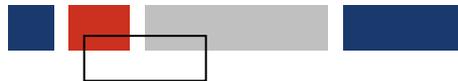


PORTUGAL experience

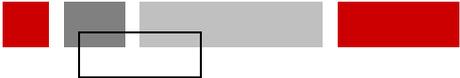
Results of the evaluation on the exploitation of scanner data



Filipa Marin Lidónio/Cecília Cardoso
Data Collection Department



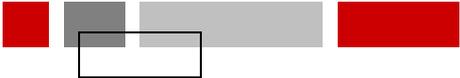
Scanner Data Workshop
26|27 September 2013
Lisbon, Portugal



Focus of presentation

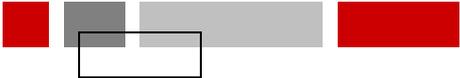
- Sharing Portuguese experience in the research on the exploitation of scanner data concerning:
 - collaboration with data providers, negotiation process, approaches adopted and data required (retrospective data for pilot project)
 - ways to link products to COICOP using EAN and/or internal store codes

- Background and context
- Contacts with data providers/collaboration process
- Exploitation of scanner data/link to COICOP
- Next developments/further work



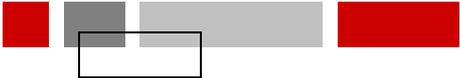
Background and context

- Statistics Portugal has been awarded in 2011 with a **Eurostat Grant** to undertake the initial research on the exploitation of scanner data
- In line with the national strategy on the **modernisation of data collection methods and processes** our focus was the study of pragmatic approaches to use scanner data as improvement of the existing price collection system
- A more ambitious goal is to extend the methods to other statistical operations



Background and context

- For grant application, Statistics Portugal has identified four main lines of action:
 - Knowledge acquisition
 - Collaboration with data providers
 - Pilot project
 - Infrastructure



Background and context

Purpose

- **Knowledge acquisition** to access data on product characteristics for products covered by scanner data, especially on the use of EAN and its linkage to COICOP and in-store codes; to learn from experiences of other countries using scanner data;

Results

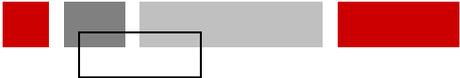
- ✓ Understand that a step-by-step approach is most suitable for Portugal, so that we can allocate resources in different phases of the project and to define the best way to obtain data, to select data providers and to prepare an effective collaborative process

Purpose

- **Collaboration with data providers** to explore and to negotiate arrangements to access scanner data from retailers chains and to select data providers for a pilot experience;

Results

- ✓ Very successful since the beginning with one of the retailers selected and also successful but more time consuming with the second retailer



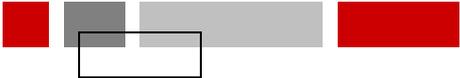
Background and context

Purpose

- **Pilot project** to establish continuous scanner data flows routines with selected retailers; to develop the necessary linking of aggregated or product-level to COICOP; to develop internal data methods for processing scanner data based on a datawarehouse approach; to implement actions to develop sample designs and weights, including methods to integrate scanner data with the existing price collection processes;

Results

- ± Some goals were achieved but our purpose was very (really very) ambitious.



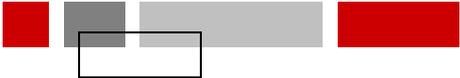
Background and context

Purpose

- **Infrastructure** to design and build an information system to support the pilot project;

Results

- ± Prototype developed to link to COICOP but not tested to work in a integrated way and with data flow routines



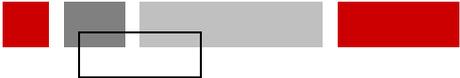
Contacts with data providers collaboration process

What we want?

- Successful collaboration
- Negotiate arrangements to access scanner data from retailer chains

How to do?

