Lowering prices of pharmaceuticals, medical supplies and equipment:

Insights from Big Data for better procurement strategies in Latin America

Mihály Fazekas (Central European University)

Alexandre Borges de Oliveira (World Bank)

and Nóra Regős (Government Transparency Institute)

Evidence-Based Policy Making - A collaborative event of WIFO, IHS and CEU, 27/05/2021

SCHOOL OF PUBLIC POLICY
purpose beyond power®
CENTRAL EUROPEAN UNIVERSITY

Main arguments

- Types of evidence needed for making good policies
- Example evidence-type
  - Descriptive and predictive, not explanatory
  - Policy uses

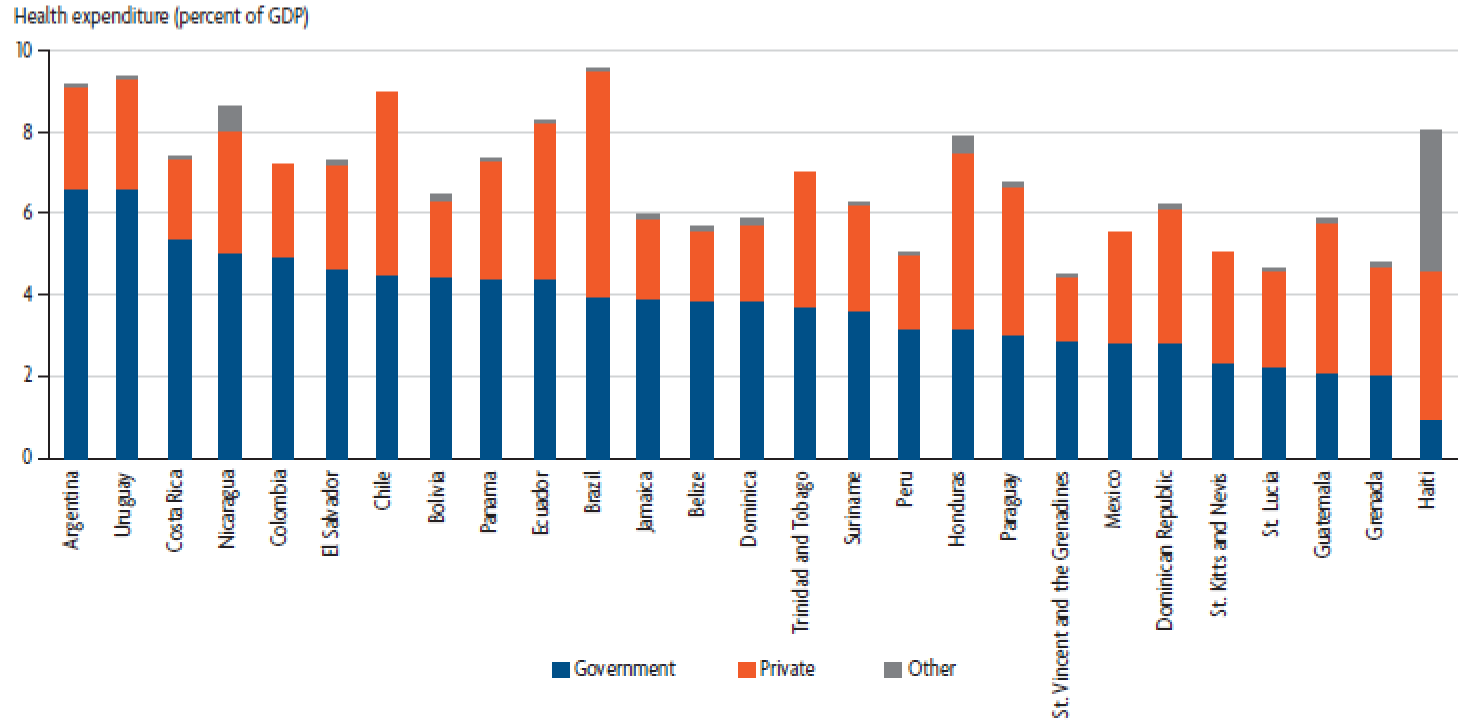
Evidence-based policy making

- Academia praises:
  - Explanatory, especially causal evidence
  - But it is hard to come by high quality, generalisable causal evidence
- BUT policy often needs:
  - Descriptive
  - Predictive
- WHY?
  - In many policy problems the scale of the problem/phenomenon is the most important question for mobilizing resources
  - In many policy problems expectations under the business as usual scenario is the baseline

The paper

Motivations

- High amounts of public spending on healthcare in LAC



Source: WHO.

