

# **1996 CENSUS DATA QUALITY: JOURNEY TO WORK**

**Euan Robertson**

**Population Census Evaluation  
July 2000**

## SUMMARY OF FINDINGS

The Journey to Work Paper evaluates the data quality of the Journey to Work (JTW) questions in the 1996 Census. Two main topics are analysed. The central issue in this paper concerns the processing procedures used to code Workplace Address responses, particularly those that did not include a full workplace address. Secondly the Journey to Work questions themselves are analysed to determine the impact of the question wording on the quality of responses. The main conclusions of the analyses are as follows:

- JTW data can be compromised by poor quality responses to the workplace address question. Although the strategies for dealing with insufficient responses were acceptable, they could be improved by assigning destination zones to responses based on the name of the company that employs the respondent.
- A low proportion of respondents were assigned to the 'State not Further Defined' or the 'Capital City Not Further Defined' non-mappable DZN codes.
- There was a large decrease in the non-response rates from the 1991 Census (12.4% of respondents) to the 1996 Census (3.6% of respondents). This was mainly attributed to an improvement in form design.
- The incidence of respondents reporting the address of a head office rather than their actual workplace address was small. To eliminate this problem completely, it is recommended that, for the 2001 Census, the question be changed to ask for 'Person's Workplace Address' rather than 'Employer's Workplace Address'.
- It was possible to cross-tabulate DZN region by method of transport to work. Unusual combinations (for example a person walking from Perth to work in the Pilbara) were highlighted. A minimal proportion of respondents had answered with such unusual combinations.
- Changes to questions, instructions and data capture criteria influence the quality of JTW data. For example changes to the instructions for those with no fixed workplace address saw a large increase in the number of these responses from 2,005 in the 1991 Census to 182,132 in the 1996 Census.

## CONTENTS

1. INTRODUCTION .....	1
1.1 Collection of Journey to Work Data .....	1
1.2 Quality Issues in Journey to Work Data .....	2
1.3 List of Acronyms used in this Paper .....	2
2. PROCESSING ISSUES .....	3
2.1 Destination Zone (DZN) Coding Procedures .....	3
2.2 Coding of Insufficient Responses .....	4
2.3 State to State Processing Differences .....	7
2.4 Validation of Coding Procedures .....	7
2.5 Processing Strategies for 2001 Census .....	9
3. QUESTION DESIGN ISSUES .....	10
3.1 'Head Office Address' Responses .....	10
3.2 Time Inconsistencies between Journey to Work Questions .....	11
3.3 Respondents with No Fixed Place of Work .....	11
3.4 Question Wording for 2001 Census .....	12
4. CONCLUSIONS AND RECOMMENDATIONS .....	13
APPENDIX 1: 1996 Census Labour Force Status, Occupation, Industry and Journey to Work Questions .....	14

## LIST OF TABLES

TABLE 1: Number of Respondents and Proportion of Applicable Respondents assigned to State/ Capital City Not Further Defined .....	6
TABLE 2: Number of Respondents failing to Answer Workplace Address Question, 1986, 1991 and 1996 Censuses .....	6
TABLE 3: Shift in Working Population after Investigation of Inconsistent Responses .....	8

## 1. INTRODUCTION

Information pertaining to Journey to Work (JTW) has been collected from the Australian Census of Population and Housing since 1971. JTW information helps in the examination of the patterns of respondents' journeys from their home to their place of work. This Working Paper evaluates the quality of the JTW data collected in the 1996 Census.

The most frequent users of Journey to Work information are State Transport Authorities (STAs) because journeys to and from work are concentrated in peak traffic periods and therefore represent the most serious challenge to transport systems. Data are used by STAs and others to identify travel patterns, model fuel usage, forecast public transport patronage and plan alterations to existing transport systems. Thus the data can assist in determining changes to the frequency and routes of public transport, the construction of new, or alteration of existing roads, and the planning of private and public facilities in employment centres. Although STAs perform their own transport surveys, which complement census data, they are the main clients for JTW data and they remain closely involved in the development of JTW aspects of the census. STAs provide the Australian Bureau of Statistics (ABS) with the necessary Destination Zones (DZNs) and are consulted on question design and processing issues.

The users of JTW data are not limited to STAs, as the data are also used to monitor land use and industry density, as well as the economic attachments of those commuting from peri-urban areas.