- Maximum simplification/minor effort for data providers
- Limiting the scope of data to Food and non alcoholic beverages
- Access retrospective data

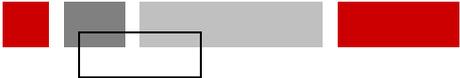


Contacts with data providers collaboration process

Who choose?

- 2 of the top 5 retailers chains that dominate the sector
 - 60% of national share
 - significant geographical distribution all over country
 - comparing to CPI collection, represent 40% of prices and outlets collected (COICOP 01)

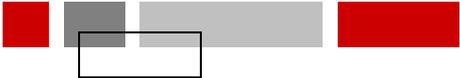
- The groups are: SONAE and Jerónimo Martins



Contacts with data providers collaboration process

Important remarks:

- The first action taken was **identifying the best key contact** in each of the two retailers in order to present the aims of the project and explain the conditions to implement a collaborative relation
- **Key challenge** → to find strong arguments to convince retailers to join the project



Contacts with data providers collaboration process

Important remarks:

The more convincing message in the first meeting was the following:

- **Scanner data is the future** and will be adopted in our country;
- We offer the opportunity to the provider **to participate in the very beginning of the project**, influencing the design of the project in order to be prepared in advance.



Contacts with data providers collaboration process

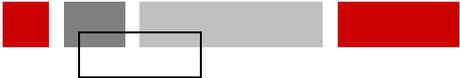
Important remarks:

- The **information is very sensitive** so it is important to explain how we intend to use it
- It should be mentioned that **we are flexible** and that data structure is not a problem and every data is important for analyses (even if redundant)
- **Choose to show possible examples of data structure;**
 - List of the outlets, items description, prices, transactions
- **This pragmatic approach** was appreciated by the providers

Contacts with data providers collaboration process

Important remarks:

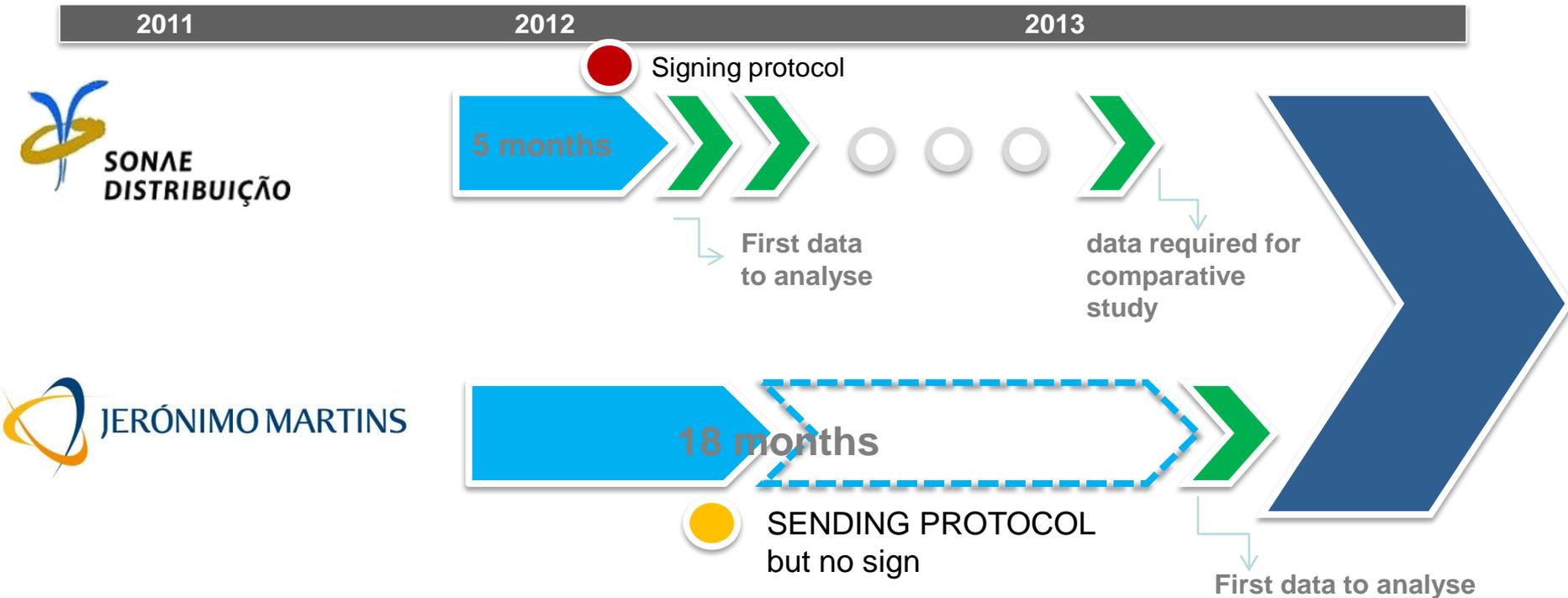
- **Be available** for additional meetings and clarifications
- Explore potential data to give in return
- If possible, adopt a formal protocol to sustain the collaboration process
- Adapt yourselves to retailers work practices instead of impose new ones
- ... and be patient!



