

Modification of the Methodology to Estimate the Value of Procurement Published in OJ/TED



Economic and Statistical Working Group
Bruxelles, April 2014

Introduction

- The total value of procurement published annually in OJ/TED is a key element to assess the size and degree of openness of the European public procurement market.
- Unfortunately, the total final value of contracts is often missing in Contract Award Notices.
- It is also usually misreported.
- The real annual value of published contracts cannot be directly recovered from the information published in OJ/TED.
- It is therefore necessary to deal with:
 - Missing values
 - Outliers
- This presentation will show the alternatives under study in order to control for these two shortcomings of the data.

- 1 The nature of the problem
 - Outliers
 - Missing Values
- 2 The current methodology
- 3 The new proposed methodology (ongoing work)
- 4 Robustness checks
- 5 Concluding remarks

Outline

- 1 The nature of the problem
 - Outliers
 - Missing Values
- 2 The current methodology
- 3 The new proposed methodology (ongoing work)
- 4 Robustness checks
- 5 Concluding remarks

What do we mean by outliers? I

- Values that can be regarded as “abnormally” large.
- They will have a disproportionate impact on any descriptive statistics we compute.
- Outliers can appear due to many reasons, as for example:
 - Values that are obviously (deliberately?) mistyped:

What do we mean by outliers? II

- Xosé-Luís VARELA-IRIMIA
- User profile
- Update ECAS information
- Preferences
- Clipboard
- Recently viewed documents
- Saved search profile
- Customised RSS feeds
- Personal address book
- Bulk Download
- Log out

News

- Site news
- Public Procurement News
- RSS feeds
- What is RSS?

e-Certis

Information on certificates
required in public
procurement in the EU

EUR-Lex

The Official Journal of the
European Union online

EU Bookshop

All EU publications you are
looking for

28/01/2012 S19 Member states - Service contract - Contract award - Open procedure
I.1.V

F.Ivry-sur-Seine: special-purpose road passenger transport services

2012/S 19 030105

Contract award notice

Services

Directive 2004/18/EC

Section I: Contracting authority

I.1) Name, addresses and contact point(s)

Ville d'Ivry-sur-Seine
esplanade Georges Marrane
For the attention of: M. Philippe Bouyssou
94205 Ivry-sur-Seine Cedex
FRANCE
Telephone: +33 149602960
E-mail: cmp@ivry94.fr
Fax: +33 149602660
Internet address(es):
Address of the buyer profile: <https://ivry94.achatpublic.com/sdm/ent/gen/index.jsp>

Section II: Object of the contract

II.1) Description

II.1.5) Common procurement vocabulary (CPV)

60130000

Description

Special-purpose road passenger transport services.

Section V: Award of contract

Contract No: 11076 Lot No: 2 Lot 1: transport scolaire adapté Ivry-sur-Seine extra-muros (d'Ivry vers d'autres villes); Lot 2: transport scolaire adapté Ivry-sur-Seine intra-muros (d'Ivry à Ivry); Lot 3: transport adapté pour adultes. Le lot 1 est ultérieurement.

V.3) Name and address of economic operator in favour of whom the contract award decision has been taken

Groupe M Service
18 rue Saint-Vincent-de-Paul
75019 Paris
FRANCE

V.4) Information on value of contract

Initial estimated total value of the contract:

Value: 422 000 EUR

including VAT

Total final value of the contract:

Value: 59 999 999 999 999 999 999 EUR

including VAT

What do we mean by outliers? III

- Mistakes due to a lapse. For example, the CN displays:

II. szakasz: A szerződés tárgya

II.1) Meghatározás

II.1.1) Az ajánlatkérő által a szerződéshez rendelt elnevezés:

„Keretmegállapodás keretében mélyépítési feladatok ellátása Budapest Főváros XVII. kerület Rákosm

II.1.2) A szerződés típusa és a teljesítés helye

Építési beruházás

Tervezés és kivitelezés

A teljesítés helye: Budapest Főváros XVII. kerület, Rákosmente Önkormányzatának közigazgatási te
NUTS-kód HU101

II.1.3) Közbeszerzésre, keretmegállapodásra és dinamikus beszerzési rendszerre (DBR) vonatkozó

II.1.4) Keretmegállapodásra vonatkozó információk

Keretmegállapodás több ajánlattevővel

A tervezett keretmegállapodás résztvevőinek száma : 3

A keretmegállapodás időtartama

Időtartam év(ek)ben: 36

A közbeszerzéseknek a keretmegállapodás teljes időtartamára vonatkozó becsült összérték

Becsült érték éfa nélkül:

50 000 000 és 3 000 000 000 között HUF

What do we mean by outliers? IV

- While the CAN shows:



E-mail: riz.levente@rakosmente.hu

Fax: +36 12533489

Internet address(es):

General address of the contracting authority: www.rakosmente.hu

Section II: Object of the contract

II.1) Description

II.1.5) Common procurement vocabulary (CPV)

45230000, 34923000, 34928400, 34928500, 34992200, 45233000, 45231220, 45231300, 45232410, 45232452, 45233120, 45233161, 45233162, 77330000, 71320000

Description

Construction work for pipelines, communication and power lines, for highways, roads, airfields and railways; flatwork.

