UNIVERSITÉ LAVAL

Usage and importance of DASP in Stata

Abdelkrim Araar, Jean-Yves Duclos and Luis Huesca

Comparisons of Stata to other software or use of Stata together with other software.

Mexico, May 12, 2011

Outline

***DASP*: a Stata package for distributive analysis**

Conclusion

Outline

DASP: a Stata package for distributive analysis

DASP in a few words

DASP features

Other *DASP* features

DASP's main menu

DASP's main variables

Using variables in *DASP*

Inputting *DASP* commands

Applications and files in *DASP*

Producing curves with *DASP*

Saving *DASP* graphs

Examples with the Mexican data

Conclusion

DASP: a Stata package for distributive analysis

DASP in a few words

- Stata enables programmers to provide specialized ".ado" routines to add to the power of the software.
- *DASP*, which stands for *Distributive Analysis Stata Package*, is mainly designed to assist those researchers and policy analysts that are interested in conducting distributive analysis with Stata.
- *DASP* uses Stata for two main reasons:
  - Stata is a powerful tool to store and manage household data surveys. Combining *DASP* and Stata allows to use the same environment for processing and analyzing data.
  - Stata easily allows adding specialized programs, making it possible for programmers to add to its power and flexibility.

DASP features

DASP allows to:

- Estimate the most popular statistics (indices, curves) used for the analysis of poverty, inequality, social welfare, and equity;
- Estimate the differences in such statistics;
- Estimate standard errors and confidence intervals by taking full account of survey design;
- Perform the most popular distributive decomposition procedures;
- Check for the ethical robustness of distributive comparisons;
- Support distributive analysis on more than one data base at the same time.

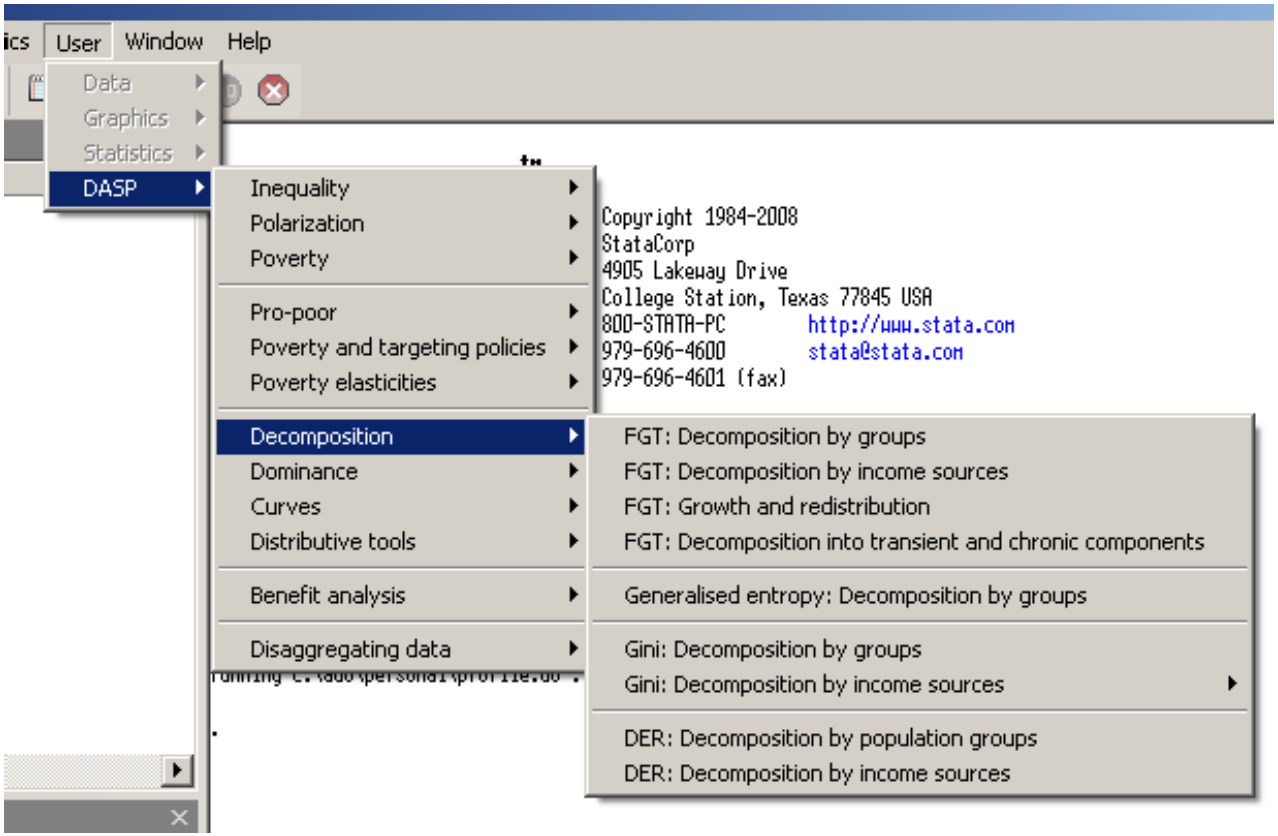
Other *DASP* features

- Contains optimized algorithms for the estimation of distributive indices;
- Unifies syntax and parameter use across various estimation procedures for distributive analysis;
- For each *DASP* module, three types of files are provided[1]:
  - ***.ado** file: contains the program of the module;
  - ***.hlp** file: contains the help material for the given module;
  - ***.dlg** file: allows the user to perform the estimation using the module's dialog box.

[1] For more information about DASP modules, see the user manual: (**?**)).

DASP's main menu

DASP's windows menu makes it possible to access quickly each of the dialog boxes. The latter are grouped by main themes.



