

WORLD BANK FIRMS SURVEY LARGE FIRM SURVEY

Background on the Material available on this site

Using the data provided

The data set at your disposal is composed of:

- the 8 modules which form the survey questionnaires. The heading of each of these documents specify the main theme of the module (e.g. GENERAL ISSUES, PRODUCTION RELATED ISSUES, etc.). This is specified on the first page of each individual document together with the respondent that was approached by the field workers. These 8 documents are in Word for Windows.
- The database of all the information gathered in the survey towards the *large* firms (with at least 50 employees) interviewed in the Greater Durban Metropolitan Area (the GDMA) ie the eThekweni Municipality of the Republic of South Africa. This is supplied to you in *SPSS v. 11.5 for Windows* format.

Though the database can be converted into *Stata* via *Stat Transfer*, the SPSS version allows you to refer directly to the relevant question number of the questionnaire. In particular, the information contained in the “Variable View” worksheet – which has been copied into a separate document complemented by the coding instruction document - specifies the variable names, their format, the value labels, as well as the unit of measurement of the variable when applicable. This is because the variable names use the question number. For example, the variable “q1_14a” refers to module 1, question 1-14. For this question, the fieldworkers asked the respondents to

*Circle (O) the **three** most important contributions that **central government authorities** can make towards promoting investment and local economic development in the DMUC area and rank them in order of importance.*

The data was coded q1_14a with “a” referring here to the rank – the most important (“b” being the second most important etc.), with the value label from the type of contribution that was reported as the “most important” e.g. value 1, described in the value label as corresponding to “policy stability”.

Note that the information provided in the “Variable View” worksheet (or the equivalent document provided in Word) is important as there is no ‘systematic’ labelling of rows and columns. For instance with q2_10, *a1*, *a2* etc. refer to information specified in rows. In contrast for q2_11, *a1* and *a2* refers to two main columns and not rows. Moreover, whilst *a1* and *a2* refers to information gathered for 2000, *b1* and *b2* for this question refer to the same information specified for 2001. However, generally, letters – that is *a*, *b*, *c*, etc. - denote the rows and numbers – that is *1*, *2* etc. – denote the column of the information matrix.

Finally, following entry error, data were recorded by staff at the School. These variables start with “*r*”. As this happened for all the questions of q3_9, these have become r3_9 in the dataset. Since the researchers have only processed some and not all (!) the data available, we recognise that the data still contain some errors. Also included in the dataset are variables which researchers computed at the School for their own purposes. We have kept some of these in the database as we believe that these might be helpful to further research on the city, South Africa and beyond.

Background information

The large firm survey is the result of an agreement entered by the Durban Unicity Council in 2000 (with USAID) to fund a World Bank technically-supported survey of firms in the greater Durban region. The survey, carried out between May 2002 – April 2003, involved 22 fieldworkers. These were managed by the Bureau of Market Research (BMR) at Unisa.

To assist fieldworkers in selecting sample units (firms) the World Bank provided the Bureau of Market Research (BMR) at Unisa with various sample frames (listings of firm names). These sample frames were used to contact firms randomly by telephone in order to set up appointments with the managing directors, managers or owners of the firms. Because no sample frame is comprehensive enough to include all firms operating in the Durban Metropolitan Area, various sample frames had to be utilized. The study was constructed to stratify industry by type, employment size group and geographic area. The sample frames had some limitations in this regard: they lacked employment size group classifications and they had a limited number of firms for certain sectors and they showed geographic location problems. In some cases information on firms was outdated (i.e. non-existing or a change in contact details).