- Cost explosion
- Large variance in unitprices within country-within market

Data: scope

Cross-country, large-scale, micro-level data

Unit of observation: contracted item

Country	Total number of health care-related tenders in country dataset	Total number of health care-related tenders in combined healthcare dataset	Years covered
Ecuador	22609	22609	2013-2017
Brazil (federal)	14108	2140	2014-2016
Amazonas (Brazil)	9030	2797	2014-2018
Santa Catarina (Brazil)	1348	948	2013-2018
Paraguay	2899	830	2012-2016
Panama	56738	5439	2014-2018
Uruguay	12319	1008	2014-2018
Peru	9217	686	2015
Costa Rica	517	154	2016-2017
Total	128785	36611	

Data: detail

Harmonized product classification

Country	Classification scheme	Level of observation	Total number of categories in country dataset (at level of obs)	Total number of categories overlapping with master file (Ecuador)
Ecuador	CPC - Clasificación Central de Productos	Level 5 - CPC Level 5	222	222
Paraguay	UNSPSC - United Nations Standard Products and Services Code	Level 4 - Descripción ítem nivel4 (Commodity)	5122	217
Panama	UNSPSC - United Nations Standard Products and Services Code	Level 4 - Nombre Rubro (Commodity)	5122	217
Uruguay	SICE - Sistema de Información de Compras y Contrataciones del Estado	Level 5 - Descripción Artículo	1097	166
Peru	CPC - Clasificación Central de Productos	Level 4 - Commodity	1763	203
Brazil (federal)	Catálogo de Materiais e Serviços	Level 3 - Padrao	2445	181
Amazonas (Brazil)	Catálogo de Materiais e Serviços	Level 3 - Classification Commodity	2211	195
Santa Catarina (Brazil)	Catálogo de Materiais e Serviços	Level 3 - Item Classification Description	5996	205
Costa Rica	UNSPSC - United Nations Standard Products and Services Code	Level 5 (free text) - DescProducto	692	92

Research goals

Introducing a comprehensive framework for:

- Predicting prices and
- Quantifying potential savings achievable through policy changes

in the procurement of standardized healthcare products, but applicable approach to government-wide puchases

Conceptual framework

- Data science-oriented, theory-driven, policy relevant framework

- Comparison of different models
  1. Simple OLS
  2. Interacted OLS; and
  3. Random forest (Ntree=175, m=6).

- Model performance assessed based on test dataset (25% of sample)
  – R2
  – MSE

Methods: regression specification

$$Pr_i = \alpha_i + \beta_1 * X_{1i} + \beta_2 * X_{2i} + \beta_3 * X_{3i} + \varepsilon_i$$

- Pr_i represents log unit price for the *i*th item purchased;
- X_{1i} stands for the set of directly policy influenceable predictors for the *i*th item purchased such as the choice of procedure type;
- X_{2i} represents the set of indirectly policy influenceable predictors for the *i*th item purchased such as the number of bidders.
- X_{3i} denotes the set of control variables accounting for structural factors not amenable to policy intervention for the *i*th item purchased such as the year of purchase, or country.
- ε_i stands for the error term of the regression model.