DASP's main variables

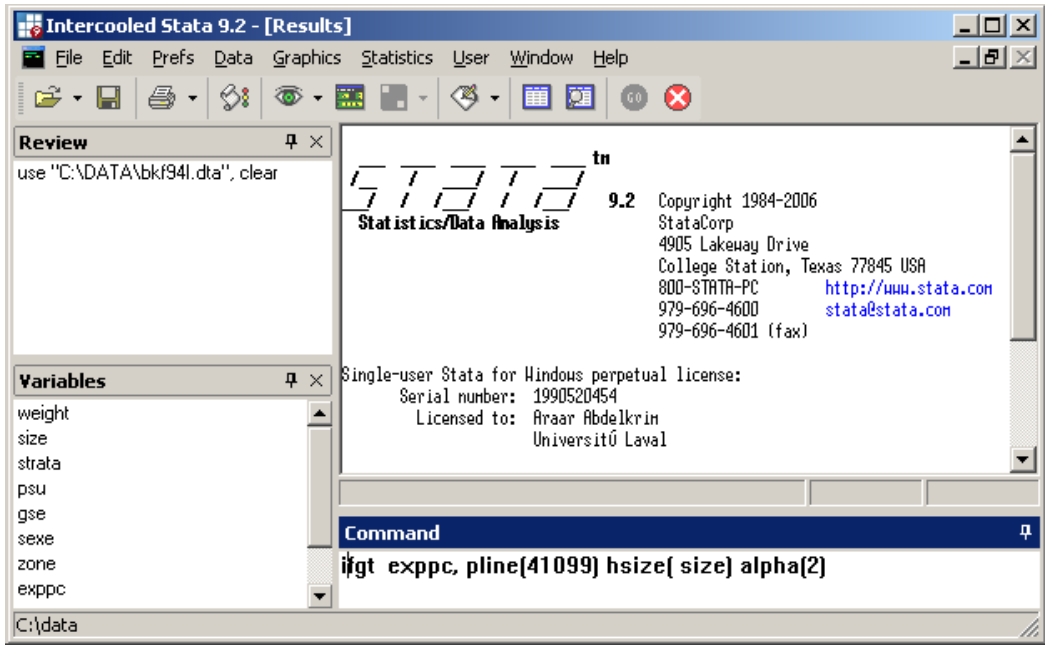
- VARIABLE OF INTEREST. This is the variable that usually captures living standards. It can represent, for instance, income per capita or expenditures per adult equivalent.
- SIZE VARIABLE. This refers to the "ethical" or physical size of the observation. This variable usually refers to the number of household members.
- GROUP VARIABLE. Say that we wish to estimate poverty within a country's rural area or within female-headed families. One way to do this is to force *DASP* to focus on a population subgroup defined as those for whom some GROUP VARIABLE(say, area of residence) equals a given GROUP NUMBER (say 2, for rural area).
- SAMPLING WEIGHT. Sampling weights are the inverse of the sampling probability. This variable should be set upon the initialization of the data set.

Using variables in *DASP*

- *DASP* makes it possible to use simultaneously more than one data file.
- The user should initialize each data file before using it with *DASP*. This initialization is done by:
  - Labeling variables and values for categorical variables;
  - Initializing the sampling design with the command *svyset*;
  - Saving the initialized data file.
- It is useful to add a character such as “I” to the names of initialized files (Example: Uganda99I.dta) in order to distinguish them.

Inputting *DASP* commands

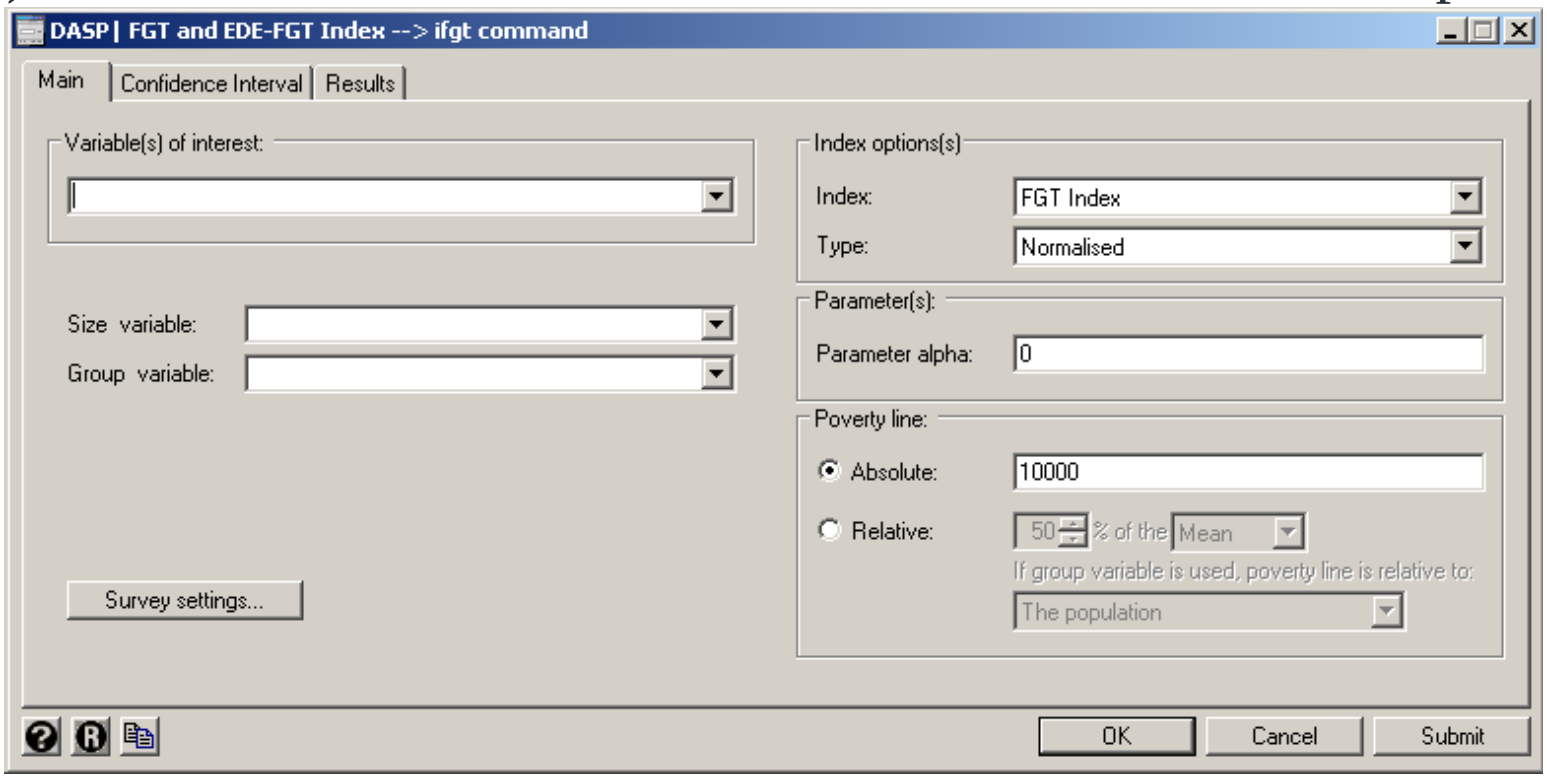
- Stata and *DASP* commands can be entered directly into a command window:



