Overview

Identification

**ID NUMBER**

zaf-ctcc-ctrlfd-2000-2006-v1

Version

**VERSION DESCRIPTION**

v1: Data accessible through the Secure Data Service

**PRODUCTION DATE**

2015-06-29

**NOTES**

Secure variables available in this dataset include:

- Post code for firms' street address (trinstpcde)
- Magisterial District where the firm is located (magdistric)
- Cadastral number of the firm from a separate GIS survey (cadastralno)

Overview

**ABSTRACT**

Until 2006 metropolitan and district councils in South Africa, which were previously called Regional Services Councils (RSC), were permitted to raise revenue by taxing firms that operated within the council area. The City of Cape Town taxed firms based on their turnover and wage bill but also used the administration of the RSC levy, as the tax was called, to create an administrative dataset of firms. The city used this data to calculate Gross Geographic product and produce a number of reports on the local economy (cf City of Cape Town 2001).

The City undertook a survey of all firms on the RSC tax database in 2000 and linked the firms to location based GIS data. The RSC data is thus a mixture of administrative and survey data. The dataset covers the period 2000 to 2006.

In theory any enterprise employing at least 1 worker or with a revenue of R10 000 a year was supposed to pay the RSC levy and thus be included in the database. Thus this dataset should be a census of all formal firms operating with the city of Cape Town during the period covered, except the very smallest self-employed operators. In practice, however those familiar with the RSC have said that there was evasion of the tax, with a possible 30% of firms evading.

Around two thirds of the active firms in 2000 responded to this survey. Some of this survey data (particularly employment) was supposed to be updated every year but in practice we only have useable survey data for the first year of the panel. Much of the employment data was imputed in subsequent years and cannot be used. New entrants were captured in the database but some of the survey information is not available for these firms.

The project to create this research dataset was made possible by an exploratory grant obtained by Andrew Kerr and Martin Wittenberg of DataFirst from the Private Enterprise Development in Low-Income Countries (PEDL) research initiative. PEDL is a joint research initiative of the Centre for Economic Policy Research (CEPR) and the UK Department For International Development (DFID). It aims to develop a research programme focusing on private-sector development in low-income countries.

Scope

**NOTES**

Data collected every year for the firm levy database included the firm type, wage bill and revenue of each firm, as well as whether the firm had exited and when this exit took place.
The 2000 survey collected data on age and size of firm, exporter status and use of IT in the firm, as well as data on sub-contracting and outsourcing.

Coverage

GEOGRAPHIC COVERAGE
The data covers the Metropolitan area of Cape Town. The data is available with firm-level GPS coordinates.

Producers and Sponsors

PRIMARY INVESTIGATOR(S)

<table>
<thead>
<tr>
<th>Name</th>
<th>Affiliation</th>
</tr>
</thead>
<tbody>
<tr>
<td>Andrew Kerr</td>
<td>DataFirst, University of Cape Town</td>
</tr>
</tbody>
</table>

OTHER PRODUCER(S)

<table>
<thead>
<tr>
<th>Name</th>
<th>Affiliation</th>
<th>Role</th>
</tr>
</thead>
<tbody>
<tr>
<td>Cape Town City Council</td>
<td></td>
<td>Original collectors and depositors of the data</td>
</tr>
</tbody>
</table>

FUNDING

<table>
<thead>
<tr>
<th>Name</th>
<th>Abbreviation</th>
<th>Role</th>
</tr>
</thead>
<tbody>
<tr>
<td>Private Enterprise Development in Low Income Countries</td>
<td>PEDL</td>
<td>Funder</td>
</tr>
</tbody>
</table>

Metadata Production

METADATA PRODUCED BY

<table>
<thead>
<tr>
<th>Name</th>
<th>Abbreviation</th>
<th>Affiliation</th>
<th>Role</th>
</tr>
</thead>
<tbody>
<tr>
<td>DataFirst</td>
<td></td>
<td>University of Cape Town</td>
<td>Metadata producer</td>
</tr>
</tbody>
</table>

DATE OF METADATA PRODUCTION
2017-03-08

DDI DOCUMENT VERSION
Version 2
Sampling

No content available
Questionnaires

No content available
Data Collection

Data Collection Dates

<table>
<thead>
<tr>
<th>Start</th>
<th>End</th>
<th>Cycle</th>
</tr>
</thead>
<tbody>
<tr>
<td>2000</td>
<td>2006</td>
<td>N/A</td>
</tr>
</tbody>
</table>

Data Collection Mode

Other [oth]

Data Collection Notes

Data was collected by the Cape Town City Council as administrative records and in 2000 with a survey. Creating this research dataset was made possible through an Exploratory Research Grant (ERG) for the project "Exploring Dynamics in South African firms" from Private Enterprise Development in Low Income Countries (PEDL).

