



NIDS and DataFirst

Martin Wittenberg

DataFirst

Presentation to Public Sector Economists' Forum

26 November 2014

Overview

- National Income Dynamics Study
 - What is it?
 - What data does it contain? What can it be used for?
 - How do you access the data?
- DataFirst
 - What is DataFirst?
 - Roles of DataFirst
 - Data Portal
 - Secure Data
 - Training
 - Research



NIDS and DataFirst

NATIONAL INCOME DYNAMICS STUDY

What is NIDS?

- NIDS is a nationally representative **panel** data set
 - Track people over time
 - Interview all people in the core panel plus co-resident household members
- Study is commissioned by the Presidency and run by SALDRU (South African Labour and Development Research Unit) at UCT
- First wave in 2008
 - 7 300 households and 28 000 individuals
- Reinterviewed in 2010
 - 9 000 households and 34 000 individuals
- Third wave in 2012
 - 10 000 households and 37 000 individuals
- Fourth wave is in progress

What information is in NIDS?

- Household information
 - Composition
 - Services, Assets
 - Consumption module
- Individual information (adult)
 - Education, Labour market, Incomes
 - Health (self-assessed), Anthropometric measurements, Subjective well-being
- Individual information (child)
 - Education, Health, Parents & family support, Anthropometric measurements

What can the data be used for?

- Data can be used cross-sectionally
 - Get nationally representative pictures for 2008, 2010, 2012
 - Also provincially representative. Not below that.
 - Will miss immigrants, but it does capture births
 - Selective attrition will become a problem over time
 - Issue mainly among Whites/Indians
- Main purpose is to look at **changes** over time
 - Do poor people stay poor?
 - How do people gain/lose employment?

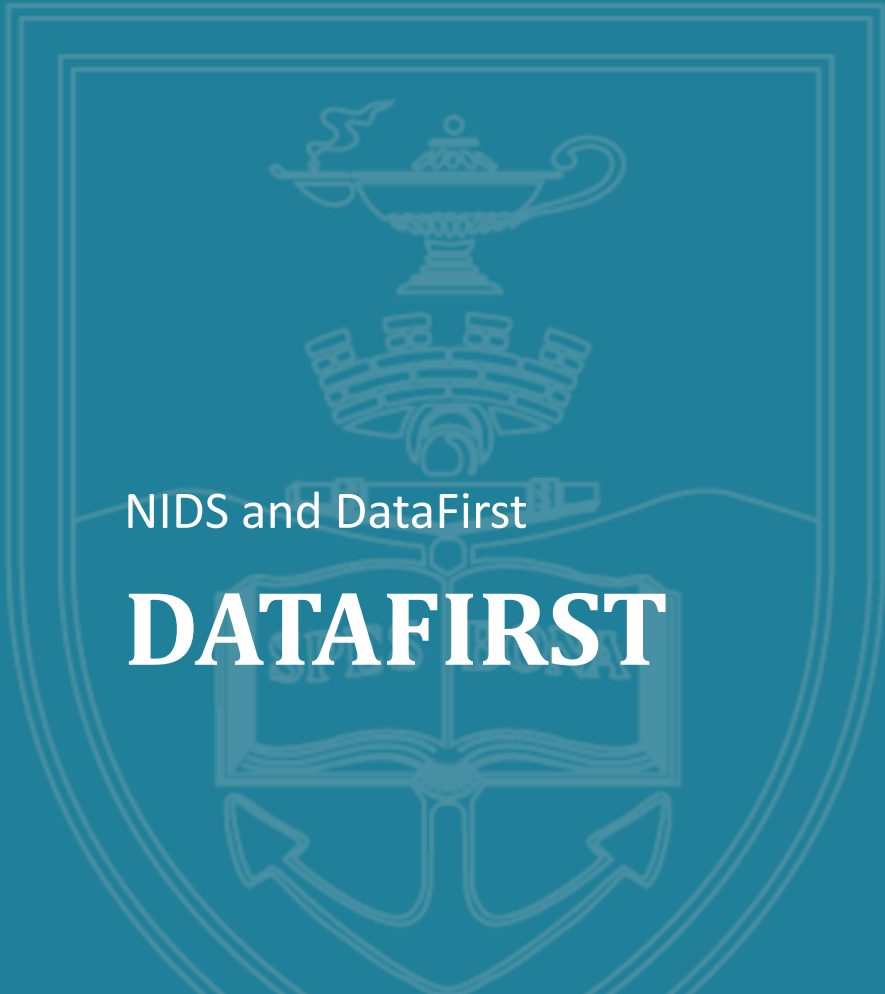
How do you access the data?

- Data is freely available from the DataFirst data portal
 - Requires a user registration process only
- But
 - Proper use will require the right sort of statistical software
- Training in data analysis techniques
 - is available from SALDRU and DataFirst

 **DataFirst**
University of Cape Town

NIDS and DataFirst

DATAFIRST



University of Cape Town
iYunivesithi YC

What is DataFirst?

