

**Supplementary material****A. Table of variables**

Description of variables used from SECOP II and other open data sources.

Variable	Stage	Type	Description
<b>Identifier variables</b>			
Contract ID	Adjudication	String	Signed contract identifier.
Buyer name	Pre-contractual	String	Name of the public entity that publishes the contract.
Buyer id	Pre-contractual	String	Tax identification number (NIT) of the public entity that publishes the contract.
Have procurement process	Adjudication	Categorical	Indicator created for those observations that report procurement process information.
Supplier ID	Adjudication	String	Unique identifier of the selected supplier.
<b>Variables for IRIC calculations</b>			
Sign date	Adjudication	Date	Date of signing the contract.
First advertisement date	Pre-contractual	Date	First date in which the procurement process is open to receive offers.
Last advertisement date	Pre-contractual	Date	Last date in which the procurement process is open to receive offers.
Number of suppliers responses	Pre-contractual	Integer	Number of suppliers with responses to offers, within the procurement process
<b>Explanatory variables</b>			
Value	Pre-contractual	Numeric	Total value of the contract in Colombian Pesos (COP).
Budget balance	Pre-contractual	Numeric	Balance of Budget availability certificate of the public entity that publishes the contract in Colombian Pesos (COP).
Number of invited suppliers	Adjudication	Numeric	Total number of invitations made for the procurement process to which the contract belongs.
Number of directly invited suppliers	Adjudication	Numeric	Total number of direct invitations made for the procurement process to which the contract belongs.
Number of answers	Adjudication	Numeric	Total number of public entity answers to offers from suppliers.
Number of processes	Adjudication	Numeric	Number of procurement processes with the same contract.
Number of SIC sanctions	Adjudication	Numeric	Number of sanctions imposed on the supplier by the <i>Superintendencia de Industria y Comercio</i> SIC at the start date of contract execution.
Value of SIC sanctions	Adjudication	Numeric	Total value of sanctions imposed on the supplier by the <i>Superintendencia de Industria y Comercio</i> SIC at the start date of contract execution in Colombian Pesos (COP).
Sign-start contract days	Adjudication	Numeric	Days between the dates of signing the contract and the start of it.
Start-end contract days	Adjudication	Numeric	Days between the dates of start and end of the contract.
Sign-start execution days	Adjudication	Numeric	Number of days between the dates of signing the contract and its start of execution.
Start-end execution days	Execution	Numeric	Number of days between the start and end of the contract execution.
Days supplier registered	Adjudication	Numeric	Number of days from the first registration of the supplier on the platform.
Advance payment percentage	Execution	Numeric	Percentage of contract value paid in advance.

Advertisement period	Adjudication	Numeric	number of days between the first and last date of publication of the procurement process on the platform.
Balance/value	Pre-contractual	Numeric	Balance of Budget availability certificate of the public entity that publishes the contract as proportion of the contract value.
Sector	Pre-contractual	Categorical	Sector to which the public entity publishing the contract belongs.
Centralized entity	Pre-contractual	Categorical	Indicator about whether the public entity that publishes the contract is centralized or decentralized.
Contract type	Pre-contractual	Categorical	Type of contract according to its judicial framework
Procurement mechanism	Pre-contractual	Categorical	The process of finding and agreeing to the terms, for the acquisition of goods, services, or works of external sources.
Procurement mechanism justification	Pre-contractual	Categorical	Justification of the scenario that defines the procurement mechanism.
Supplier ID type	Adjudication	Categorical	Type of adjudicated supplier identification document.
Is group	Adjudication	Categorical	Indicator about whether the supplier is, or not, a group of companies.
Is pyme/sme	Adjudication	Categorical	Indicator about whether the supplier is, or not, a <i>PYME/SME</i> , small or medium company ( <i>PYME</i> is its acronym in Spanish).
Allow pay in advance	Pre-contractual	Categorical	Indicator about whether the contract allows a percentage of payment in advance or not.
Sold out	Execution	Categorical	Indicator about whether the contract is, or is not, liquidated.
Environmental obligation	Pre-contractual	Categorical	Indicates whether the contract has, or has not, commitments to comply with environmental obligations.
Post-conflict	Pre-contractual	Categorical	Indicates whether the contract is, or is not, associated with some event of the peace agreement.
Expenditure purpose	Pre-contractual	Categorical	Expenditure destination at budget level: investment or operation.
Product segment to contract	Pre-contractual	Categorical	Code of the main category in the UNSPC in which the object belongs.
Company type	Adjudication	Categorical	Type of company registered by the supplier on the platform.
Month of sign	Adjudication	Categorical	Month of the year in which the contract was signed.
Month of publication	Adjudication	Categorical	Month in which the first publication of the process on the platform was made.
Product group to contract	Pre-contractual	Categorical	Code of the group in the UNSPC to which the object to be contracted belongs.
Company type aggregation	Adjudication	Categorical	Aggregation of the company type registered by the supplier on the platform.
<b>Target variables</b>			
Cost overruns	Execution	Categorical	The contract did or did not present cost overruns.
Delivery delays	Execution	Categorical	The contract did or did not present delivery delays.
Additions	Execution	Categorical	The contract did or did not present cost overruns or delivery delays.

**B. Tables of explanatory variables descriptive statistics**

Descriptive statistics of variables used in the inefficiencies models and irregularities indices.

