 2012. IPD is a subsidiary of MSCI and a world leader in performance analysis for the owners, investors, managers, lenders and occupiers of real estate. IPD provides critical business intelligence including analytical services, indices, and market information regarding performance and risks to the real estate industry worldwide.

IPD has a simple portfolio data collection process as the properties are held in investment funds and trusts, which are sampled in the IPD UK monthly index. These are valued according to Royal Institute of Chartered Surveyors (RICS) guidelines. Where property funds and trusts are coded based on their geographical location and the sector in which these properties operate, this is done using the [RICS red book](#).

[The IPD UK monthly index](#) coverage tracks the performance of a stable sample of 3,011 property investments, with a total capital value of £45.5 billion as of July 2017. These are then collated and aggregated to the sector breakdowns produced by the index.

The data are collected and submitted electronically through a secure portal in either a standardised Excel submission template or Pisces file from all providers such as: lenders, real estate investment trusts, consultancies, investors, fund managers, property companies and occupiers. The templates are sent electronically and assisted with concise guidance notes and for further guidance they provide clients with the [Data Standard Guidelines](#) document.

Data providers create a record for every property within a portfolio and collect information on: location; value; sector covering retail, office, industrial, residential and other; size; age; and tenancy details. All property records are aggregated to UK level and the data are not protected by law.

IPD do not survey the whole market; instead they rely on a stable sample, which consists of clients that are subscribed to their Portfolio Analysis Service, and as such the sample may not be fully representative. IPD estimate the monthly index coverage to be around 10% compared with the quarterly index of under 30% and annual index of under 40%. IPD is unable to provide 100% market coverage due to the voluntary basis of data submission meaning that the level of coverage will vary; the data are re-weighted to reflect the UK market sizes as accurately as possible. IPD data are accessible on their website though most of their data require a subscription to access it.

Data submitter code of conduct

The quality of data inputs are reinforced through data providers' compliance with a [Data Submitter Code of Conduct](#) developed and published by MSCI. The code of conduct covers internal systems and controls within the contributing firm for preparing data submissions to comply with IPD quality and integrity standards.

Those who submit data are responsible for compliance with the standards set out in the data submitter code of conduct. Non-compliance with the standards set out in the code of conduct may result in the rejection of a data submission or the removal of a submitter's data from the IPD database – procedures are in place to handle breaches of the code of conduct through the IPD technical committee.

IPD have a change management process in place, which encompasses not only methodology and data changes but it also covers corrections and consultation policies. For more information please see the Appendix: Morgan Stanley Capital Investment (MSCI) policies for the IPD index and benchmark determination in the [MCSI Global Methodology Standards for Real Estate Investment](#) guide (pages 43 to 48).

Table 3: Strengths and weaknesses of operational context and administrative data collection

Strengths	Weaknesses
Coverage covers 3,000 plus properties capital value of £45.5 billion	Sample only taken from clients that are subscribed to IPD Portfolio Analysis Service
Timely data source	Monthly index has a lower coverage (10%) than longer term indexes, for example, annual, under 40%
Secure submission portal	Data submission is not statutory
Consistent valuation process based on Royal Institute of Chartered Surveyors (RICS) guidelines	
Code of conduct	
Change management policy	

3.2 Communication with data supply partners (QAAD matrix score: A1)

This relates to the need to maintain effective relationships with suppliers (through written agreements such as service level agreements or memoranda of understanding), which include change management processes and the consideration of statistical needs when changes are being made to relevant administrative systems.

Investment Property Databank communication with data providers

Each month the Investment Property Databank (IPD) sends a standardised data collection template to the sample of data providers selected – many of these data providers of real estate data are also their clients. The template sent is accompanied by concise guidance notes, which explain how to fill in the template, to maximise quality of data submitted and for further guidance they provide clients with the Data Standard Guidelines document and Data Submitters Code of Conduct mentioned in Section 3.1. IPD meet their data providers regularly on an individual basis and in some cases as a group to discuss future product developments and gather feedback with regards to products and services.

In the UK IPD is advised by its Index Consultative Group, which meets twice a year. This group reviews the composition and construction rules for the UK monthly index. It is an independent non-executive body whose members are drawn from property investment industry, universities, major research and chartered surveying firms, and index publication and investment measurement organisations.

NAES communication with IPD

National Accounts and Economic Statistics (NAES) is a subscriber to IPD's Multinational Market Intel product and are therefore allowed to download market data from their secure portal and use it in accordance with the [legal agreement and terms and conditions](#) that are in place.

IPD do not provide evidence for changes in data trends, however, the exception to this relates to instances where they have introduced methodology changes – when this occurs web notices are published on their website to inform users. There is a mechanism for NAES to query data with IPD should anything require further explanation. This is done via email correspondence and ensures an audit trail record of all communication.

Outside of the communication channels mentioned previously, there is no regular direct contact between IPD and NAES and no formal agreements, such as a service level agreement, are in place. Managing such formal arrangements are considered both disproportionate and prohibitive considering the contribution of this data source to gross domestic product (GDP) and Index of Services (IOS) and that the current data provision arrangement has consistently resulted in a monthly data sample since 1987.

Table 4: Strengths and weaknesses of communication with data supply partners

Strengths	Weaknesses
IPD regularly communicate with data providers	Changes to methodologies and policies are not communicated directly to NAES
Independent non-executive consultation group drawn from stakeholders and experts	
Long-term data provision and relationship between ONS and IPD	
Audit trail of records	

The next steps are for us to talk to IPD about communicating changes to methodologies and changes to NAES.