Data Collectors

<table>
<thead>
<tr>
<th>Name</th>
<th>Abbreviation</th>
<th>Affiliation</th>
</tr>
</thead>
<tbody>
<tr>
<td>Cape Town City Council</td>
<td>CTCC</td>
<td></td>
</tr>
</tbody>
</table>
Data Processing

No content available
Data Appraisal

No content available
File Description
Variable List
rscdataforSRDC-s12

Content
Cases 352789
Variable(s) 21
Structure Type: Keys: ()

Version
Producer
Missing Data

Variables

<table>
<thead>
<tr>
<th>ID</th>
<th>Name</th>
<th>Label</th>
<th>Type</th>
<th>Format</th>
<th>Question</th>
</tr>
</thead>
<tbody>
<tr>
<td>V19</td>
<td>firmid</td>
<td>firm identifier</td>
<td>contin</td>
<td>numeric</td>
<td></td>
</tr>
<tr>
<td>V20</td>
<td>year</td>
<td>Financial year, running from July 1999 to June 2000 etc</td>
<td>discrete</td>
<td>numeric</td>
<td></td>
</tr>
<tr>
<td>V21</td>
<td>surveyed</td>
<td>dummy, =1 if firm was surveyed in 2000/2001</td>
<td>discrete</td>
<td>numeric</td>
<td></td>
</tr>
<tr>
<td>V22</td>
<td>weight</td>
<td>Weight to account for survey non-response</td>
<td>contin</td>
<td>numeric</td>
<td></td>
</tr>
<tr>
<td>V23</td>
<td>revenue</td>
<td>firm revenue in the financial year</td>
<td>contin</td>
<td>numeric</td>
<td></td>
</tr>
<tr>
<td>V24</td>
<td>wagebill</td>
<td>firm wage bill in the financial year</td>
<td>contin</td>
<td>numeric</td>
<td></td>
</tr>
<tr>
<td>V25</td>
<td>siccode</td>
<td>SIC code of major product/good/service provided</td>
<td>discrete</td>
<td>character</td>
<td></td>
</tr>
<tr>
<td>V26</td>
<td>totemploy</td>
<td>Total number of employees within the city, 2000 only</td>
<td>contin</td>
<td>numeric</td>
<td></td>
</tr>
<tr>
<td>V27</td>
<td>trinpcde</td>
<td>Post code for firm's street address</td>
<td>contin</td>
<td>numeric</td>
<td></td>
</tr>
<tr>
<td>V28</td>
<td>magdistric</td>
<td>Magisterial district where firm is located</td>
<td>discrete</td>
<td>character</td>
<td></td>
</tr>
<tr>
<td>V29</td>
<td>typebusine</td>
<td>Type of business</td>
<td>discrete</td>
<td>numeric</td>
<td></td>
</tr>
<tr>
<td>V30</td>
<td>estypeorg</td>
<td>Type of organisation</td>
<td>discrete</td>
<td>character</td>
<td></td>
</tr>
<tr>
<td>V31</td>
<td>exportgoods</td>
<td>Does the firm export goods, from survey</td>
<td>discrete</td>
<td>character</td>
<td></td>
</tr>
<tr>
<td>V32</td>
<td>exportservices</td>
<td>Does the firm export services, from survey</td>
<td>discrete</td>
<td>character</td>
<td></td>
</tr>
<tr>
<td>V33</td>
<td>essubconctr</td>
<td>Does company usually sub-contract or outsource work</td>
<td>discrete</td>
<td>character</td>
<td></td>
</tr>
<tr>
<td>V34</td>
<td>cadastralno</td>
<td>cadastral number from separate GIS survey</td>
<td>discrete</td>
<td>character</td>
<td></td>
</tr>
<tr>
<td>V35</td>
<td>liquiddatestr</td>
<td>Date of liquidation</td>
<td>discrete</td>
<td>character</td>
<td></td>
</tr>
<tr>
<td>V36</td>
<td>suspendatestr</td>
<td>Date of suspension</td>
<td>discrete</td>
<td>character</td>
<td></td>
</tr>
<tr>
<td>V37</td>
<td>exit</td>
<td>exit indicator based derived from wage bill+turnover</td>
<td>discrete</td>
<td>numeric</td>
<td></td>
</tr>
<tr>
<td>V38</td>
<td>age</td>
<td>Firm age, from survey</td>
<td>contin</td>
<td>numeric</td>
<td></td>
</tr>
<tr>
<td>V39</td>
<td>entry</td>
<td>an indicator of entry derived from wage bill+turnover</td>
<td>discrete</td>
<td>numeric</td>
<td></td>
</tr>
</tbody>
</table>
### firm identifier (firmid)
**File:** rscdataforSRDC-s12

**Overview**
- **Type:** Continuous
- **Format:** numeric
- **Width:** 6
- **Decimals:** 0
- **Range:** 11-805372

| Valid cases | 352788 |
| Invalid | 1 |
| Minimum | 11 |
| Maximum | 805372 |
| Mean | 223638.5 |
| Standard deviation | 106749.5 |