- An alternative is to use dialog boxes. For this, the command *db* should be typed and followed by the name of the relevant *DASP* module. Example: db ifgt.

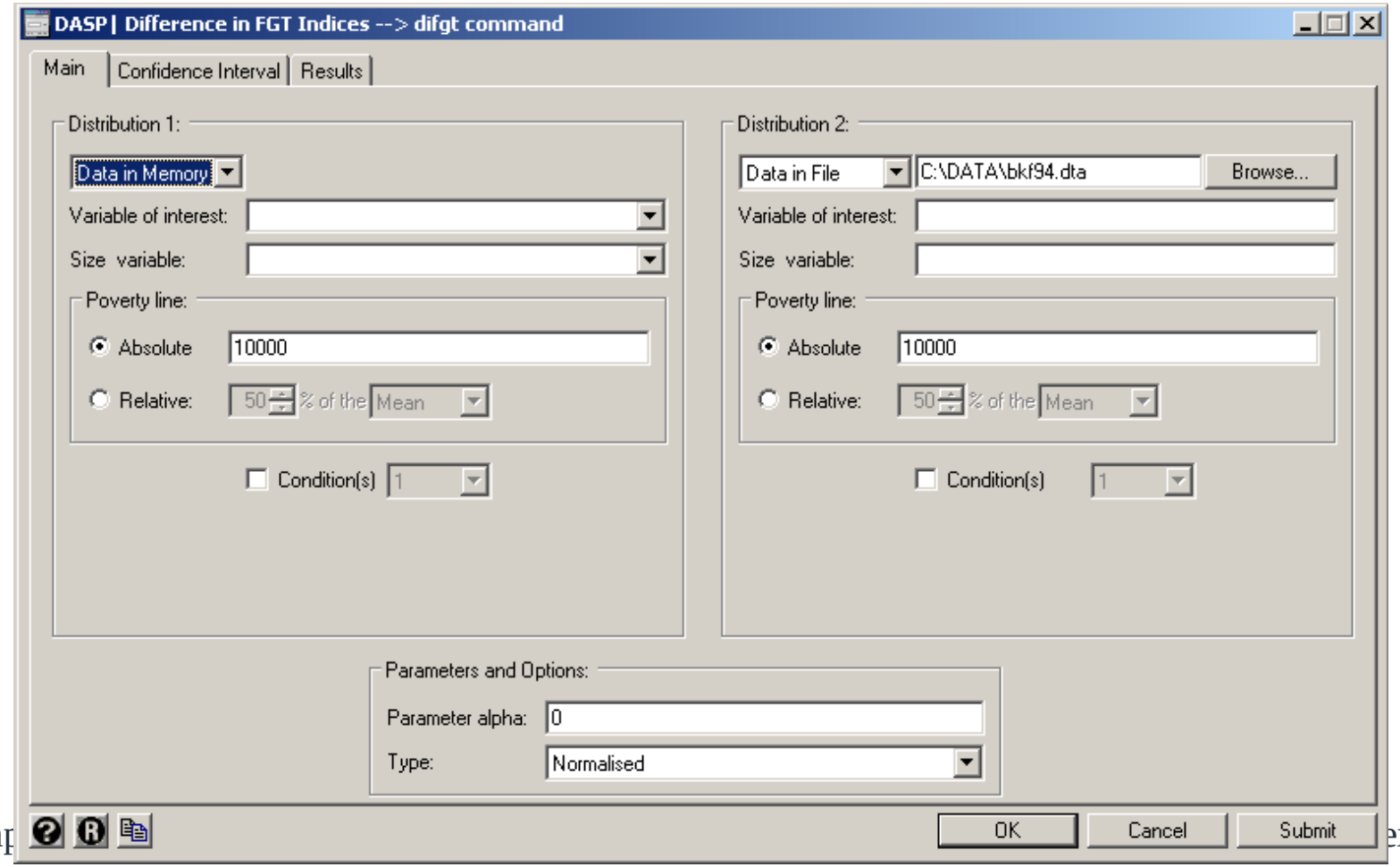
Applications and files in *DASP*

Two main types of applications are provided in *DASP*. For the first one, the estimation procedure uses only one data file, the data file in “memory” (or “loaded”). It is from that file that the relevant variables must be specified.



Applications and files in *DASP*

Two main types of applications are provided in *DASP*.
For the second type of applications, two distributions are needed. For each of these two distributions, the user can specify the currently-loaded data file (the one in memory) or one saved on disk.