Road traffic-control equipment.

Urban furniture.

Street-lighting equipment.

Road signs.

Construction, foundation and surface works for highways, roads.

Construction work for gas pipelines.

Construction work for water and sewage pipelines.

Sewerage work.

Drainage works.

Road construction works.

Footpath construction work.

Cycle path construction work.

Floral-display services.

Engineering design services.

II.2) Total final value of contract(s)

II.2.1) Total final value of contract(s)

Value: 500 000 003 000 000 000 HUF

Excluding VAT

Section V: Award of contract

Contract No: 1 Lot No: 1

V.3) Name and address of economic operator in favour of whom the contract award decision has been taken

Duna Aszfalt Kft.

Béke u. 150.

6060 Tiszakécske

HUNGARY

V.4) Information on value of contract

Initial estimated total value of the contract:

Value: 500 000 003 000 000 000 HUF

Excluding VAT

Total final value of the contract:

Value: 500 000 003 000 000 000 HUF

Excluding VAT

What do we mean by outliers? V

- However, other may be “right” outliers:

Sezione V: Aggiudicazione dell'appalto

V.1) **Data della decisione di aggiudicazione dell'appalto:**
15.2.2012

V.2) **Informazioni sulle offerte**

V.3) **Nome e recapito dell'operatore economico in favore del quale è stata adottata la decisione di**
ATI Astaldi SpA – Impregilo SpA
via G.V. Bona n° 65
00156 Roma
ITALIA

V.4) **Informazione sul valore dell'appalto**
Valore totale inizialmente stimato dell'appalto:
Valore: 961 904 356,99 EUR
IVA esclusa
Valore finale totale dell'appalto:
Valore: 790 878 752,00 EUR

- Mistyped values that do not appear to be outliers represent an even bigger problem!

OJ/TED vs MaPPS

- Sometimes the information appears on the OJ/TED form, but perhaps not in the correct field.
- The OJ/TED format is not suitable for structured statistical analysis of data.
- MaPPS (DG Markt local database) does automated reading of OJ/TED to build a database, from xml files.
- Values not inserted in the correct field will be missing, even if strictly speaking their are present within the form.
- These missings will just add to the “real” ones (information which is missing in the form).
- Defining as missings the notices where the final value is not stated in the fields foreseen for that purpose:

The incidence of missing values - Classical (% of CANs)

country	2006	2007	2008	2009	2010	2011	2012	Total
AT	24.75	19.83	28.42	34.81	22.03	26.05	39.45	28.60
BE	23.82	22.42	23.96	33.75	41.15	38.19	33.37	32.89
BG	.	1.30	1.35	5.24	1.02	1.15	1.16	1.69
CY	6.68	6.34	3.35	7.32	5.49	0.10	0.45	4.00
CZ	1.69	1.15	1.50	2.72	3.74	4.81	5.05	3.39
DE	21.85	21.34	28.01	31.81	42.86	38.07	43.40	34.54
DK	30.34	22.54	28.57	37.62	46.12	47.21	43.77	38.79
EE	2.68	7.34	0.00	0.30	0.00	9.19	0.00	2.61
ES	4.19	6.02	6.18	7.86	10.93	10.34	8.88	8.12
FI	87.63	71.05	68.54	43.48	34.78	34.85	43.19	50.01
FR	22.48	38.05	43.46	45.96	46.37	46.36	49.03	42.81
GR	3.39	7.15	7.19	10.41	8.13	10.04	16.40	9.50
HU	2.73	6.24	3.31	3.42	3.93	2.50	3.96	3.65
IE	36.95	43.14	51.89	60.68	71.09	75.22	74.13	62.38
IT	7.68	8.31	7.26	9.19	7.90	9.11	11.78	8.95
LT	0.00	0.15	0.27	0.03	0.11	0.60	0.02	0.19
LU	2.03	17.06	25.84	29.65	30.84	35.48	42.32	26.82
LV	1.03	3.11	1.70	2.11	0.92	1.09	12.99	4.35
MT	10.34	0.00	1.00	0.00	4.66	0.56	26.49	7.62
NL	44.52	52.04	59.41	68.13	66.44	69.51	73.31	63.06
PL	2.26	1.57	2.23	1.91	2.41	2.20	2.58	2.22
PT	2.13	5.03	17.87	23.19	18.30	17.64	21.27	16.88
RO	.	0.98	0.28	0.01	0.03	0.96	1.10	0.62
SE	48.64	51.75	57.76	64.06	77.06	83.03	90.22	69.09
SI	11.56	31.34	27.84	20.74	18.66	22.97	30.58	24.56
SK	3.72	10.67	2.85	3.82	2.72	1.62	6.20	4.46
UK	29.09	34.23	35.96	45.47	43.86	45.76	48.38	41.92
Total	16.63	23.54	25.05	27.25	26.91	25.94	27.07	25.28