**Numeric variables**

Variable	Mean	Std	25-th perc	50-th perc	75-th perc	Max
Value	72454998.58	797445592.2	10025890	20444050	41753400	85858937325
Budget balance	4648927200.4	51430017925.6	0	25000000	116975500	3.3E+12
Number of invited suppliers	5.75	34.91	0	0	0	1866
Number of direct invited suppliers	0.16	2.13	0	0	0	155
Number of answers	0.02	0.70	0	0	0	153
Number of processes	1.03	0.18	1	1	1	3
Number of SIC sanctions	0.20	7.21	0	0	0	687
Value of SIC sanctions	13382622.59	539294765.33	0	0	0	64304254243
Sign-start contract days	-9.84	37.53	-4	0	1	335
Start-end contract days	217.44	126.46	110	187	333	3704
Sign-start execution days	-1.76	53.87	-3	0	2	742
Start-end execution days	211.79	126.6	104	181	331	3722
Days supplier registered	333.33	302.64	67	281	474	1969
Advance payment percentage	0.0001	0.01	0	0	0	0.5
Days process open	0.61	4.01	0	0	0	300
Balance/value	83205.57	23171818.75	0	1	1.37	6566250000

**Categorical variables variables**

Variable	Unique	Top category	Top relative freq
Sector	14	Health and social services	44.5%
Centralized entity	2	No	76.0%
Contract type	10	Professional services	90.9%
Procurement mechanism	12	Direct contracting	59.0%
Procurement mechanism justification	14	Professional services	58.3%
Supplier ID type	4	Colombian Identity Card	88.4%
Is group	2	No	99.3%
Is pyme	2	No	87.0%
Allow pay in advance	3	No	99.8%
Sold out	2	No	91.3%
Environmental obligation	2	No	94.4%
Post-conflict	2	No	99.8%
Expenditure purpose	3	Investment	52.7%
Product segment to contract	55	80 (Management, Professional Business and Administrative Services)	78.3%
Company type	32	Natural person	87.3%
Month of sign	12	Feb	30.5%
Month of publication	12	Jan	28.4%
Product group to contract	6	Administrative services	78.3%
Company type aggregation	10	Natural person	87.3%

### C. Distributions in advertisement and decision periods

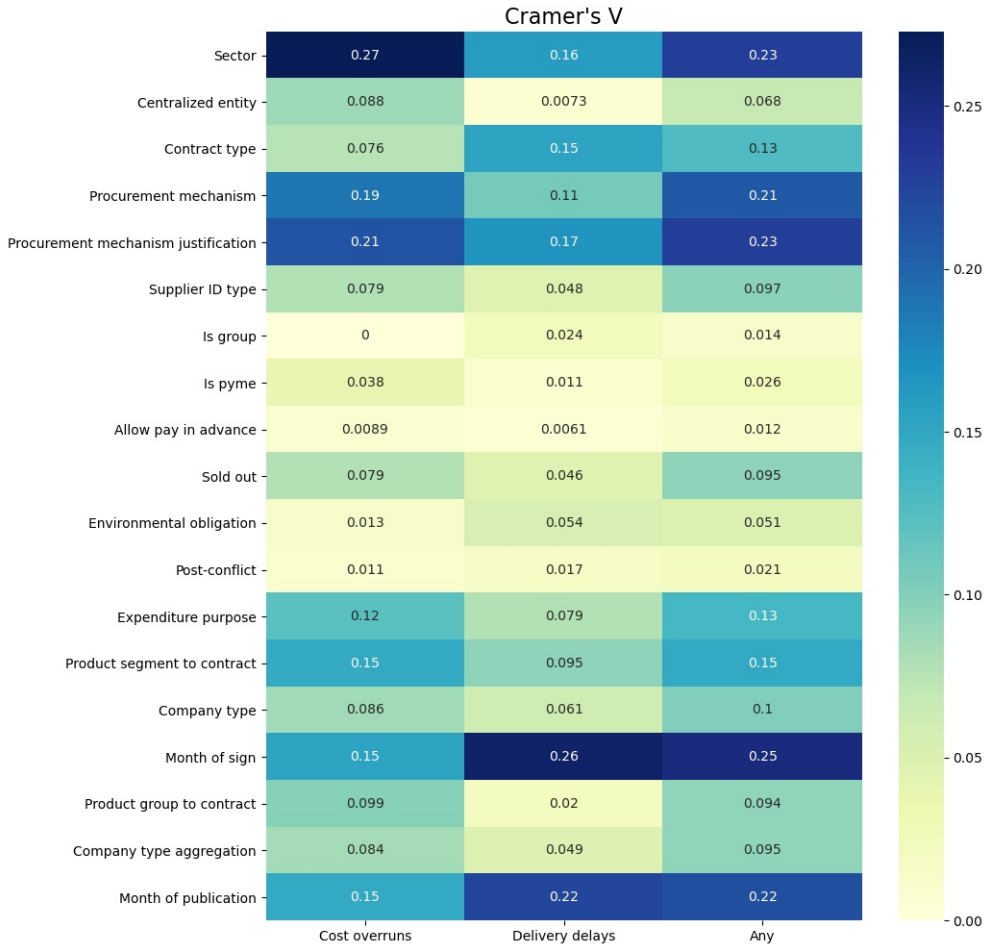
Descriptive statistics of the procurement processes advertisement period and decision period by contract type (professional services vs non professional services). These distributions were used to find the thresholds that define the red flags of extreme advertisement and decision periods.

Type of contract	Mean	Std	Min	10%	50%	95%	97.5%	99%	Max
<b>Advertisement period (days)</b>									
Professional Services	0.25	2.43	0	0	0	0	0	14	188
Non Professional Services	2.98	8.82	0	0	0	19	22	31	300
<b>Decision period (days)</b>									
Professional Services	9.97	26.14	0	0	2	43	97	125	785
Non Professional Services	19.29	33.53	0	0	8	55	71	125	988

### D. Explanatory variables and inefficiencies correlations

Pearson correlation and Cramer's V between each explanatory variable, and the targets cost overruns and delivery delays.





**E. SHAP values cost overruns model**

Beeswarm plot to summary how the features impact the probability of cost overruns using all variables and all contract types.