In the 1996 Census, due to the relative importance of journey and transport issues in densely populated areas, JTW data were analysed for the following eight study areas:

- Sydney/ Wollongong/ Newcastle;
- Melbourne/ Geelong/ LaTrobe Valley/ Bendigo/ Ballarat;
- Brisbane/ Gold Coast/ Tweed Heads/ Cairns/ Townsville;
- Adelaide;
- Perth/ Pilbara;
- Hobart/ Launceston;
- Darwin/ Alice Springs; and
- Canberra/ Queanbeyan.

### ***1.1 Collection of Journey to Work Data***

There is no specific census question addressing Journey to Work. Rather, data were collected in the 1996 Census via Question 1, the respondents' enumeration address, Question 35, 'For the main job held last week what was the employer's workplace address?', and Question 38, 'How did this person get to work on Tuesday 6 August 1996?' (Refer to Appendix 1 for the wording and sequencing of JTW questions in the 1996 Census). JTW data were therefore captured and processed for employed respondents who were enumerated in a JTW study area. This information can be used to map each respondent's departure point and destination and his/her means of transport from one to the other. Note that users could still request information on a Usual Residence basis rather than Place of Enumeration if this information was more relevant to their needs. However, this data was only available at the Statistical Local Area (SLA) level.

Changes in data capture criteria with important implications for time series analysis were evident from previous censuses. In both 1986 and 1991, JTW data were coded only for respondents whose enumeration address was in the same study area as their JTW destination. In 1996, data were coded for all employed persons enumerated in JTW study areas irrespective of their work address. However, inconsistencies appeared in the 1996 data when a person responded from somewhere other than his/her usual residence. Consider a Victorian holidaying in Queensland on census night. This person's place of enumeration (origin zone) was in Queensland, but his/her destination zone was in the Victorian study area so he/she would have been coded as 'Worked Outside Study Area'.

This coding was accurate given the provided information, but misleading given the intention of JTW analysis. As a result of the change in capture criteria in 1996 to code respondents independently of their work address, the number of 'Worked Outside Study Area' responses rose from 2,081 in 1991 to 122,723.

DZN information was also coded for respondents who were not at their usual residence on census night. Thus if a respondent both lived and worked in a JTW study area, and stayed with a friend on census night in the same study area, details of his/her journey and mode of transport may not be reflective of their usual JTW behaviour. However, as previously stated, users could still request JTW information on a Usual Residence basis.

There were also changes made to the wording of the instructions for the workplace address question. Specifically, respondents with no fixed workplace address were instructed to write 'no fixed address' rather than to provide the address of their depot or office. The implications of these changes for data output and time series analysis are discussed later in the paper.

## ***1.2 Quality Issues in Journey to Work Data***

Journey to Work data are subject to the usual quality constraints imposed by a self-enumerated questionnaire. Data quality relies on the ability of respondents to understand each question and answer in the appropriate manner and with an appropriate amount of detail. Collectors have no opportunity to probe a respondent for more information in order to obtain a complete response. Inevitably instances of insufficient responses complicate the processing of the data and can significantly detract from the quality of the resultant information. The first major issue discussed in this paper is the coding and processing procedures used, particularly when respondents provide insufficient information about their place of work. The second main discussion point is an evaluation of question wording in order to maximise the quality of information provided by respondents.

## ***1.3 List of Acronyms Used in This Paper***

JTW - Journey to Work

STA - State Transport Authorities

ABS - Australian Bureau of Statistics

DZN - Destination Zone

SLA - Statistical Local Areas

CD - Collection District

CBD - Central Business District

## **2. PROCESSING ISSUES**

The most important issue in evaluating the Journey to Work (JTW) aspects of the census is the quality of Destination Zone (DZN) data. DZN information provides the endpoint of respondents' journeys to work. In order to appropriately appraise the quality of JTW data, it is important to understand the coding procedures used to ascribe a DZN code to a particular workplace address.

### **2.1 *Destination Zone (DZN) Coding Procedures in 1996***

Each JTW study area comprised a number of four-digit codes in the range of 0001 to 9999 called DZNs. These DZNs are geographical units designed to represent areas with working populations of at least 100 persons. Although DZNs aggregate to Statistical Local Area (SLA) boundaries, they have no relationship to Collection Districts (CDs). DZNs are based upon an area's working population, while CDs are designed according to the number of residential dwellings in the area. Thus an area like a Central Business District (CBD) may contain many DZNs, because the working population is great, but a small number of CDs because few people live there.