The incidence of missing values - Utilities (% of CANs)

country	2006	2007	2008	2009	2010	2011	2012	Total
AT	54.66	53.44	62.64	63.45	56.54	68.11	68.26	61.62
BE	20.04	28.41	28.69	39.58	43.10	49.17	26.48	34.60
BG	.	0.00	0.99	2.41	1.92	2.16	0.83	1.49
CY	0.00	0.00	0.00	7.89	0.00	0.00	20.83	3.49
CZ	0.00	2.30	2.75	5.96	5.25	13.74	16.39	7.46
DE	44.55	57.33	59.70	58.74	65.03	70.21	71.44	61.94
DK	58.78	35.93	52.99	51.20	49.70	63.75	69.01	56.62
EE	3.70	8.33	2.13	1.60	0.00	0.00	0.00	1.78
ES	7.35	9.31	9.89	20.27	45.89	50.29	44.85	31.35
FI	72.33	67.03	62.22	19.20	12.03	20.08	10.86	37.90
FR	51.15	45.42	52.69	53.97	60.94	68.60	70.08	59.24
GR	3.27	14.03	9.74	2.26	11.17	25.84	13.62	10.39
HU	1.62	10.13	5.58	1.83	3.26	0.71	2.80	3.47
IE	72.02	82.64	76.07	87.54	94.04	98.85	94.97	89.68
IT	17.17	17.61	16.31	20.25	20.61	25.04	19.50	19.61
LT	0.62	0.00	0.25	0.00	0.47	0.00	0.00	0.17
LU	98.55	84.21	67.24	68.57	30.43	71.43	100.00	78.60
LV	3.95	3.68	7.83	0.77	0.45	1.14	100.00	14.95
MT	0.00	0.00	0.00	0.00	0.00	15.38	12.50	3.03
NL	62.16	60.00	78.12	80.13	74.60	58.78	57.36	68.40
PL	0.42	4.50	2.98	2.93	3.87	2.95	6.93	3.83
PT	3.17	1.98	3.91	11.03	4.15	8.74	12.86	6.74
RO	.	1.36	0.29	0.18	0.00	0.00	0.27	0.30
SE	26.32	39.60	48.38	64.06	70.59	72.70	73.41	56.06
SI	5.22	14.06	15.49	11.18	6.67	32.31	35.11	20.95
SK	1.22	9.69	10.54	13.75	7.18	7.69	0.00	6.76
UK	46.70	43.90	45.98	47.07	49.83	38.31	48.78	45.62
Total	23.42	24.46	26.77	27.61	33.45	35.39	35.55	30.37

Outline

- 1 The nature of the problem
 - Outliers
 - Missing Values
- 2 The current methodology**
- 3 The new proposed methodology (ongoing work)
- 4 Robustness checks
- 5 Concluding remarks

Imputation of missing values

- Differentiate between Works, Supplies and Services
- Separate treatment of “large” notices (> 100 Million)
- Compute the average of CANs (both classical and utilities) values during the relevant year, for notices between €4.500 and 100 M.
 - For smaller countries in terms of published notices, the average is based on an average of the last 7 years (BU, CZ, DK, EE, IE, EL, CY, LV, LT, LU, MT, NL, PT, RO, SI, SK, FI).
- Compute the number of contract notices (CNs) < 100 M.
- Multiply the average value by the number of CNs \Rightarrow Value of contracts < 100 M.
- Add the value of greater or equal than 100 M. \Rightarrow Estimated value per category.
- Sum the total estimated values for Works, Supplies and Services \Rightarrow Estimated value of tenders published in TED.

