Profiling green jobs in South Africa: An occupational tasks approach

Data statement

We use three datasets for this analysis

1. Labour force surveys from PALMS and the South Africa Quarterly Labour Force Survey (QLFS)
The labour market data we use in this analysis is the Post Apartheid Labour Market Series 1993-2019 (PALMS) version 3.3. Our period of focus is 2012-2019 and PALMS contains Quarterly Labour Force Surveys for this period i.e. QLFS 2012 Q1- QLFS 2019 Q2. Since we needed all the four quarters of 2019, we downloaded QLFS 2019 Q3 and Q4 from the same website and merged these to the PALMS dataset. We provide a dofile merging the files and cleaning the variables.

This data is publicly available on the DataFirst website(https://www.datafirst.uct.ac.za/) however one must register to get login information. See the links below.

*Download links
PALMS V3.3 LINK
https://www.datafirst.uct.ac.za/dataportal/index.php/catalog/434/get-microdata
QLFS Q3 LINK
https://www.datafirst.uct.ac.za/dataportal/index.php/catalog/794/get-microdata
QLFS Q4 LINK
https://www.datafirst.uct.ac.za/dataportal/index.php/catalog/795/get-microdata

2. Tasks data from Granata, J. and Posadas, J. (2024).
Granata, J. and Posadas, J. (2024). Why Look at Tasks When Designing Skills Policy for the Green Transition? A Methodological Note on How to Identify Green Occupations and the Skills They Require (English). Policy Research working paper; no. WPS 10753; PEOPLE Washington, D.C.: World Bank Group.
http://documents.worldbank.org/curated/en/099507304172419760/IDU16ad6160713dd714cb81ba4f109325371b13.

This data cannot be republished for public use. No portion of the paper can be reproduced without this data.

3. ISCO88-ISCO08 crosswalk from the ILO
Correspondence between ISCO88 and ISCO08 classification tasks. Data is available here: https://www.ilo.org/resource/correspondence-table-isco08-isco88.
This dataset is public and can be redistributed as per ILO's Rights and Permissions site (https://www.ilo.org/rights-and-permissions)

Executing the analysis.
STEP 1: Load the *mastergreenjobs.do file* on to Stata.
Before executing the file, change the master directory to that of your computer. The next step is to install some user written commands that can't be installed using *ssc install*. These are *grc1leg* and *plotplainblind*. Instructions are in the *mastergreenjobs*.do file.

STEP 2: Creating the green task intensity (GTI) variables for South Africa
The *greentasksSA.do* file creates the Green Tasks Intensity (GTI) scores.

The datasets used for these are ***Dictionary of green terms 11 11 21 SA.xls*** and ***ISCOGroups_en.xls*** found in the data folder under GTI. The resulting data is ***greentasksSA.dta.***
This data is not publicly available but can be requested from Granata and Posadas (2024)
Table 1 in the draft paper is an excerpt of ***Dictionary of green terms 11 11 21 SA.xls***.

STEP 3: Crosswalk the GTI scores from ISCO08-to ISCO88
* The ***greenOCCS_isco88.do*** crosswalks the GTI scores from ISCO-08 to ISCO-88 consistent with the Statistics South Africa Quarterly Labour Surveys.

*Please not that the ISCO (International standard classification of occupations) is an ILO (International Labour Organisation) classification for organising jobs into clearly defined set of groups according to the tasks and duties undertaken in the job. The first version was known as ISCO-58 and was adapted in 1958, subsequent versions are ISCO-68, then ISCO-88 and finally ISCO-08 and was adapted in 2007.

*The GTI scores are computed using ISCO08 occupational classifications (that is the classification Granata and Posadas 2024 used to define green tasks). However, the South African labour market data is collected using ISCO-88 occupational classification codes. Therefore, after calculating the GTI scores, we had to crosswalk the scores from ISCO-08 to ISCO-88 to merge the scores to South African labour market surveys.

* The ILO provides corresponding tables to move from ISCO-08 to ISCO-88. We provide this corresponding table in excel format (***crosswalk_isco08_isco88.xlsx***). The crosswalk file is also provided in Stata format (***crosswalk_isco08_isco88.dta)*** in the folder ISO08crosswalk in the Data folder. The crosswalk was originally sourced from the link
https://www.ilo.org/public/english/bureau/stat/isco/isco08/ , which is not available anymore. A pdf version of the table can be accessed at *https://www.ilo.org/resource/correspondence-table-isco08-isco88*

The resulting data file is ***greenOCCS_isco88.dta.***

STEP 4: Cleaning PALMS 2012-2019
The file "***mergeqlfs_palms.do***" merges palms v3.3 and QLFS 2019 Q3 and Q4.
The raw data is found in the palms folder within the data folder. The raw data files are named as follows:
  i)   palmsv3.3-stata11.dta
  ii)   qlfs-2019-q3-worker-1.0-stata11.dta
  iii)   qlfs-2019-q4-worker-1.0-stata11.dta
The data is publicly available on the DataFirst website website(https://www.datafirst.uct.ac.za/) however, one must register to get login information. See the links below.

*Download links
PALMS V3.3 LINK
https://www.datafirst.uct.ac.za/dataportal/index.php/catalog/434/get-microdata
QLFS Q3 LINK
https://www.datafirst.uct.ac.za/dataportal/index.php/catalog/794/get-microdata
QLFS Q4 LINK
https://www.datafirst.uct.ac.za/dataportal/index.php/catalog/795/get-microdata

* The resulting data is saved as ***PALMS_q3q4qlfs2019.dta***

STEP 5: QLFS variable creation and cleaning
The file "palms4waves.do" cleans the merged labour force survey data and creates some labour market variables for the analysis. The resulting data file is ***palms4waves.dta***

STEP 6: Merging PALMS to the Green Tasks Intensity (GTI) scores.
The file "*PALMS4_GTI_CLEAN.do*" is used merge the labour force surveys data file ***palms4waves.dta*** to GTI scores ***greenOCCS_isco88.dta.*** The resulting data is ***PALMS4_GTI_CLEAN.dta***

STEP 7: Creating Appendix Tables containing all green occupations.
The file "occupationcounts_1.do" creates tables A1 and A2 in the Appendix. The data used in this analysis is ***PALMS4_GTI_CLEAN.dta*** in the palms folder within the data folder. The results are saved in the **output** folder.
Table 3 in the draft paper is an excerpt of Table A1 in the Appendix.

STEP 8: Creating figures (Figure 1-7) and regression tables (Table 4-6) in the paper.
The file "***Analysis.do***" performs the descriptive and regression analysis. The exact lines that produce the figures and tables are labelled in the dofile itself. The output (Figures and Tables) is stored in the output folder.

Data sets approximate access dates.
 1. Dictionary of green terms 11 11 21 SA.xls and ISCOGroups_en.xls (September 2022 from the authors)
 2. palmsv3.3-stata11.dta   accessed December 2022
 3. qlfs-2019-q3-worker-1.0-stata11.dta   accessed December 2022
 4. qlfs-2019-q4-worker-1.0-stata11.dta   accessed December 2022
 5. crosswalk_isco08_isco88   accessed April 2020