Responsibility for defining DZN indexes lies with the relevant STA. This is a mutually advantageous situation, as the ABS can rely upon the local knowledge of STAs and the STAs, as the main users of JTW data, can compose the indexes to reflect their own needs. Before processing JTW information, ABS carefully ensures that universal coverage of JTW study areas is present in the indexes. Specifically, this means that ABS ensures that a valid code exists for every relevant locality or street name, and that a mappable DZN or a dump code (a code used when there is insufficient information to allocate a mappable DZN) exists to receive these poorly defined addresses. Additionally there is the need to confirm that as many possible colloquial names, initials and acronyms are available for each locality, street or building so that the maximum number of valid responses can be coded. Finally it is necessary to check that the list of localities for JTW processing concord to the National Locality Index.

Coding of DZN information took place through Computer Assisted Coding in two stages. Coders were instructed to code DZNs based primarily on the workplace address provided by the respondent. Only if a match could not be made using workplace address information could the DZN coding of that respondent be based upon the provided building or business name if supplied in the index.

The first stage of coding in 1996 used a locality index. Locality, in this case, indicated a city, town, suburb, large employer name or shopping centre. When a coder sought to allocate a DZN code for the stated workplace address (as given in Question 35), they firstly entered the locality and State provided by the respondent. If the entered locality was entirely contained within a DZN, the coder was prompted to assign that DZN code to this response. If the locality overlapped two or more destination zones, the coder was instructed to proceed to Level 2 (street level) coding and to match the street name and number of the workplace address to a DZN code. In CBDs, the entry of a street name alone was frequently insufficient to assign a DZN, as many streets overlapped two or more DZNs. Thus it was critical that respondents provided accurate street number information.

When circumstances arose in which the provided information was ambiguous or required close inspection before a correct DZN could be assigned, coders sent the response to a Query Resolution Team. Such a referral might take place if a respondent provided insufficient workplace address information to be successfully coded, but included an employer's business name. The query resolution team then searched telephone books for the address of that particular business and used the additional information to get the right index and to assign a DZN. The comprehensiveness of JTW data was therefore increased because partially complete responses were assigned to a DZN as if they contained complete information.

The use of Query Resolution also allowed for frequent updating of indexes. Common examples of changes made during processing involve the updating of indexes to include an acronym or colloquialism for a particular locality, street, building or employer. For example, in 1996 the DZN index entry for the Civil Aviation Authority was updated to include the acronym CAA.

## **2.2     *Coding of Insufficient Responses***

Responses to Question 35, 'For the main job held last week, what was the employer's workplace address?' are frequently of insufficient detail to assign a DZN code with complete accuracy. For example, the NSW Transport Study Group reported from a 1991 survey that 23 per cent of respondents answered in insufficient detail to be coded with certainty. One quarter is widely accepted as a reasonable estimation of imprecise responses, although as many as 45 per cent of respondents in a 1989 Census Test failed to provide street number details. The specifics of an employer's workplace address (particularly street number) are not at the forefront of many respondents' minds. Further, imprecise responses in 1996 JTW data may have been more frequent as the workplace address question was located towards the end of the census form, by which point the ability of fatigued or uninterested respondents to answer is diminished. This represents a serious issue, particularly in CBDs and areas of high density employment where address specifics are required.

Given the frequency of imprecise responses, the ability of the indexes to cope with the range and quality of employer addresses provided by respondents determines to a large extent the quality of JTW data.

### **2.2.1     *Incomplete Address Responses***

Procedurally, there are two means of dealing with imprecise responses. The first is to assign a mappable DZN to every incomplete response based upon a decision rule during the index building process. Such an action gives the illusion of uniformly high quality data, which may be more acceptable to users. Alternatively, imprecise responses may be left as stated and assigned to various non mappable 'dump' codes. Thus the data might be incomplete, but accurately reflects the responses provided by respondents. In 1996, both types of methodology were used. To understand the strategies used to process imprecise responses, NSW is used in the case study below. Note that other States used similar approaches but did not assign to mappable zones and dump code in exactly the same proportions.