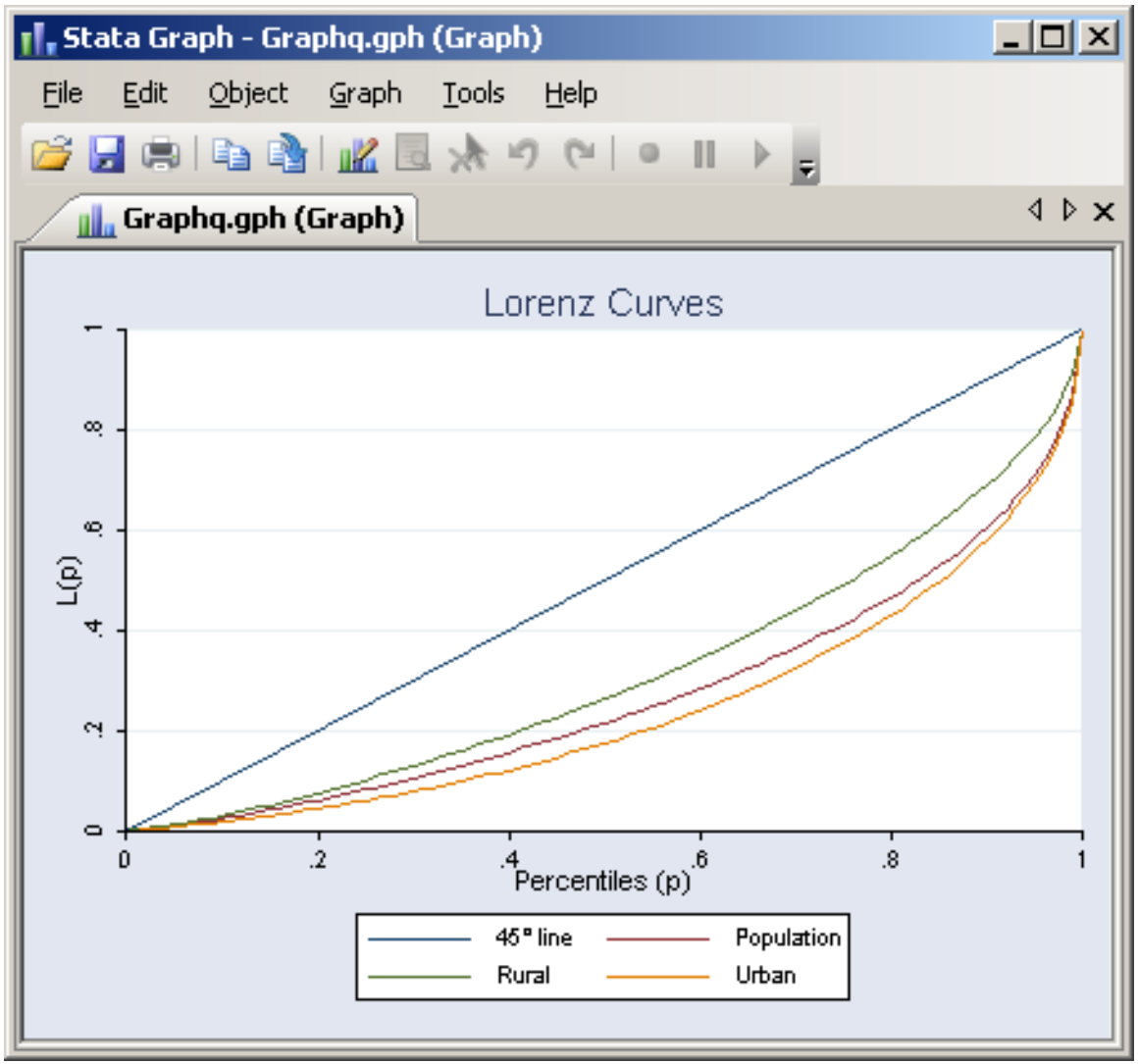


Producing curves with *DASP*

- *DASP* was designed to facilitate the use of curves to display distributive information.
- For instance, if we wish to graph Lorenz curves to compare inequality between rural and urban areas, the following command line can be typed:
  `clorenz` `exppc, hgroup(zone) hsize(size)`
  where in this example `exppc` is *per capita* expenditures , `size` is household size and `zone` is the zone variable (1 = rural / 2= urban).

Producing curves with *DASP*

- After executing this command the following window appears:

