USING GRID TECHNOLOGIES FOR THE MAPPING OUT OF TAXATION POLICY

Glavelis Themistoklis, Margaritis Konstantinos, Ploskas Nikolaos, Samaras Nikolaos

Department of Applied Informatics, University of Macedonia
Email: {mai086, kmarg, mai084, samaras}@uom.gr

January 2008

APPLICATION PROBLEM – AREA DESCRIPTION

This presentation aims to introduce an application which is a powerful tool for the tracing of taxation policy.

More generally, the scope of this tool is to present the ongoing opportunities that grid technologies provide to many sectors, such as the sector of Public Administration.

Why grid technologies can be used in the Public Administration?

Public Administration needs to have:

- Reliable elements
- Alternative models
- Alternative scripts

Grid Technologies provides:

- process of big volume of data
- simultaneous control of different approaches, models and scripts

effective application of tax policy

How grid technologies can be used in the Public Administration?

Economic tools:

- Vast list of historical elements
- Different models

Using grid technologies, we can:

- Seek for the most effective model
- Advance in control of various affairs, altering
  - Values of entries
  - Values of parameters of models
- Make forecasts
- Handle alternative scripts

The model - Variables

The regression which we appreciated the tax policy of government associates the Taxes (T) with:

1) the Government budget deficit/surplus (S_{govt})
2) the Transfer Payments (TR)
3) the Net Interest Payments (INT) and
4) the Government Purchases (G)

The model

The regression that was estimated is the following :

$$S_{govt} = a_0 + a_1 T + a_2 TR + a_3 INT + a_4 G$$

where a_i, i=1, …, 4 are the coefficients of the regression and a_0 is the constant term.

The model - Interpretation

With this model a government can forecast its future surplus or deficit.

So, if governments' goal is to increase its surplus by 5%, can create different approaches to achieve this goal, such as:

- Increase only the taxes
- Increase the taxes and decrease the Government Purchases

The model - Data

- Problem:

  Lack of real elements for many years.

- Solution:

  1. Sample of elements for the past eighteen years from the databases of OECD, Eurostat and National Statistical Service of Greece.
  2. Creation of a vast list of instances with random generated numbers.

The model – Creation of instances

The application creates:

- A huge volume of instances
- Daily data from 1988 to 2006 (6570 elements each instance)
- In each execution are created at least 20 instances in order to have more reliable forecasts.
- Elements are as much as possible more representatives using the range of values of the real elements that were collected from databases

Application – Tools being used

The application was developed with the high-level open source language Gnu Octave (edition 2.9.12).

Why Octave:

- Open source
- Available in the Grid Infrastructure that we used
- Statistical and econometric packages that provides as with various functions to estimate our regression
- Many mathematical functions which help as to accomplish linear algebra computations that we used in the application

Application - Resources Demands (1)

- Our model has 5 variables
- For 18 years we have 6570(18x365) daily data
- As a result we have an array with 5 columns (one for each variable) and 6570 rows
- Finally we have 32850(5x6570) elements
- All these only for input

Application - Resources Demands (2)

- Our application make use of the Ordinary Least-Square method (OLS) to estimate the regression.
- Various statistics are extracted from the models' estimation.
- As a result, it uses many complicated linear algebra computations.
- Problem:

  Exhausting demands for memory and computational resources.

Application - Resources Demands (3)

- The size of the produced data set was about 20 gigabytes
- Consequently, it is infeasible to be executed locally in a typical computer with a single processing node.
- Obviously, the solution is given by the grid technologies only

Why a Grid (technical explanation) ?

- Due to the demand for resources
- If there wasn't the solution of Grid technologies, this kind of applications would be extremely complicated and difficult to accomplish a project with this size of data.

Application – Execution in grid

- The application was executed through the User Interface of Athens and was managed by the European Mediterranean Virtual Organization (http://www.eumedgrid.org).
- The time duration that the application needs in order to be executed varies depending on the number and the size of instances that will be created.
- Obviously, as more as possible instances will be created by the application, so much more representative will be the estimation of our model.

Application – Report (1)

- At the beginning the coefficients of our variables are estimated.
- Next, we check if the coefficients are in agreement with the economic theory.
- Finally, the application examines if the coefficients are statistically significant.

Application – Report (2)

General Statistics

- R-square statistic
- Adjusted R-square statistic
- F-statistic
- Correlation at the first level (order)
- Total cputime (secs)

Application - Example of Report (1)

REPORT GENERATION

- Our model is

  $S_{govt} = a_0 + a_1T + a_2TR + a_3INT + a_4G + e_i$

- CONSTANT

  The Constant is : -102,759,720.71582535

- TAXES (T)

  The coefficient of Taxes(T) is in agreement with the economic theory

  The coefficient of Taxes(T) is significant and it is:

  0.01625334

Application - Example of Report (2)

- TRANSFERS (TR)

  The coefficient of Transfer Payments(TR) is in agreement with the economic theory

  The coefficient of Transfer Payments(TR) is not significant and it is:

  -0.00847457

- NET INTEREST PAYMENTS (INT)

  The coefficient of Net Interest Payments(INT) is in agreement with the economic theory

  The coefficient of Net Interest Payments(INT) is significant and it is:

  -0.04973043

- GOVERNMENT PURCHASES (G)

  The coefficient of Government Purchases(G) is in agreement with the economic theory

  The coefficient of Government Purchases(G) is significant and it is:

  -0.00894575

Application - Example of Report (3)

GENERAL STATISTICS

- R-square statistic : 0.8192591
- Adjusted R-square statistic : 0.8356110
  Our sample interprets the population at 83.56 %
- The coefficients are not all significant at the same time (f-statistic).
- There is correlation at the first level (order).
- Total cputime to run the regression : 3505.30000000

Social Impact

- Government can have alternative scripts for the taxation policy.
- Useful information can be collected for the taxation policy.
- With these tools in the possession of a government, an effective mapping out of taxation policy is more possible.

Scientific Impact

- Reliability

  A bigger sample of data leads to more accurate results.

- Forecasts

  Forecasts can be more adequate.

Importance of results

- A powerful tool for forecasts.
- Grid technologies are solving real-world problems.
- Effectiveness due to the range of the samples.
- More information for the tracing of taxation policy.

Future Work (1)

- With real elements for many years we would be able to make more reliable, realistic and effective forecasts.
- Public Administration could be able to manage more alternative scripts for the taxation policy.

Future Work (2)

- Government can be able to see how it can change its taxation policy according to the current circumstances, for example when there is an economical crisis due to the increase of oil.

- Our model is just an illustrative example to present the ongoing opportunities that grid technologies provide to many sectors, such as the sector of Public Administration. So, each government can adjust the models and scripts according to its needs and goals.

END OF PRESENTATION

Thank you for your attention…

Questions?