Variables in the analysis

Type	Group	Variable name	Types	mean/ most frequent values*	std.div	N(non- missing
DV	**-**	Unit price	continuous	8512.9	453443.4	284,872
Structural	**Market characteristics**	Market ID: reflecting product code (1…235)	categorical	159, 194		287,041
		Year of contract award (2012- 2018)	categorical	2014		287,041
	Buyer characteristics	Buyer type (independent agency, ministry, etc)	categorical	National gov.		286,634
		Buyer location (region)	categorical	EC Pichincha		287,041

Variables in the analysis

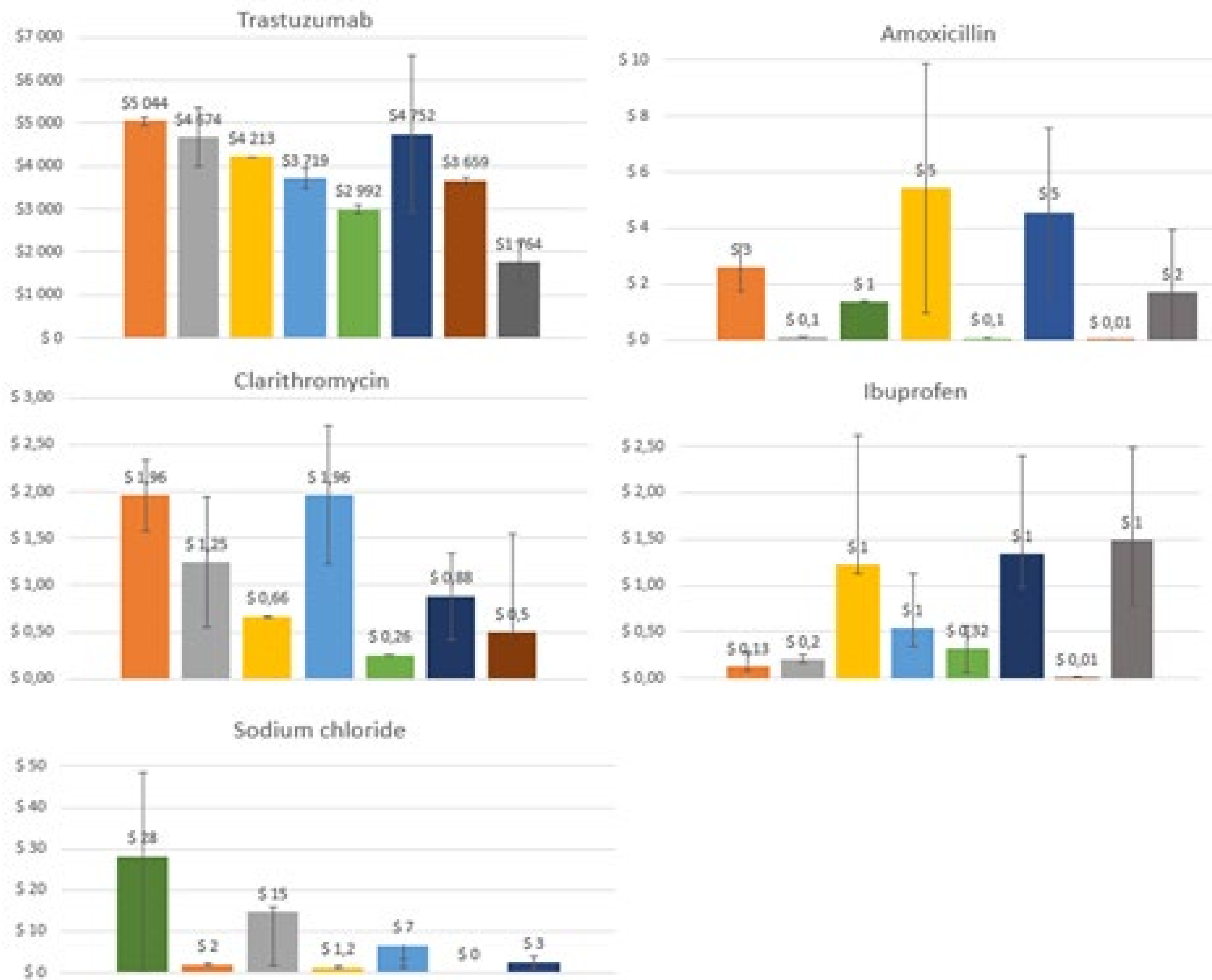
Directly policy influenceable	**Tender specifications**	Month of spending (January, February, etc.)	categorical	July		33,155
		Procedure type (1 - fully competitive; 2-restricted, etc.)	categorical	Fully competitive		279,720
		Advertisement period length (days)	continuous	15.88	17.6	34,923
		Decision period length (days)	continuous	29.8	33.07	30,491
		Failed tenders (%)	continuous	0.306	0.461	22,398
		Framework agreement (Y/N)	binary	Yes		287,041
		Product bundling	continuous	34.8	38.5	287,041
		Quantity of purchased goods (number of units)	continuous	23198.3	105542.8	287,041