- DataFirst is a data service based at UCT
 - Started by disseminating surveys run under auspices of SALDRU and the Centre for Social Science Research
 - Project for Statistics on Living Standards and Development (1993)
 - First publicly released nationally representative household survey
 - Cape Area Panel Study (2000 – 2009)
 - Made other surveys (mainly SA) available to researchers

Role of DataFirst

- Data dissemination
 - Data portal
 - Secure data centre
 - User feedback & support
- Training
 - Data curation
 - Survey analysis methods
- Research
 - Data quality issues
 - Substantive
- National/International networks

Role of DataFirst

DATA DISSEMINATION

Data Portal

- Key elements
 - Searchable database of surveys
 - Detailed metadata
 - Questionnaires
 - Citations module

<http://www.datafirst.uct.ac.za>

DataFirst

- DataFirst Home
- Data Dissemination
- Data Portal
- Citations
- Contact us



Home > Data Portal

Search by Keyword

in study description

in variable description

Search

Reset

Filter by Year

Show studies conducted between

1960 and 2014

Filter by Data Access

All

- Public use data files
- Licensed data files
- Data enclave
- Data available from external repository

Filter by Country

All

View / Select More

Filter by Topic

All

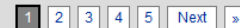
Data Portal

- About
- Datasets
- Citations

Found 287 studies out of 287

Sort results by: Country | Year | Title | Popularity

Showing 1-15 of 287 studies



Afrobarometer Survey 2008, Merged 20 Country

Africa, 2008

By: Institute for Democracy in South Africa (IDASA), Michigan State University (MSU), Ghana Centre for Democratic Development (CDD-Ghana)

Created on: Jan 16, 2012 Last modified: Nov 14, 2013 Views: 11672 Downloads: 428 Citations: 119



SACMEQ III Project 2005-2010

Africa, 2007

By: Southern and Eastern Africa Consortium for Monitoring Educational Quality

Created on: Nov 20, 2012 Last modified: Nov 14, 2013 Views: 14937 Downloads: 1296



Afrobarometer Survey 2005-2006, Merged 18 Country

Africa, 2005-2006

By: Institute for Democracy in South Africa (IDASA), Michigan State University (MSU), Ghana Centre for Democratic Development (CDD-Ghana)

User support

- On-line help
- Queries range from data access to data quality issues
- If there is a data quality query we try to reproduce it/resolve it before referring back to the original data producer

What data do we carry?

- **Microdata**
 - Mainly household surveys
 - Some firm data (tables from Manpower surveys)
- **Next frontier**
 - Admin data
 - Agreement with DBE to carry school surveys
- **Confidentiality issues**

Secure Data Services

- DataFirst established the first secure data centre in Africa
 - Room that is physically secured
 - Fingerprint reader
 - Surveillance camera
 - Only dumb terminals
 - People can work on the data, but results are saved on the server
 - DataFirst staff check the output to see that it is not disclosive and then e-mail output back to researcher (audit trail)
 - Emphasis on “secure researchers”
 - Protocols have to be signed off by institution
 - Training session

What secure data do we have?

- NIDS geospatially referenced data

Data that we have permission to release (but that still needs to be cleaned/processed)

- UCT admissions data
- DBE schools data
- RSC levy data from City of Cape Town



Role of DataFirst

TRAINING

DataFirst Training

- Survey Data Analysis Skills (with SALDRU)
 - Developing a Postgraduate Diploma
 - Short courses
- Data Curation Skills

Short courses 2015

- January: Cross-sectional econometrics
- February: Introduction to Panel Data
- June: Analysis of Complex Survey Data
- June/July: Measuring Economic Welfare
- November: Labour Statistics
- Nov/December: Consumption Data

Data curation training

- Run under the auspices of the “Accelerated Data Program” of the OECD
- Designed to make National Statistics Offices in Africa comfortable with the NADA software
- Teach metadata creation skills
- In last 12 months have run workshops in Botswana/Lesotho/Rwanda

Role of DataFirst

DATAFIRST RESEARCH

Types of Research

- Data quality issues
 - e.g. how to work with breaks in the October Household Survey/Labour Force Survey series
 - What to do about missing data
 - Released PALMS (post-Apartheid Labour Market Series)
- Best practice for working with survey data and noisy data
 - Weights, calibration, calculation of standard errors
- Scoping studies: what data is out there?
- Substantive issues
 - Wage trends



DataFirst

University of Cape Town



Role of DataFirst

NETWORKS

Types of Relationships

- International Data Services
 - UK Data Archive
 - International Consortium for Political and Social Research
 - World Data Service
- International data initiatives
 - World Bank, OECD – accelerated data program
 - International Household Survey Network
- National agencies
 - Statistics SA
 - DST, National Planning Commission, DBE
- Academics
 - Internationally: Population Studies Centre, University of Michigan
 - Nationally: Applied microeconomists, 4 SARChI chairs,