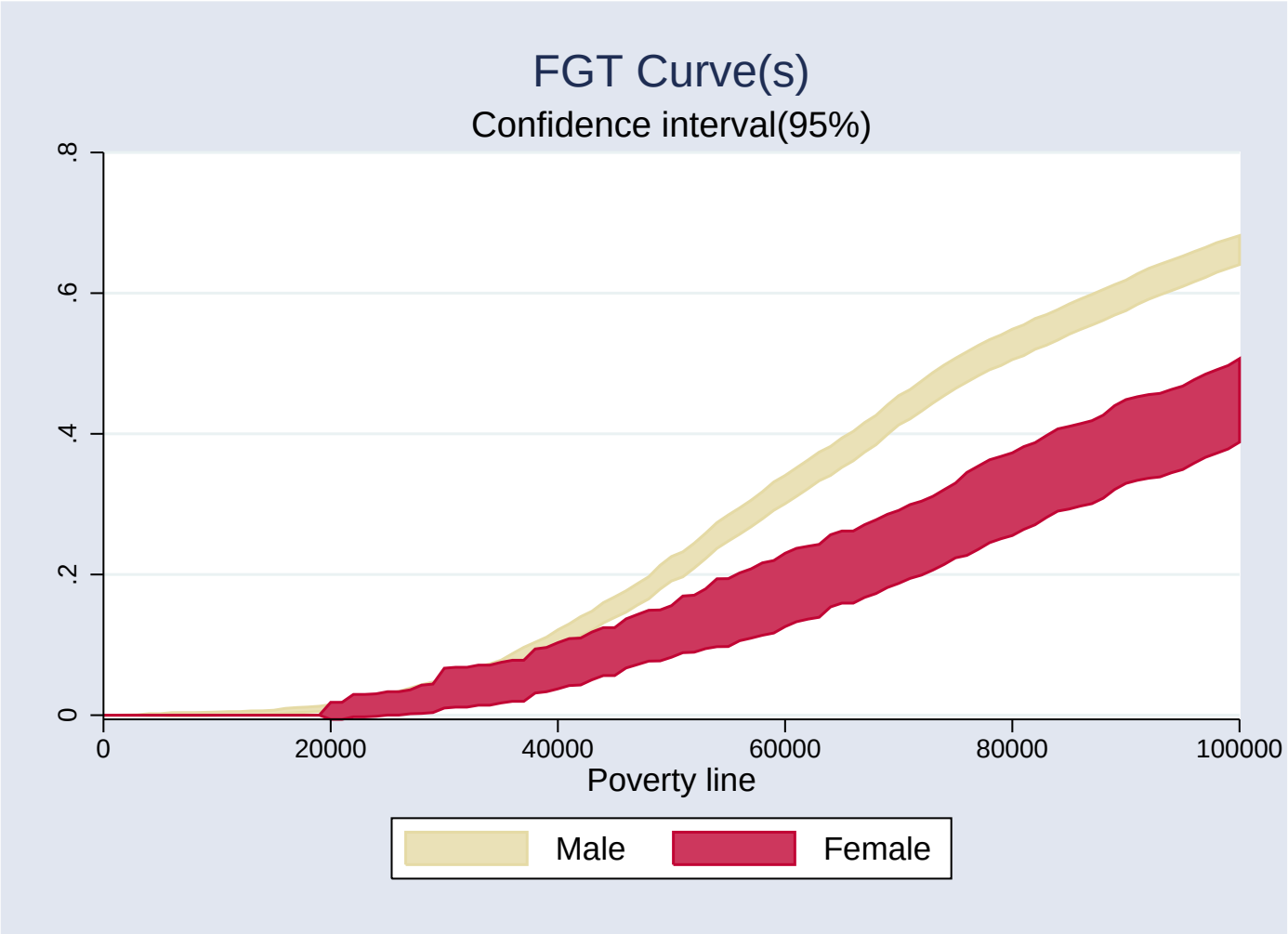


Producing curves with *DASP*

- For many curves, *DASP* allows showing their confidence intervals according to selected levels of statistical significance (this value is by default set to 95%).
- For instance, to draw confidence intervals around FGT curves, we can use the **`cfgtsm`** *DASP* module: **`cfgtsm`** `exppc, alpha(0) hsize(size) hgroup(sex) max(100000)`

Producing curves with *DASP*

- After executing this command the following window appears: Drawing the confidence interval of distributive curves (FGT curves)



Saving *DASP* graphs

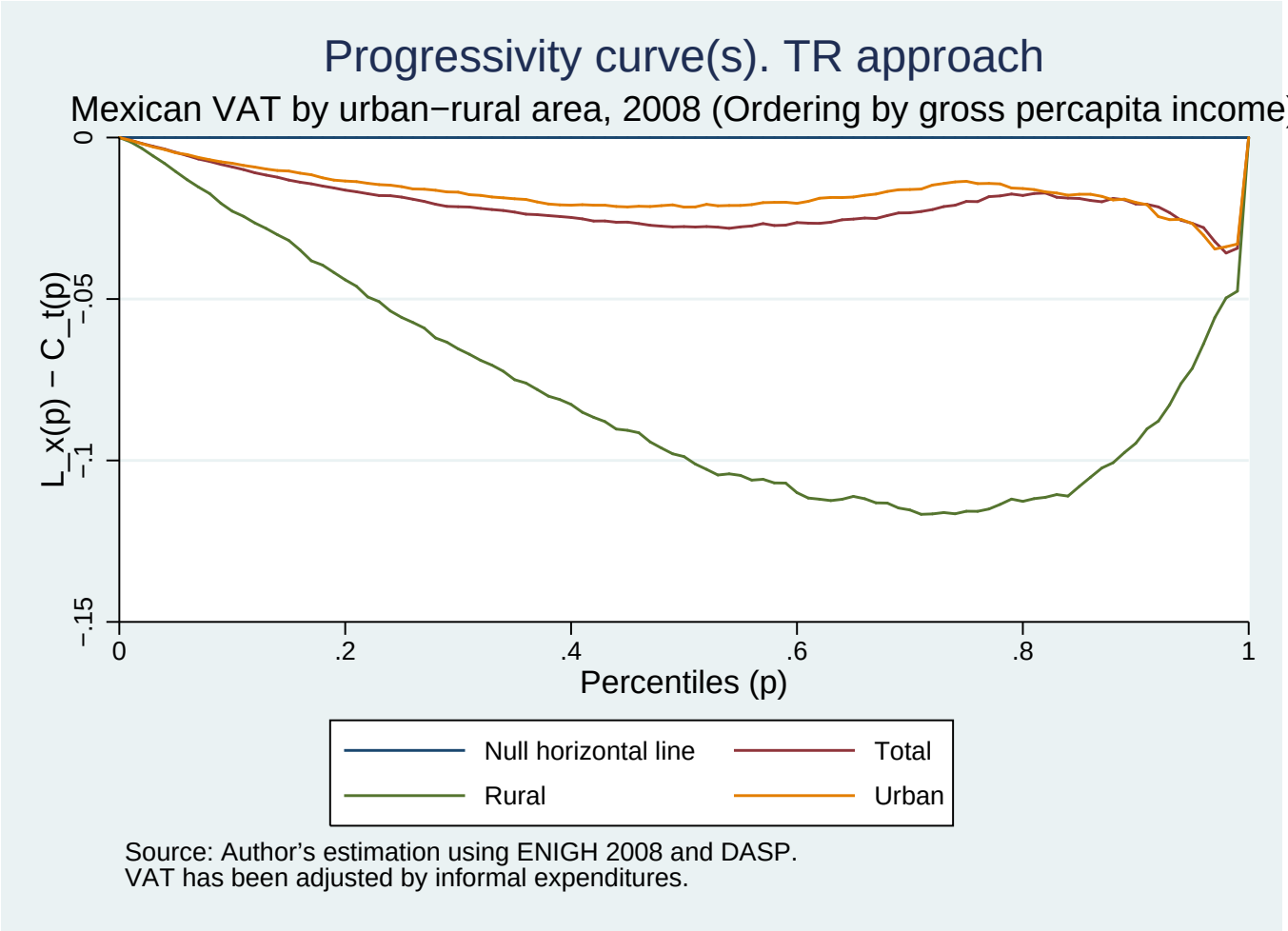
- Graphs produced with *DASP* or Stata can be saved in many different formats. Among them:

*.gph is Stata's graphical format. It is useful to allow re-editing the graph (with Stata 10 or higher).