In NSW, if a respondent worked in a CBD region and provided insufficient address information, that person would be assigned to a non-mappable dump zone. The DZN codes for CBD dump zones fell in the range 6000-7999. As an example, 'George St' in

Sydney runs through a number of different mappable DZNs. If a respondent describes his/her workplace address as merely 'George St, Sydney', with no street number information, then that respondent cannot be classified with any accuracy as working in any of the DZNs that George St runs through. Therefore that respondent will be assigned to the appropriate dump zone: 'George St, no number'.

In 1996, 71,534 respondents in the NSW Study Area (3.7 per cent of the Study Area's applicable population) were allocated to non-mappable dump zones in CBD regions. For comparison purposes, 33,080 respondents (or 2.1 per cent) of the Victorian Study Area and 18,608 respondents (or 1.7 per cent) of the Queensland Study Area were dump coded in CBD regions.

Outside CBD areas, respondents who provided insufficient information were assigned to a mappable DZN based upon an imputation rule devised by the Transport Data Centre of the NSW Department of Transport. For example, a respondent might have provided his/her workplace address as 'Hurstville Rd, Hurstville' (which overlaps three DZNs) with no street number information. The NSW RTA 'built' an imputation rule into the indexes which automatically assigned 'no number' responses to one of the three mappable zones, in this case DZN 0265. ABS relies upon the major data users (the STAs) to use their local knowledge to appropriately allocate these codes.

Users should be aware that when a response was allocated to a DZN based on these definitions, it was indistinguishable from a complete response, therefore it gave a false impression of the completeness of the respondent's answer. Secondly, since these imputations became part of the normal DZN data it is not possible to know how many responses were treated in this way.

### *2.2.2 Responses of State or Capital City Only*

The most frequently used dump codes were those to which a respondent was assigned when he/she provided some street level information, but failed to provide an exact street number. Less frequent were those in which only the locality of a respondent's address was provided, with no street name or number information. These dump codings have been discussed above. However there were two other common types of insufficient response.

The most general level of insufficient response occurred where a respondent provided only the State or Territory in which he/she worked (for example, simply answered 'New South Wales' or 'ACT'). In such an instance it was impossible to assign the respondent to a DZN. These responses were therefore assigned to DZN code 9979, 'State or Territory Not Further Defined'. Secondly, there existed an available code when a respondent provided only the Capital City in which he/she worked (for example 'Sydney' or 'Melbourne' with no street level detail). 'Capital City Not further Defined' had the DZN code of 9985.

Table 1 shows the number of respondents who were coded in 'State Not Further Defined' or 'Capital City Not Further Defined' for each JTW Study Area and the proportion of respondents who answered in this way. Note that such responses comprised a low proportion of the total number of responses.



**TABLE 1: NUMBER OF RESPONDENTS AND PERCENTAGE OF APPLICABLE RESPONDENTS ASSIGNED TO STATE/ CAPITAL CITY NOT FURTHER DEFINED**

	<i>Frequency</i> <i>'State NFD'</i>	<i>%</i> <i>'State NFD'</i>	<i>Frequency</i> <i>'Capital City NFD'</i>	<i>%</i> <i>'Capital City NFD'</i>
NSW	18,717	1.0	8,804	0.5
VIC	15,718	1.0	6,632	0.4
QLD	10,765	1.0	2,057	0.2
SA	2,214	0.5	3,060	0.7
WA	3,713	0.6	1,529	0.3
TAS	812	0.5	2,230	1.4
NT	528	0.9	1,254	2.2
ACT	842	0.5	7,072	4.3

### 2.2.3 Non-Responses to Workplace Address Question

In addition to the figures quoted in Table 1, 222,661 employed respondents over the age of 15 across Australia were coded to 'Not Stated' for failing to answer the employer's workplace address question altogether (3.6 per cent of the employed population). This represented a significant improvement on the 1986 (549,855 or 11.0 per cent of applicable respondents) and 1991 (678,589 or 12.4 per cent of applicable respondents) non-response rates. These data can be seen in Table 2.

The reduction in non-response rate from 1991 can be attributed to an improvement in form design. In the 1991 Census, some respondents answered that they had a full or part-time job (Question 30) but ignored the sequencing instruction 'Now go to 32'. They then responded to Question 31 that they had not actively looked for work in the previous 4 weeks and were sequenced to Question 40, thus not completing the workplace address question. The likelihood of this error occurring was compounded by the fact that Questions 30, 31 and 32 were at the bottom of a page on the census form and the fact that the sequencing instructions required the respondents to skip the subsequent page entirely. Therefore respondents would not read the workplace address question at all, and would not realise that this question may apply to them. To reduce this confusion, the 'Looking for Work' and 'Hours Worked' questions were moved to the end of the employment related questions for the 1996 Census (see Appendix 1 for the question wording and sequencing of JTW questions for the 1996 Census).