Imputation method: Non-parametric estimation of the TFV

- For each $X = \text{Works (W), Supplies (G) and Services (S)}$:

$$TFV_{Xt} = \#CN_X * \frac{\sum_{TFV_X \in (4500, 10^8)} TFV_X}{\#CN_X < 10^8} + \sum_{TFV_X > 10^8} TFV_X$$

- Sum all partial values to get the total:

$$TFV_t = TFV_{Wt} + TFV_{Gt} + TFV_{St}$$

Outline

- 1 The nature of the problem
 - Outliers
 - Missing Values
- 2 The current methodology
- 3 The new proposed methodology (ongoing work)**
- 4 Robustness checks
- 5 Concluding remarks

The imputation model: General view

- The estimated value (EV) of a tender is a natural proxy for the total final value (TFV) of a contract.
- We can estimate how close, on average, the EV and the TFV are, for CANs where both values are known.
- The model can also control for different types of:
 - Procedure (open, restricted,...)
 - Contract (works, supplies, services)
 - Contracting authority (national, regional, local, utility,...)
 - Sector (CPV division or CPV category,...)
 - etc.
- This **imputation model** can be used to obtain an expected value (the *imputed value*) for missing TFVs, based on the EVs and the control variables.

Imputation methods: Model specification

- Estimate TFV for each contract i at time t :

$$TFV_{it} = \beta_0 + \beta_1 EV_{it} + \beta_2 Procedure_{it} + \beta_3 Contract_{it} \\ + \beta_4 Sector_{it} + \beta_5 CAE_type_{it} + \varepsilon_{it}$$

- The TFV at time t is the sum of all estimated values:

$$TFV_t = \sum_i TFV_{it}$$

Controlling for outliers

- Outliers could have an exaggerated effect over the imputed values \Rightarrow They have to be treated apart / excluded.
- We intend to use a conservative definition of outlier \Rightarrow Exclude as little as possible:
 - Values which are intuitively “large” will be excluded and set apart (above approx. three billion Euros).
 - Values which have a disproportionate effect (from an statistical point of view) on the imputation model will be treated separately.
- The imputation model will be applied over the data cleaned from outliers.

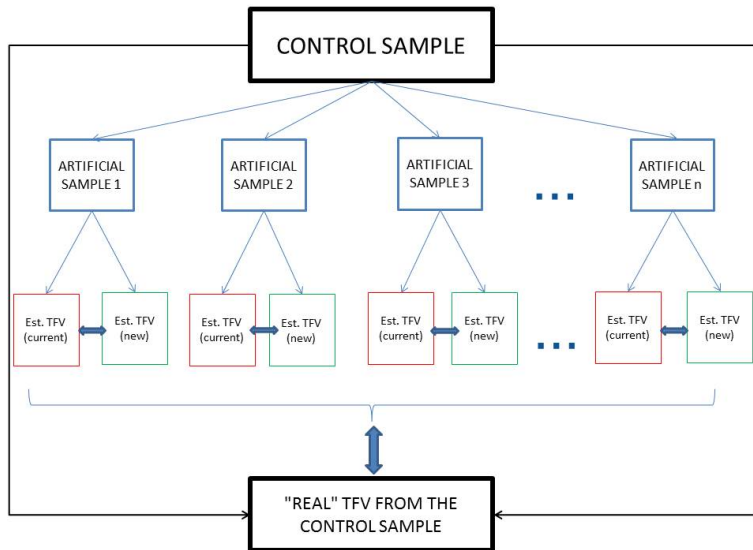
Outline

- 1 The nature of the problem
 - Outliers
 - Missing Values
- 2 The current methodology
- 3 The new proposed methodology (ongoing work)
- 4 Robustness checks**
- 5 Concluding remarks

The effects of changing the methodology

- Change of method \Rightarrow jump in the historical series of estimated values.
- Different impact across countries/years (due to tender-specific factors).
- Sensitivity analysis on the different assumptions of the imputation model.
- Ambitious and computationally intensive exercise. Many elements to play with:
 - The number and type of controls to be included in the imputation model.
 - The number of years to estimate the model.
 - Alternative definitions for outliers.
 - Alternative definitions for sectors.
 - Alternative statistical/econometric methods for the estimation of the model.
 - ...

The robustness check exercise: Comparing the methods



Outline

- 1 The nature of the problem
 - Outliers
 - Missing Values
- 2 The current methodology
- 3 The new proposed methodology (ongoing work)
- 4 Robustness checks
- 5 Concluding remarks**

Concluding Remarks

- Missing values, outliers is a relevant problem in TED data.
- Good reporting practices are fundamental in order to correctly compute the value of procurement published.
- Imputation methods are required to deal with missing information on contract values.
- Different methods will produce different results.
- The choice of method is not about “right” vs “wrong”.
- The current methodology takes a wider, macroeconomic approach, as it is looking at values across sectors within a category of contracts.
- The new methodology takes a more microeconomic approach as it cares more about tender-specific factors to predict the expected value of each contract.
- The method will have to be fine tuned to ensure quality and to check the robustness of estimated value of procurement.

Links

- Contact: Xosé-Luís Varela-Irimia
Jose-Luis.VARELA-IRIMIA@ec.europa.eu
- Public Procurement website:
http://ec.europa.eu/internal_market/publicprocurement/index_en.htm