### Financial year, running from July 1999 to June 2000 etc (year)
**File:** rscdataforSRDC-s12

**Overview**
- **Type:** Discrete
- **Format:** numeric
- **Width:** 4
- **Decimals:** 0
- **Range:** 2000-2006

| Valid cases | 352789 |
| Invalid | 0 |

### dummy, =1 if firm was surveyed in 2000/2001 (surveyed)
**File:** rscdataforSRDC-s12

**Overview**
- **Type:** Discrete
- **Format:** numeric
- **Width:** 1
- **Decimals:** 0
- **Range:** 0-1

| Valid cases | 35992 |
| Invalid | 316797 |

### Weight to account for survey non-response (weight)
**File:** rscdataforSRDC-s12

**Overview**
- **Type:** Continuous
- **Format:** numeric
- **Width:** 1
- **Decimals:** 0
- **Range:** 1-4

| Valid cases | 140799 |
| Invalid | 211990 |
| Minimum | 1 |
| Maximum | 4 |
| Mean | 1.5 |
| Standard deviation | 0.3 |

### firm revenue in the financial year (revenue)
**File:** rscdataforSRDC-s12

**Overview**
- **Type:** Continuous
- **Format:** numeric
- **Width:** 11
- **Decimals:** 0
- **Range:** 0-13481176518

| Valid cases | 339522 |
| Invalid | 13267 |
| Minimum | 0 |
| Maximum | 13481176518 |
| Mean | 5903078.9 |
| Standard deviation | 74225106.2 |
firm wage bill in the financial year (wagebill)
File: rscdataforSRDC-s12

Overview
Type: Continuous
Format: numeric
Width: 10
Decimals: 0
Range: 0-2839252425
Valid cases: 339545
Invalid: 13244
Minimum: 0
Maximum: 2839252425
Mean: 1014276.3
Standard deviation: 15862033

SIC code of major product/good/service provided (siccode)
File: rscdataforSRDC-s12

Overview
Type: Discrete
Format: character
Width: 5
Valid cases: 352789
Invalid: 0

Total number of employees within the city, 2000 only (totemploy)
File: rscdataforSRDC-s12

Overview
Type: Continuous
Format: numeric
Width: 6
Decimals: 0
Range: 0-420000
Valid cases: 39976
Invalid: 312813
Minimum: 0
Maximum: 420000
Mean: 21.8
Standard deviation: 2103.4

Post code for firm's street address (trinstpcde)
File: rscdataforSRDC-s12

Overview
Type: Continuous
Format: numeric
Width: 4
Decimals: 0
Range: 0-9999
Valid cases: 352788
Invalid: 1
Minimum: 0
Maximum: 9999
Mean: 7438.3
Standard deviation: 1360.5

Magisterial district where firm is located (magdistric)
File: rscdataforSRDC-s12

Overview
Type: Discrete
Format: character
Width: 2
Valid cases: 352769
Invalid: 0

Type of business (typebusine)
File: rscdataforSRDC-s12
Type of business (typebusine)
File: rscdataforSRDC-s12

Overview
Type: Discrete
Format: numeric
Width: 1
Decimals: 0
Range: 1-7
Valid cases: 352785
Invalid: 4

Type of organisation (estypeorg)
File: rscdataforSRDC-s12

Overview
Type: Discrete
Format: character
Width: 3
Valid cases: 206565
Invalid: 0

Does the firm export goods, from survey (exportgoods)
File: rscdataforSRDC-s12

Overview
Type: Discrete
Format: character
Width: 1
Valid cases: 206565
Invalid: 0

Does the firm export services, from survey (exportservices)
File: rscdataforSRDC-s12

Overview
Type: Discrete
Format: character
Width: 1
Valid cases: 206565
Invalid: 0

Does company usually sub-contract or outsource work (essubcontr)
File: rscdataforSRDC-s12

Overview
Type: Discrete
Format: character
Width: 1
Valid cases: 206565
Invalid: 0

cadastral number from separate GIS survey (cadastralno)
File: rscdataforSRDC-s12

Overview
Type: Discrete
Format: character
Width: 21
Valid cases: 165312
Invalid: 0
Date of liquidation (liquiddatestr)
File: rscdataforSRDC-s12

Overview
Type: Discrete
Format: character
Width: 8
Valid cases: 352789
Invalid: 0

Date of suspension (suspendatestr)
File: rscdataforSRDC-s12

Overview
Type: Discrete
Format: character
Width: 8
Valid cases: 352789
Invalid: 0

exit indicator based derived from wage bill+turnover (exit)
File: rscdataforSRDC-s12

Overview
Type: Discrete
Format: numeric
Width: 1
Decimals: 0
Range: 1-1
Valid cases: 8117
Invalid: 344672

Firm age, from survey (age)
File: rscdataforSRDC-s12

Overview
Type: Continuous
Format: numeric
Width: 3
Decimals: 0
Range: -2-206
Valid cases: 126814
Invalid: 225975
Minimum: -2
Maximum: 206
Mean: 15.2
Standard deviation: 16.8

an indicator of entry derived from wage bill+turnover (entry)
File: rscdataforSRDC-s12

Overview
Type: Discrete
Format: numeric
Width: 1
Decimals: 0
Range: 1-1
Valid cases: 11782
Invalid: 341007