**TABLE 2: NUMBER OF RESPONDENTS FAILING TO ANSWER WORKPLACE ADDRESS QUESTION, 1986, 1991 AND 1996 CENSUSES**

<i>Census Year</i>	<i>No. Of Applicable Respondents</i> <i>Coded to 'Not Stated'</i>	<i>% of Total Applicable</i> <i>Respondents</i>
1986	549,855	11.0
1991	678,589	12.4
1996	222,661	3.6

A further change was made to the processing of census forms which may also have helped to reduce the non-response rates. Specifically, prior to 1996, responses which were uninterpretable to coders were coded as 'Not Stated'. However in 1996 such cases were more frequently referred to query resolution. This increased the likelihood that these answers would be treated as a stated response and some level of DZN allocated.

### **2.3     *State to State Processing Differences***

As discussed above, States and Territories in 1996 used different indexes to assign responses to DZNs. Specifically, some States were more likely to ‘dump code’ incomplete responses, while others were more likely to assign insufficient responses to a mappable DZN based upon a decision rule built into the indexes. However differences between States were also present in the processing strategies used.

Users should be aware that processing of 1996 JTW data took place State by State. Thus processing began with smaller States and Territories (Tasmania, the Northern Territory and the Australian Capital Territory), before concluding with the larger States, Queensland, New South Wales and Victoria. Due to this sequential processing there were systematic differences in processing between States, as the procedures were refined over time. Broadly speaking, these changes involved a simplification of procedures based on system and procedural improvements introduced over time.

In larger States, query resolution found difficulties when respondents entered a business name but neglected to provide a locality. Due to employment density, Query Resolution required a locality in order to know which city to search. Thus coders in the larger States were instructed to include the locality of the respondent’s usual residence when referring to query resolution.

Further changes in the use of query resolution were also made. When processing smaller States queries were referred if locality information, but not street information, was provided, or if there was street name but no number. This strategy was changed for the larger States. Typically there were only two DZNs involved (that is, the respondent’s workplace address could only be in one of two DZNs) so the non-mappable DZN code was deemed to be sufficient with no query resolution necessary. This was decided because the dump code provides some information about the respondent’s workplace destination, while the loss of specific information was considered justified in terms of the reduction in processing delays.

### **2.4     *Validation of Coding Procedures***

Once JTW processing has been completed, a number of strategies are available to validate the overall integrity of the data. At the most basic level this involves a confirmation that respondents under the age of 15 have been coded to Not Applicable, or that all respondents coded (for example) to the Perth/Pilbara study area have their enumeration address in Western Australia. More complicated cross tabulation can also reveal unusual combinations of enumeration address, DZN and mode of travel to work. A region like the Queensland study area is particularly vulnerable to exposures of this sort because it is made up of a number of distinct geographical regions rather than contiguous areas. For example, responses in Townsville, with workplace addresses in Brisbane, warrant closer inspection.

Careful analysis of respondents whose responses cause such unusual combinations can reveal systematic miscoding through shortcomings in indexes. For example, there are a number of localities which share a name with another locality in the same State, such as Cooks Hill in Newcastle (Postcode 2300) and Cooks Hill in Yass (2582). Often, if a respondent does not provide a postcode, query resolution is required to correctly identify the locality to which the respondent refers. Similarly there is frequently difficulty

when a respondent gives his/her employer's address as 'Westfield Shopping Centre', as there is a number of shopping centres of this kind in most States. Many of these inconsistent responses can also be ascribed to coder error, misinterpretation of JTW questions, or inaccurate information provided by the respondent. Further discussion of these errors is given in Chapter 3.

From the number of identified 'inconsistent' responses, a sample was closely investigated. As a result of this investigation, it became possible to adjust the overall DZN figures to compensate for any errors. For example, 104 respondents were coded as living in Wollongong and working in Newcastle. A sample of 25 of these respondents was investigated, revealing that 7 of them (28 per cent) should have been coded as working in Wollongong. There was no evidence that the remaining 18 responses were correct or incorrect. Applying this proportion to the total number of 104 inconsistent responses, 29 respondents were adjusted to working in Wollongong. However, the same principles applied in reverse (i.e. living in Newcastle and working in Wollongong) suggested that 23 persons should be shifted from working in Wollongong. Therefore the net change from Newcastle to Wollongong is 6 people. This is less than 0.01 per cent of the working population of either city.