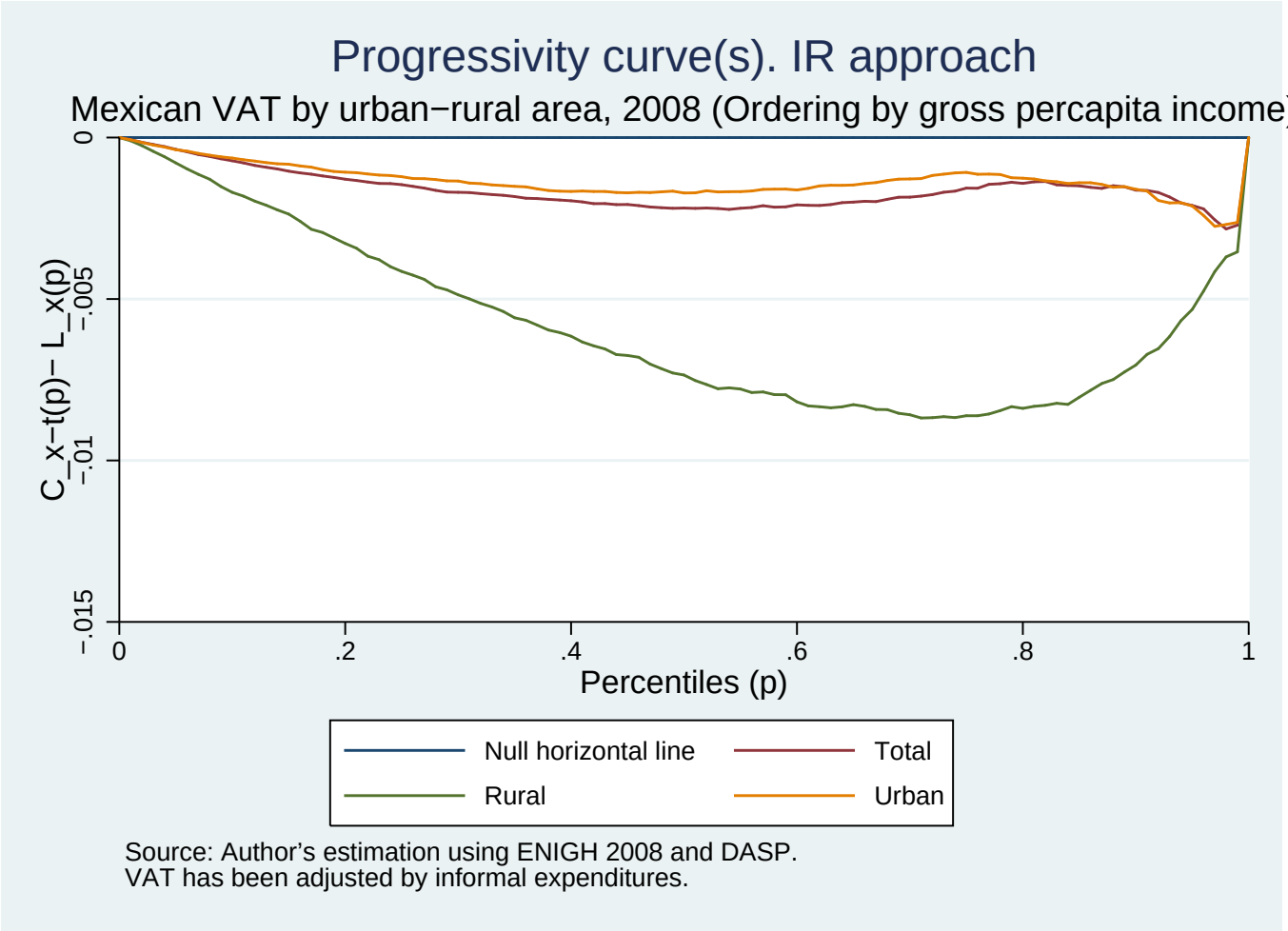
*.wmf is the Windows metafile format. This format may be easily inserted into Word documents. The user can also copy a Stata graph and paste it directly into a Word document.

*.eps is the encapsulated postscript format. This format can easily be inserted in Latex documents.