Contacts with data providers collaboration process

- Although the results of our collaboration process were extremely positive (success in obtaining retrospective data for pilot project), the negotiation process was different with retailers
- **Collaboration with SONAE** was very easy, fast, cooperative since the beginning but worried with support conditions, like security (hand delivery files)
- **Collaboration with Jerónimo Martins** was more time consuming, less cooperative in the beginning, several months with no developments but very uncomplicated with security issues (sending data by e-mail)

Contacts with data providers collaboration process

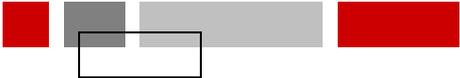


 NEGOTIATION PROCESS (several meetings)

 OBTAINING DATA

 Critical phase: no developments/reminders

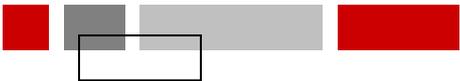
 Retailers available to continue participating



Contacts with data providers collaboration process

Retailers reaction

- Very positive and interested to participate
- Big companies have a social responsibility policy that covers these kind of goals, like building partnerships, costs rationalisation, efficiency, modernisation
- The possibility to provide data without additional effort or previous treatment looks also important because our position was to adopt to retailers work practices instead of impose new ones



Contacts with data providers collaboration process

- **In summary**, to build sustainable partnership with data providers, specially the target groups chosen, we must be patient, adapt to their work schedule and routines and define new approaches that best suit both interests.

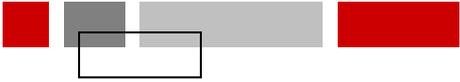
Exploitation of scanner data link to COICOP

Pilot experience goal:

- Analyse data structure and understand internal classification and items description

Note: for pilot project we only had data from one retailer

- Evaluate ways to link products to COICOP classification: using EAN and internal store codes;
- Explore items characteristics and identify attributes than can support link process and also future harmonisation with other data files;
- Identify a method to select products for comparative study with CPI



Exploitation of scanner data link to COICOP

Approach followed with first data received from SONAE group (not yet for comparative study):

- Prices and quantities sold in the two most important outlets: one in Lisbon region (Lisbon city, Colombo Shopping Centre) and one in North Region (Oporto City, Matosinhos), on items related with food and non alcoholic beverages for two consecutive months
- About 14.000 products

Exploitation of scanner data link to COICOP

The data provided is based on a hierarchical structure related with the internal code:

ITEMS	
Field	Example
Outlet Name	1 - CNT Matosinhos
Day	20121231
Business Unit	08 - Lacticíni/Congelad
Class	0802 - Gorduras
Subclass	080201 - Manteigas
Internal Code (SKU)	██████████
Item Denomination	MANT.S/SAL PRIMOR 250GR
EAN	5601051420028
Day price	██████
Day quantities sold	███

