Data infrastructure for business statistics in Portugal: Stats Business

8 September 2022

Cristina Neves
Economic Statistics Department
• Multidimensional perspective
• Data sources
• Data integration: Stats Business
Multidimensional perspective
✓ Multidimensional perspective of enterprises’ economic performance
✓ Micro data linking
✓ Integration of data from different data sources

Sample surveys
Administrative data

Combination of several data sources
✓ Focus on harmonisation (concepts, definitions)

✓ Creation of a new Administrative Data Unit (Data Collection and Management Department – Business Data Unit)

✓ Making data integrable:
  
  Quality control
  
  Outlier detection
  
  Estimations (non-response, also on administrative records)
Data sources
Data sources

IES/SBI

Conventional surveys
(quantitative)

New surveys
(qualitative; economic sentiment; forecasts)

Other administrative data
Data from enterprises:

• IES – SIMPLIFIED BUSINESS INFORMATION

• OTHER ADMINISTRATIVE DATA (growing use)

• CONVENTIONAL SURVEYS (Intrastat; Short term statistics - turnover)

• NEW SURVEYS (capture mainly qualitative aspects that influence business activity (Survey on Framework Regulation Costs; Management Practices Survey; Survey on Information and Communication Technologies Usage in Enterprises; Community Innovation Survey); get indications on economic sentiment and forecasts (Survey on Perspectives of Exports of Goods))
Four legal requirements from different entities on enterprises, fulfilled in a single report:
Main advantages for Statistics Portugal:

Units covered: 50,000 (≈15%) 400,000 (≈100%)

Availability: 12 Months 6.5 Months

Number of variables: 800 5,000
Other administrative data:

Microdata or highly detailed data from other sources:

- **Social Security**: Single Report; Monthly Statement of Remunerations
- **Tax Authority**: “e-invoice” system; VAT; fiscal declarations (Income taxes, corporation taxes, property taxes...) – business statistics, housing prices and rents (information in high detail for society)
- **Ministry of Justice**: creation and dissolution of enterprises; NACE classification (SICAE).
Data integration – Stats Business
Reengineering Business Statistics

SBS/ Integration of business statistics’ domains

Bank of Portugal
(Financial enterprises)
ISP (Insurances)

IES (Companies)

Fiscal information
(Sole proprietors and Independent workers)

IFATS

Financial Business Statistics
Non financial Business Statistics

Integrate

Business Demography
Globalization
Entrepreneurship

Main input for National Accounts
Integration of statistical data from surveys and administrative data:

**Surveys:**
- Integrated Business Accounts System
- Survey on Framework Regulation Costs; Management Practices Survey; Survey on Perspectives of Exports of Goods
- Intrastat, Short Term Statistics, CIS, IVNE, CIS, Survey on ICT Usage in Enterprises, ...

**Other data sources:**
- Single Report – Social Security
- Monthly Statement of Remunerations
- E-invoice system, VAT, fiscal declarations (income and corporation taxes, property taxes...)
- Creation and dissolution of enterprises; NACE classification (SICAE)

Enable Statistics Portugal:
- Data security
- Statistical confidentiality

Powerful enterprise database

Stats Business
Stats Business

National Data Infrastructure

ICT

CIS

Other surveys

Single Report

Other sources

PRODCOM

ITGS
One of the most advanced enterprise databases in international terms;

Its analytical exploration is still under development from three perspectives:

✓ exploitation for a more efficient production of existing official statistics (e-invoice: useful during pandemic, for Intrastat and Tourism);

✓ design of new official statistics which are made available at an early stage as experimental statistics (*Stats Lab*) (innovation and size; management and performance; new: gender analysis);

✓ support for scientific research (contrafactual analysis).
Main economic indicators of innovative and non-innovative enterprises (2018)