Variables in the analysis

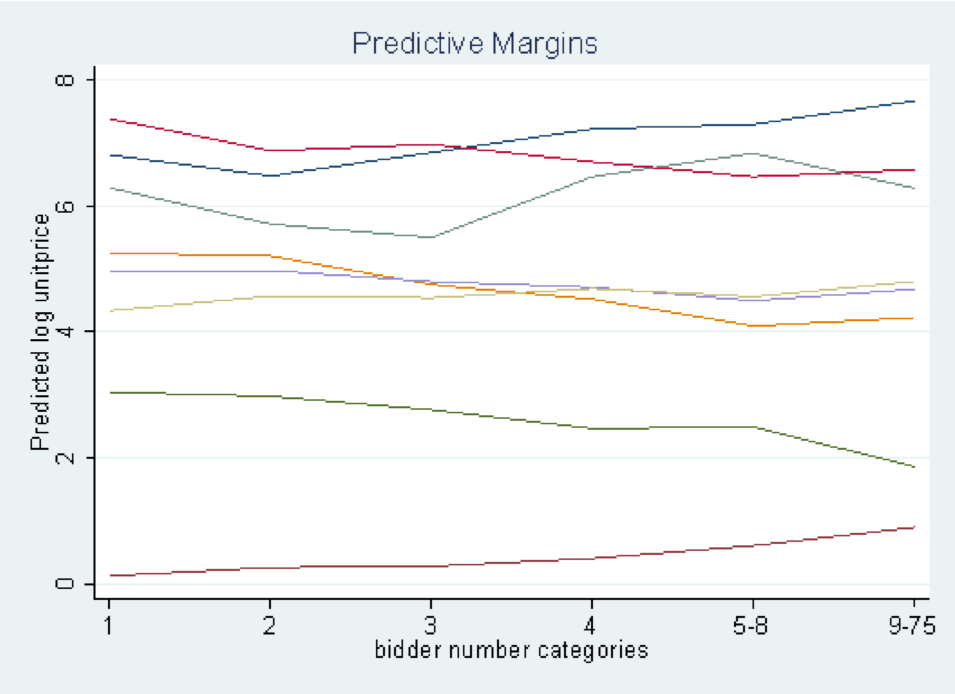
Indirectly policy influenceable	**Bidder/ supplier characteristics**	Buyer-supplier from the same state (Y/N)	binary	No		282,028
		Supplier size (micro, small, large company)	categorical	large		258,341
		Supplier specialisation: number of markets the company supplies	continuous	83.8	39.6	284,718
	Bidding outcomes	Number of bidders	continuous	4.3	7.2	48,703
		Annual winner market share (%)	continuous	0.570	0.384	284,614
		Annual winner share in buyer spending (%)	continuous	0.070	0.147	284,614