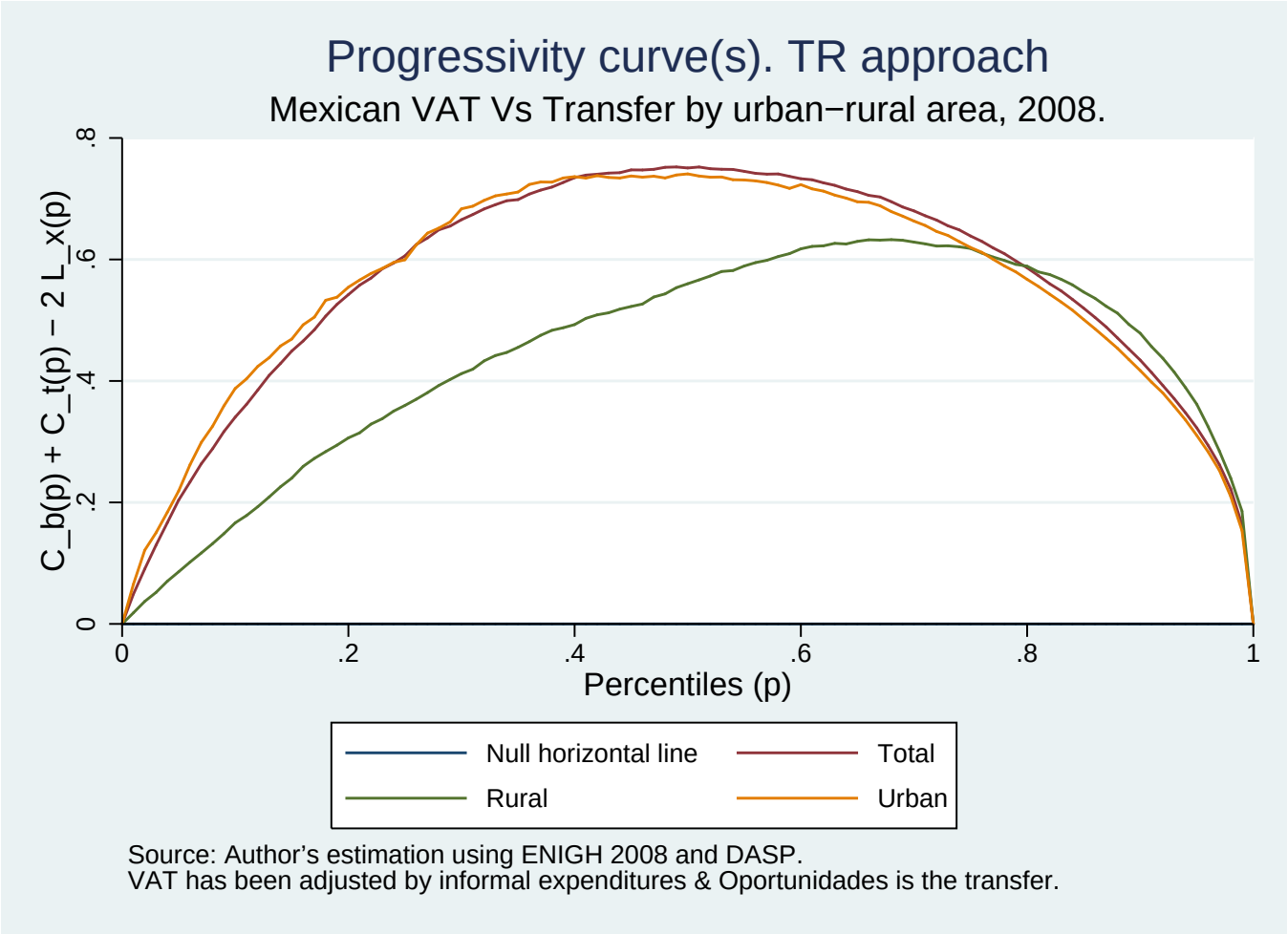
Examples with the Mexican data



Examples with the Mexican data

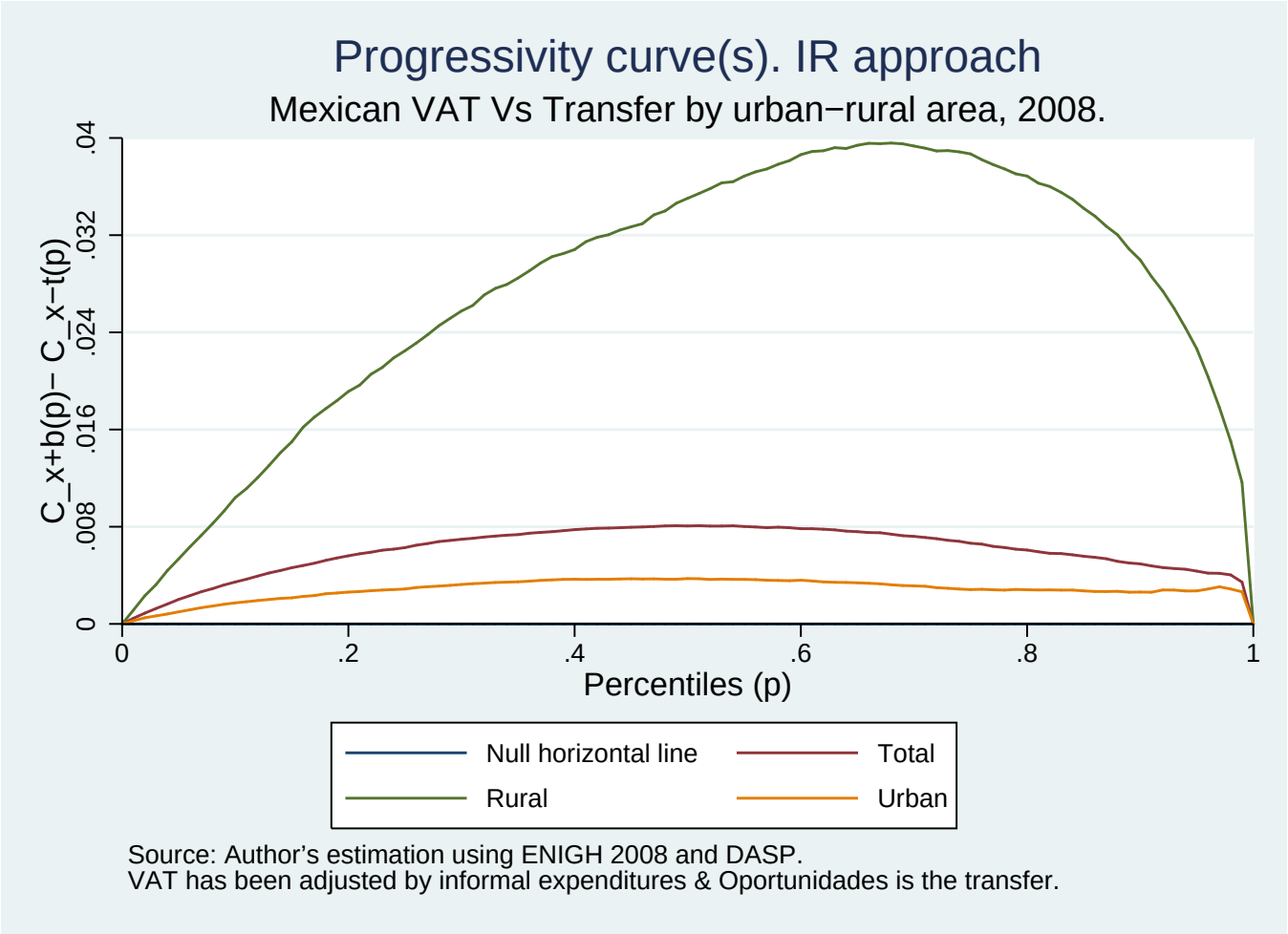


Examples with the Mexican data

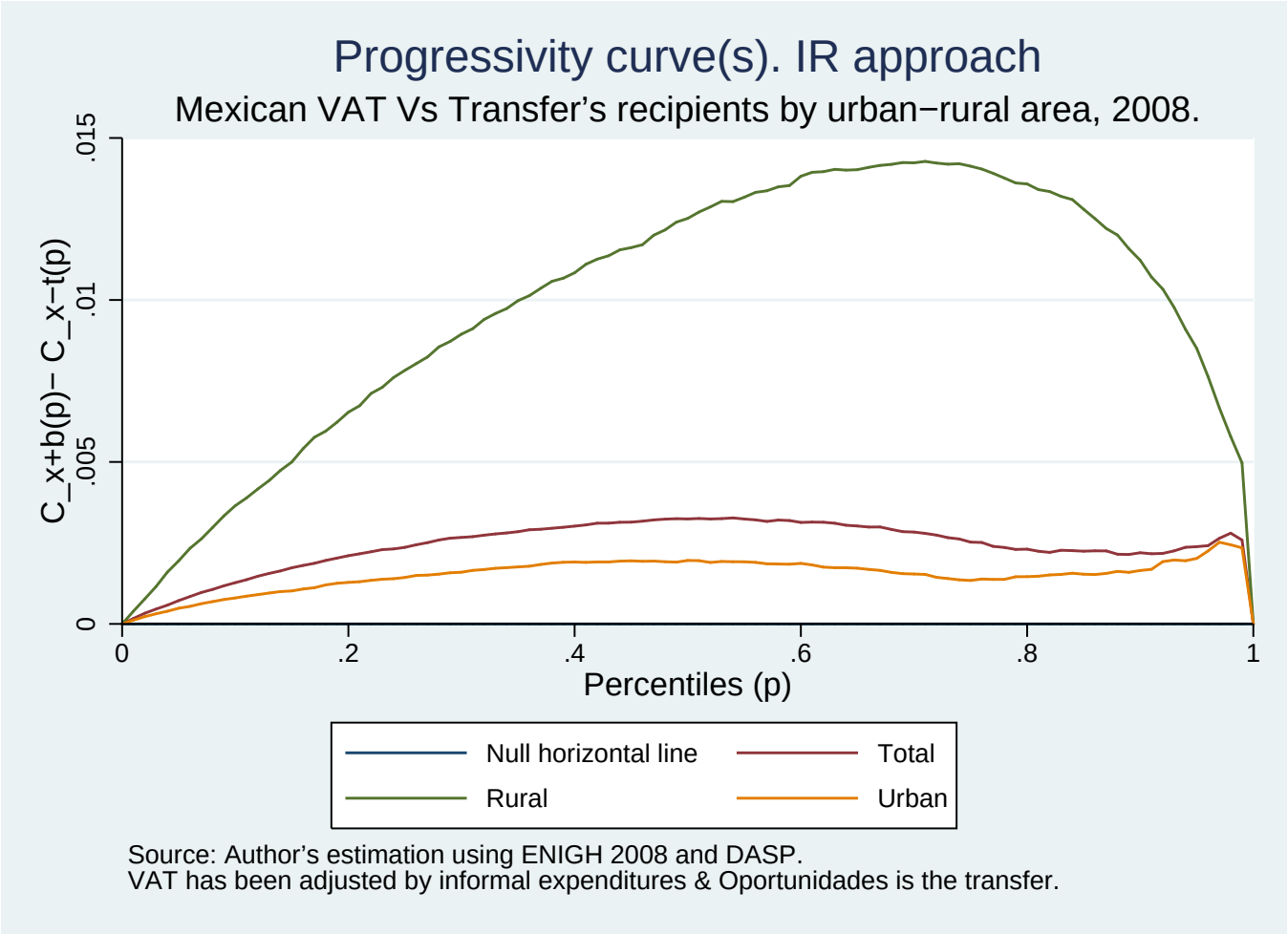


Source: Author's estimation using ENIGH 2008 and DASP.
VAT has been adjusted by informal expenditures & Oportunidades is the transfer.

Examples with the Mexican data



Examples with the Mexican data



Outline
DASP: a Stata package for distributive analysis
Conclusion
Summary
Relevant DASP commands
Exercises with Stata and DASP
Additional literature on DASP and DAD
References

Conclusion

Summary

- Stata is a popular software that provides powerful statistical applications and that is simple to use.
- Stata commands can be inputted through dialog boxes, do files, or commands windows.
- *DASP* facilitates the estimation of the most popular statistics used for the analysis of poverty, inequality, social welfare, and equity, and provides various sophisticated statistical tools to check for the robustness and the precision of such statistics.
- *DASP* unifies syntax and parameter use across various estimation procedures for distributive analysis.
- *DASP* allows the use of two distributions at the same time, and simplifies the production of tables and graphs.