Table 3 illustrates the net adjustment, as a percentage of the working population, for four study areas which often contained inconsistent responses. The most important conclusions to be drawn from Table 3 are that the overall adjustments required to be made to the data are very small, and that the integrity of 1996 JTW data is robust.

**TABLE 3: SHIFT IN WORKING POPULATION AFTER INVESTIGATION OF INCONSISTENT RESPONSES**

<i>Removed From</i>	<i>% Change in Working Population</i>	<i>Moved To</i>	<i>% Change in Working Population</i>
Newcastle	Less than -0.01	Wollongong	Less than +0.01
Pilbara	-0.54	Perth	+0.02
Hobart	-0.11	Devenport/ Launceston	+0.12
Darwin	-0.02	Alice Springs	+0.08

Victoria in particular performed extensive validation of JTW data. By cross tabulating DZN region by method of travel to work 555 unusual combinations were highlighted. After investigation, it was concluded that only 20 per cent of these respondents (or 111 persons) were incorrectly coded. Once again this represents an insignificant percentage of the total applicable respondents in Melbourne (1,592,748).

Finally, feedback was accepted from STAs, who frequently examined the distribution of industry by DZN for data anomalies. In 1996, for example, high figures for health sector employment were noted in Adelaide DZN 125, although only one nursing home and one spastic centre could conceivably employ health sector workers in this DZN. However, the neighbouring DZN 121 contained relatively few health workers, despite the presence of the Queen Elizabeth Hospital. This apparent miscoding of respondents employed at the hospital was attributed to the fact that the hospital was located in 'Woodville South' (DZN 121) but was frequently reported by respondents as being in 'Woodville' (and therefore coded to DZN 125).

## **2.5     *Processing Strategies for 2001 Census***

STAs, anxious to reduce the proportion of respondents assigned to non-mappable dump codes or allocated to mappable DZNs based on a decision rule, have expressed a desire that alternatives to workplace address coding be found. Rather than using the name and number of the street in which respondents work, it was suggested that the assignment of DZNs should increasingly use a facilities index, or index of business names. Thus in 2001, if no street name or number (where required) can be entered, coders will be prompted each time to enter a business name and to allocate a DZN based upon this list of business names.

This strategy will also reduce the likelihood of widespread miscodings of an institution to an incorrect DZN. For example, the Queen Elizabeth Hospital (see Section 2.4) would be listed on the facilities index and would be coded to the *correct* DZN in Woodville South.

The increased use of facility coding, as well as the use of Intelligent Character Recognition technology to scan completed census forms is also expected to reduce the workload of census processors. This is expected to have important implications in saving time and increasing the accuracy of JTW coding.

### 3. QUESTION DESIGN ISSUES

Respondents commonly provide incorrect or insufficient information when answering the employer's workplace address question. It is therefore relevant to consider the wording of JTW questions so that the maximum amount of accurate information can be obtained from respondents. Examples follow in which the wording of JTW questions influenced the quality of obtained data. Additionally, some possible changes to question design that may improve the quality of JTW data will be discussed.

#### 3.1 *'Head Office Address' Responses*

The employer's workplace address question served two distinct functions in 1996. The first was to provide a DZN code as part of the JTW data, the second was to assist with the coding of industry by matching the name and address of the business employer against the ABS business register. In some respects, the workplace address question was worded in the interests of industry coding to the detriment of JTW data.

The question relating to the respondents' workplace address (Question 35) asked 'For the main job held last week, what was the employer's workplace address?' (See Appendix 1 for sequencing instructions). This workplace address was used to match businesses with their industry on the ABS Business Register. However this question also sought the address of the place where respondents went to work in the week prior to the census. Due to the question wording, a number of respondents replied not with the address of their actual place of work, but with the address of their employer's head office. Responding with the address of a head office was particularly common in the case of large employment organisations with a centralised authority (for example police officers or teachers).

Potential data anomalies resulting from respondents providing a head office address may not have been immediately obvious. Coders were unlikely to have raised a query, since the address was complete and a DZN easy to assign. Also, the employer's head office was commonly in an area of high density employment in which persons incorrectly reporting that as their place of work would be difficult to detect. However by cross-tabulating enumeration address, DZN and method of travel to work, some unlikely combinations could be revealed.