3.3 Quality assurance principles, standards and checks by data supplier (QAAD matrix score: A1)

This relates to the validation checks and procedures undertaken by the data supplier, any process of audit of the operational system and any steps taken to determine the accuracy of the administrative data.

Investment Property Databank quality assessment process

The Investment Property Databank (IPD) runs a monthly internal data quality assessment process during every data update period to identify errors that may have been missed by data providers, as well as unusual or unexpected changes in values over the monthly period. For this process, each quality assessment process has a set of exception criteria appropriate for its market, against which the data are checked. Exceptions include omitted data, illogical data and data outside specified numerical ranges. The acceptable ranges are based on local market conditions and previous results. If any of the exception criteria are triggered in the quality assessment process, the data will not be added to the database until the issue is explained or resolved.

Once each portfolio has been processed in the database, a checking report is produced, which is checked for accuracy. IPD checks that headline results are in the expected range or that valid explanations have been provided if not. Performance measures for the portfolio's history are also checked with the aim of ensuring that no unexpected changes have occurred.

The checking report is then distributed to the client to review the report, so they can verify that the results are accurate and complete. After the clients review, the individual portfolios are compiled into a dataset, which are checked for extreme observations in the context of historic trends. Data quality and consistency are both integral to the IPD indexes and market information. Individual assets may be excluded from IPD indexes when they display performance attributes or abnormalities that are manifestly different from underlying market trends, which may therefore contribute to misleading aggregate reported results.

IPD data quality review

Following each monthly update, IPD operates a quality scoring mechanism based on several criteria including accuracy, completeness and timeliness. Portfolios with low scores are selected for a data quality review.

The aim of the review is to allow an open discussion between IPD and clients to understand why perceived areas of weak data supply are occurring and provide recommendations to help bring clients in line with best practice. It also allows IPD to examine clients' internal data submission processes and assess how well they are adhering to the data submitter code of conduct.

In addition to improving client's understanding of IPD data requirements and methodologies, the review process helps them to better understand clients' difficulties in reconciling source data with their template. The outcome from the review is an action plan outlining recommendations with target dates and the possibility for input and comment from the client. IPD then tracks these plans with the aim of ensuring follow-up at subsequent updates.

IPD revisions policy allows providers to correct errors once they have been identified; IPD will restate indexes if errors were made over the previous 36 months that were material to the index. Such restatements are published on the MSCI website with an explanation of the corrections made.

This method has a number of strengths:

- established quality assessment process
- queries exceptional data with clients
- fund checking report, including historical checks
- redistribution to client for final review
- quality scoring mechanism and data quality review
- clear revisions policy

3.4 Producers quality assurance investigations and documentation (QAAD matrix score: A1)

This relates to the quality assurance conducted by the statistical producer, including corroboration against other data sources.

Investment Property Database (IPD) has produced the real estate index since January 1987. The index is a monthly indicator and the timely data are an important benefit in the compilation of a monthly Index of Services.

The index is supplied in an Excel spreadsheet, which is picked up from the IPD subscription section of their website. A confirmation email is sent to National Accounts and Economic Statistics (NAES) from IPD once it is uploaded. It is delivered in sufficient time to ensure NAES can investigate, analyse and query the data in time for publication within the preliminary estimate of gross domestic product (GDP).

NAES download the data from the IPD secure portal. Validation checks are then in place, which use Excel spreadsheets to carry out summary checks and investigate historical trends, including monthly revisions. NAES believe these limited checks are sufficient when seen in context on top of the checks already carried out by IPD.

Data are then loaded into internal systems, which use one calculation to aggregate the figure for the whole of Standard Industrial Classification 2007 (SIC 2007) 68. In the event of missing data, a statistical method that forecasts a figure based on previous months data is forecasted for the team. It should be noted that the team has never had to use the statistical method due to completeness of the data. The team record an audit trail from when the data come in and have clear regularly updated desk instructions in place, which clearly explain the process. Finally, the final output is quality assured and signed off by two senior managers within the team.

User engagement is continual and the feedback tends to relate to the overall impact of the statistics rather than to the individual data source used, and to date no specific feedback on the use of real estate has been provided

This method has a number of strengths, including:

- the timeliness of data delivery
- established quality assurance checks in place
- data are compared by looking at long historic time series
- desk instructions are regularly reviewed

4 . Summary

In investigating the administrative data source for the activities of real estate, National Accounts and Economic Statistics (NAES) consider the main strengths of the data for our purpose to be the:

- timeliness of data delivery, as the index is a monthly indicator
- detailed knowledge of the subject by the Investment Property Database (IPD)
- comprehensive quality assurance principles in place
- long historical time series
- compliance with the code of conduct

We believe current limitations of this data source are:

- changes to methodologies and policies are not communicated directly to NAES

In constantly seeking to improve our data sources we will be taking next steps to address these limitations and these will be communicated to users in the future quality assurance of administrative data (QAAD) report updates for this topic.

However, despite these slight limitations, based on the low risk of quality concerns and small contribution that the real estate statistics feed into Index of Services (1.8%) and gross domestic product (1.4%), NAES consider this data source to fulfil the requirements of an A1 assurance rating.