Relevant DASP commands

- FGT and EDE-FGT poverty indices (ifgt).
- FGT CURVE with confidence interval (cfgts).
- Lorenz and concentration curves (clorenz).

Exercises with Stata and DASP

- Exercises 1.1, 1.2, 1.3

Additional literature on DASP and DAD

- The material from this presentation draws largely from Araar and Duclos (2009b) and Duclos and Araar (2006). See also Araar and Duclos (2009a), et Christian Toft (2006) and Zhang (2003) for the DAD software.

References

ARAAR, A. AND J.-Y. DUCLOS (2009a): "DAD: a Software for Poverty and Distributive Analysis," *Journal of Economic and Social Measurement*, 43, 175–189.

——— (2009b): "*DASP*: Distributive Analysis Stata Package," PEP, CIRPÉE and World Bank, Université Laval.

DUCLOS, J.-Y. AND A. ARAAR (2006): *Poverty and Equity Measurement, Policy, and Estimation with DAD*, Berlin and Ottawa: Springer and IDRC.

ET CHRISTIAN TOFT, M. B. (2006): "Poverty and Equity. Measurement and Analysis Jean-Yves Duclos, Abdelkrim Araar and Carl Fortin," *Économie publique/ Public economics*, 18-19, 5–12.

ZHANG, Q. (2003): "DAD, an innovative tool for income distribution analysis," *Journal of Economic Inequality*, 1, 281–284.