For example 502 West Australian respondents were coded as living in the Pilbara and commuting to work in Perth - an unlikely situation, although the potential remained that some of these workers (probably oil workers or miners) were periodically flown from their homes in Perth to the Pilbara for work. A sample of these 502 respondents was then selected for closer analysis. Analysis suggested that 30 per cent of respondents were, indeed, being flown from Perth to Pilbara, 36 per cent (or 180 respondents) could be attributed to coder's error and 34 per cent (or 170) of these respondents were mistakenly reporting the address of their employer's head office rather than their place of work. In a similar situation, 722 respondents were identified as commuting from Devonport or Launceston to Hobart. Closer inspection of a sample of these respondents suggests that 21 per cent (or 151 respondents) were incorrectly reporting the address of a depot or of their employer's head office.

Though these figures (180 respondents for Perth and 151 for Hobart) do not reflect the total number of respondents who incorrectly stated their head office address, it remains a

small proportion of the 577,829 applicable respondents in the Perth/ Pilbara region and the 162,662 respondents for Hobart/ Launceston. Thus the overall quality of the data should not be inferred to have been seriously compromised.

### **3.2 *Time Inconsistencies between Journey to Work Questions***

Further data anomalies were caused by differences in the time periods used in JTW questions. Firstly, the place of enumeration was the residence at which respondents were enumerated on census night and did not necessarily represent their usual residential address. In fact, it did not necessarily represent the address from which they departed for work on census day. Secondly the method of transport to work question asked 'How did the person get to work on Tuesday, 6 August 1996?' (in most instances this will refer to census morning). Thirdly the employer's address referred not to the day of the census but to the employer's address the previous week.

There is an apparent methodological inconsistency in mapping respondents journey to work, when their point of departure was based upon a particular night, their method of transport based (usually) on that morning and their destination was based upon the previous week. A number of problems result from this inconsistency. Instances where a respondent has changed jobs since last week, or if they used an unusual method of transport work on census morning, or if their place of enumeration was not the same as their departure point on census morning, affect data quality. The main issue is not the *absolute* date on which the data should be collected, but rather the *relative* dates of the three pieces of JTW data. Given inconsistent dates, JTW data may not reflect a 'snapshot' of JTW behaviour.

However two issues should be emphasised. Firstly JTW data are still likely to reflect a respondent's actual JTW behaviour. The probability of a person's employment situation changing during this one week period remains low, and the few respondents whose situations did change should have little impact upon the overall quality of data in a DZN. Only in the case of the closing down of a large employment centre in the week before the census would there have been a systematic contamination of JTW data. However such an event would be well known to STA's. Similarly, most respondents will have departed for work from the same address as their place of enumeration. Although JTW data are still reliable, users should be aware of the time inconsistencies.

Secondly it is important to understand the practical constraints which necessitate the inconsistencies between JTW dates. The rationale behind the employer's address referring to the previous week rather than census day was so that census labour force data could be compared with other ABS labour force data. Enumeration address is used as a point of departure because it is impractical to insert a further question asking the address from which the respondent left for work on census day. Method of Transport to Work is likely to refer to census morning because journeys *to* work are the source of interest and the reference point is therefore likely to be census morning.

### **3.3 *Respondents with No Fixed Place of Work***

In 1996 a notable change in the instructions for those with no fixed place of work (for example, taxi drivers or couriers) took place. Such a change had important implications for the type of data that was collected from JTW questions, and also for JTW data analyses, particularly time series comparisons of 1991 and 1996 data.

In 1991, respondents with no fixed place of work were instructed to provide the address of the depot or office of their employer. However in 1996 these instructions were changed and respondents were instructed to write 'No Fixed Address'. Such instructions were conceptually similar to the 1986 Census in which these respondents were instructed to write 'Not Applicable'. Such a change was necessary because the 1991 wording contaminated JTW data since the address of this office or depot was unlikely to be the respondent's work destination. Thus, from a JTW perspective, the change represented a great improvement to the wording of the question. Most importantly, the change in instructions meant that the data collected for JTW, particularly for respondents with no fixed workplace address, was closer to the actual journey patterns of respondents.