Exploitation of scanner data link to COICOP

- Linking SD to COICOP at the lowest level, elementary product - 8 digit COICOP (adopted in Household budget survey) or 6 digit COICOP
- it was possible to link about 80% of products and at the same time analyze item description in order to identify attributes like: brand, package size of product, price reduction classification and unit quantity. When necessary, we also consult retailers website information

Note: using internal code instead of EAN was chosen because this code was always filled and unique (allows integration within all stores of the same retailer)

Exploitation of scanner data link to COICOP

- Very time consuming but important to understand data and define specification for linking process
- The prototype tool to support link process was been developed both to search and select the codes in COICOP, type items characteristics and, very important, is improved with the actions undertaken.
- The typology of attributes is improved as new products are analyzed and classified.
- The search actions (key words search) are also saved and can allow understanding the process and ensuring harmonization

Exploitation of scanner data link to COICOP

Interface application

Lista COICOP - SKU

Enviar pesquisa para xls Sair

Sub categoria

Sku

Unid

Marca

Embalagem

Promoção

OBS

Quantidade:

Pesquisa: massa

Mostrar Todos os Registos

FOTO

Nível	Designação	Código COICOP	Descrição do Produto (COICOP 2010)	Unidade
5º	Variedade	01.1.1.3.1	Massas Alimentícias por cozinhar, não frigorificadas nem congeladas	
6º	Produto	01.1.1.3.1.01	Massas miúdas	(Kg)
6º	Produto	01.1.1.3.1.02	Massas graúdas - Outras n.d.	(Kg)
6º	Produto	01.1.1.3.1.02	Massas graúdas - Aletria	(Kg)
6º	Produto	01.1.1.3.1.04	Massas alimentícias - por cozinhar - Não Frigorificadas nem Congeladas - Outras n.d.	(Kg)
6º	Produto	01.1.1.3.1.04	Lasanha (Massa de Forno) - por cozinhar - Não Frigorificada nem Congelada	(Kg)

search

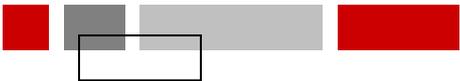
Attributes types

Link
6 or
8
digit

Exploitation of scanner data link to COICOP

Approach followed with second data received from same retailer SONAE group (for comparative study):

- 60 outlets from north region (20 common with CPI collecting)
 - 24 months information (2011 and 2012) : month average prices and monthly quantities sold
 - 2 months (November and December 2012) daily prices and quantities sold
 - Almost 420.000 prices
- Complete link to COICOP process for most representative products sold in each outlet (for each month, select 5% varieties with most expenditure)
 - 729 products were linked with COICOP in a smaller outlet and 1790 in the biggest outlet



Exploitation of scanner data link to COICOP

Approach followed with data received from the other group Jerónimo Martins/test integration:

- 2 months (November and December 2012) for one of major outlets
 - similar hierarchical structure but much more sub-classes
 - about 8.200 products
-
- 45% products automatically linked to COICOP (same EAN of Sonae)

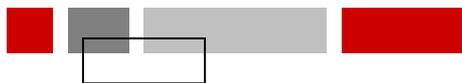
Exploitation of scanner data link to COICOP

Approach followed with data received from the other group Jerónimo Martins/test integration:

- Some products don't match (specially absence of products or outlet brand)
- Some products with the same EAN are not equal, most of them in quantities type attribute, specially packs: JM uses conversion

Next developments/further work

- Continue with retrospective analyses and integrate more data from the second retailer in order to define new strategies in linking process
- Explore other uses of data to support products replacement in traditional CPI collection (based in most representative) and for quality control procedures
- Maintain contact with data providers and establish conditions to obtain data in a monthly basis (not retrospective)
- Define specifications for integration system
- Develop contacts with other retailers



THANK YOU!

filipa.lidonio@ine.pt

cecilia.cardoso@ine.pt