### CIS/IBAS (IES)

<table>
<thead>
<tr>
<th>Category</th>
<th>Innovative Enterprises</th>
<th>Non-Innovative Enterprises</th>
</tr>
</thead>
<tbody>
<tr>
<td>Turnover</td>
<td>EUR 27,655 thousand</td>
<td>EUR 7,249 thousand</td>
</tr>
<tr>
<td>GVA</td>
<td>EUR 6,306 thousand</td>
<td>EUR 1,827 thousand</td>
</tr>
<tr>
<td>Persons employed</td>
<td>65 Women, 82 Men</td>
<td>29 Women, 39 Men</td>
</tr>
<tr>
<td>Apparent labour productivity</td>
<td>EUR 42,710</td>
<td>EUR 26,990</td>
</tr>
<tr>
<td>Personnel costs per person employed</td>
<td>EUR 24,104</td>
<td>EUR 18,788</td>
</tr>
</tbody>
</table>

**Source:** Statistics Portugal, CIS and IBAS
Distribution of innovative and non-innovative enterprises, by ticscore, persons employed and total (2018)

**Example 2**

**ticscore** – Information and communication technology use intensity indicator based on community survey on ICT usage and e-commerce in enterprises

\[ Y_i = AK_i^\alpha L_i^\beta e^{\gamma g_i} \prod_j e^{\delta_j x_{i,j}} \]

\[
\ln \left( \frac{Y_i}{L_i} \right) = \ln(A) + \alpha \ln \left( \frac{K_i}{L_i} \right) + (\beta + \alpha - 1) \ln(L_i) + \gamma g_i + \sum_j \delta_j x_{i,j}
\]

**Management Practices Survey:**

**G-score:** the G-score is obtained for each company through the simple average of the scores attributed to the answers to a set of specific questions in the Management Practices Survey.

The score for each answer varies between 0 and 1, with the maximum value assigned to the answer option corresponding to the most structured practice and the minimum value to the least structured.

<table>
<thead>
<tr>
<th>Variable</th>
<th>Coefficients</th>
<th>t Stat</th>
<th>P-value</th>
</tr>
</thead>
<tbody>
<tr>
<td>ln(A)</td>
<td>8.81</td>
<td>94.83</td>
<td>0.00</td>
</tr>
<tr>
<td>gscore</td>
<td>1.68</td>
<td>14.51</td>
<td><strong>0.00</strong></td>
</tr>
<tr>
<td>young</td>
<td>-0.02</td>
<td>-0.37</td>
<td>0.71</td>
</tr>
<tr>
<td>adult</td>
<td>0.06</td>
<td>1.82</td>
<td>0.07</td>
</tr>
<tr>
<td>group</td>
<td>0.20</td>
<td>6.86</td>
<td>0.00</td>
</tr>
<tr>
<td>micro</td>
<td>-0.52</td>
<td>-12.47</td>
<td>0.00</td>
</tr>
<tr>
<td>large</td>
<td>0.52</td>
<td>11.33</td>
<td>0.00</td>
</tr>
<tr>
<td>p_export</td>
<td>0.09</td>
<td>3.38</td>
<td>0.00</td>
</tr>
<tr>
<td>ln(asset/NPS)</td>
<td>0.13</td>
<td>21.23</td>
<td>0.00</td>
</tr>
<tr>
<td>% hab_sup</td>
<td>0.92</td>
<td>15.70</td>
<td>0.00</td>
</tr>
<tr>
<td>antig</td>
<td>0.02</td>
<td>6.79</td>
<td>0.00</td>
</tr>
<tr>
<td>ln(NPS)</td>
<td>-0.27</td>
<td>-18.52</td>
<td>0.00</td>
</tr>
</tbody>
</table>
Management Practices Survey - Relationship between gscore and tic_score

\[ x = \text{gscore} \]
\[ y = \text{tic\_score} \]

Source: Statistics Portugal, MPS and IUTICE
Thank you for your attention!

Cristina Neves
Economic Statistics Department
cristina.neves@ine.pt

UNCEBTS
8 September 2022