As a result of the change of instructions from 1991 to 1996, 'No Fixed Address' responses rose significantly from 2,005 in 1991 to 182,132 in 1996. Attention should also be drawn to the fact that this created a generalised decrease in employment density across mappable DZNs due to people of no fixed address, who were no longer assigned a mappable code.

### **3.4     *Question Wording for the 2001 Census***

In an attempt to minimise the number of respondents reporting the address of a head office, rather than their actual workplace destination, the wording of the Workplace Address question will be changed from 'Employer's Workplace Address' to 'Person's Workplace Address'. This change is likely to remove any ambiguity surrounding this question.

Secondly it was noted that different time reference points were present for the origin (census night), destination (last week) and method of travel to work (census morning). However it is necessary that the workplace address question refer to last week so that census labour force information can be compared to other ABS labour force data. Therefore it is not practicable to change these reference points for 2001.

Finally, although instructions in 1996 to persons with no fixed workplace address were an improvement on 1991, further improvements will be made. For the 2001 Census, persons with no fixed place of work will be instructed to respond to the workplace address question with 'no fixed address' unless the person usually travels to a depot to start work, in which case he/she should provide the address of that depot. This captures the maximum possible JTW information because the journeys to work of those with no fixed address who regularly travel to the same destination to begin work are recorded. However, respondents who genuinely have no fixed destination point in journeying to work are dump coded.

#### 4. CONCLUSIONS AND RECOMMENDATIONS

This paper has examined the quality of Journey to Work (JTW) data from the 1996 Census. It has been concluded that the JTW data are of a high standard, and are close to a true reflection of the actual journeys to work made by the population in the week before census date.

In conclusion:

- JTW data can be compromised due to poor responses to the workplace address question. This mainly involves respondents failing to provide the street name or number of their workplace. The proportion of respondents who were 'dump coded' in the 1996 Census was acceptable given the nature of the responses, but is still worthy of attention. A means of reducing the amount of dump coding might be to assign respondents to DZNs based on the name of the company that employs them. In this way the impact of insufficient address information and consequently the number of dump coded respondents would be reduced. The workload of coders would also be reduced.
- JTW data might also be compromised in the indexes of Destination Zones, to which respondents are assigned. However it was observed that pre-processing validation sought the universal coverage of these indexes, and also that as processing took place these indexes were constantly updated to be as accurate as possible. Thus sufficient validation is made of the DZN indexes provided by State Transport Authorities.
- In an attempt to eliminate cases where respondents report the address of a head office rather than their workplace address, this question will be changed so that it asks for the 'Person's Workplace Address' rather than 'Employer's Workplace Address' in the 2001 Census.
- JTW data might also be contaminated in the processing procedures and the strategies used for assigning respondents to DZN codes. However, coders used a series of clear, specific procedures to code responses. A number of validation procedures remained in place to isolate errors. After analysing inconsistent responses in output validation, it was found that these constituted a small proportion of the overall population, so reflected consistent coding procedures.
- A final salient issue was the way in which the question instructions and the capture criteria influenced the data obtained. Specifically, the changes to the 'No Fixed Address' instructions in Question 35 saw the number of these responses rise from 2,005 in the 1991 Census to 182,132 in the 1996 Census. Similarly changing the capture criteria so that responses were coded even if the workplace address was in a different study area to the enumeration address, led to an increase in the number of 'Worked Outside Study Area' codings from 2,081 in 1991 to 122,723 in 1996.







## Reference List

New South Wales Department of Transport (1998) *Journey to Work- User Guide*.

## Census Working Papers

- 96/1 1991 Census Data Quality: Income
- 96/2 1996 Census Form Design Testing Program
- 96/3 1996 Census of Population and Housing: Digital Geography Technical Information Paper
- 97/1 1996 Census: Homeless Enumeration Strategy
- 99/1 1996 Census: Industry Data Comparison
- 99/2 1996 Census: Labour Force Status
- 99/3 1996 Census Data Quality: Housing
- 99/4 1996 Census: Review of Enumeration of Indigenous Peoples in the 1996 Census
- 99/5 2001 Census: Indigenous Enumeration Strategy Draft
- 99/6 1996 Census Data Quality: Occupation

If you would like a copy of any of these papers, or have any other queries, please contact Rosa Gibbs on (02) 6252 5942 or Email: [rosa.gibbs@abs.gov.au](mailto:rosa.gibbs@abs.gov.au)