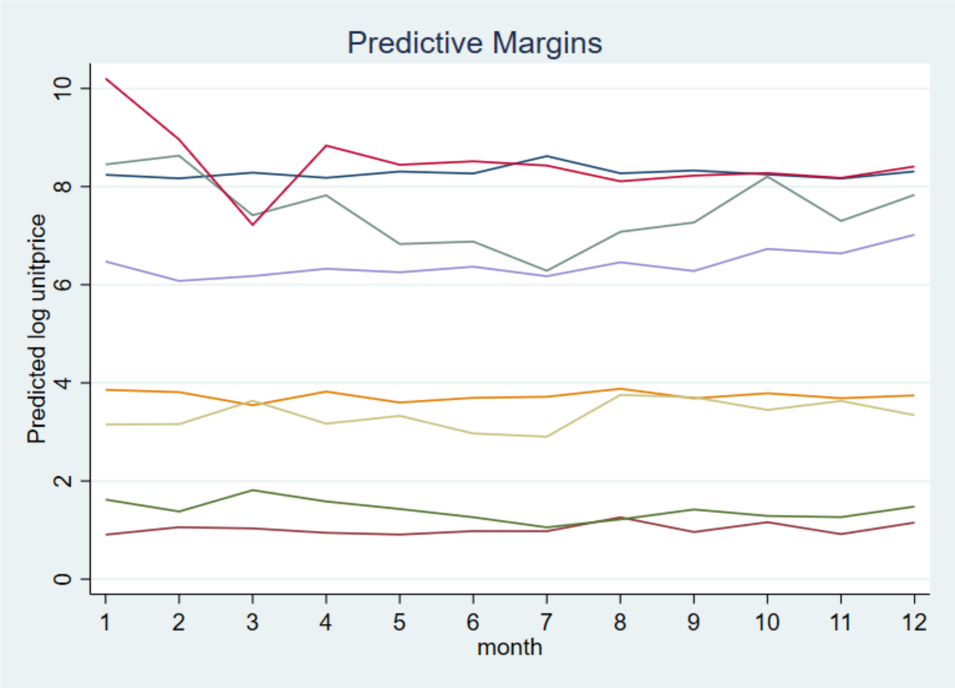
Results I: Surprising unit price variation

Average unit prices and interquartile ranges by country and territory, selected pharmaceuticals, USD

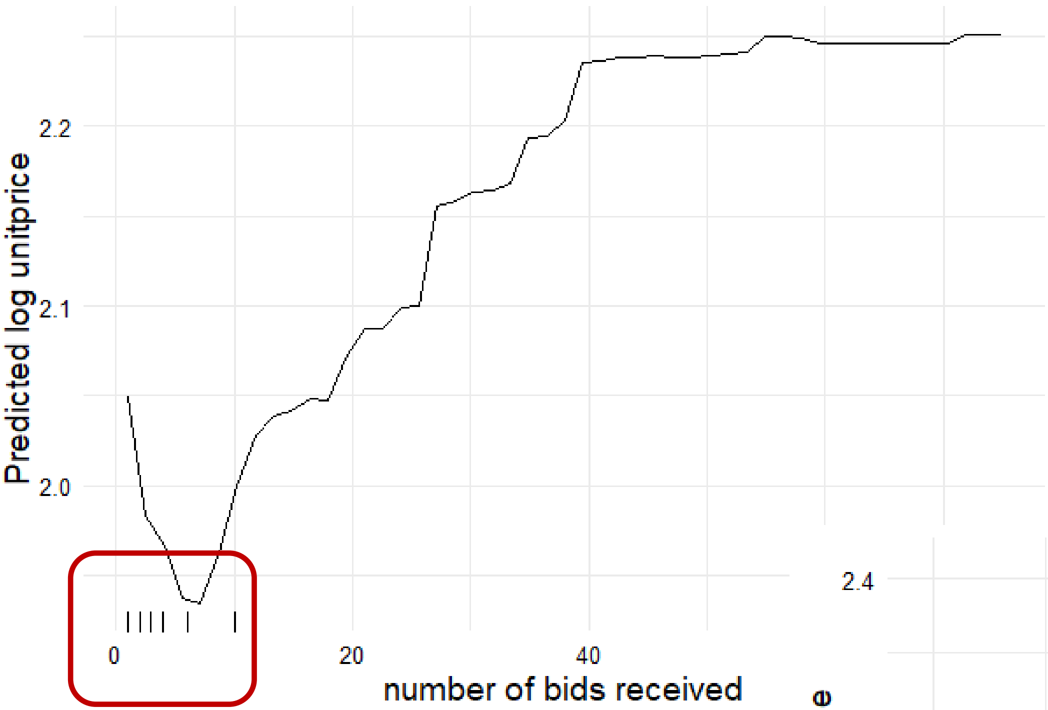


Results II: OLS regressions (interacted)

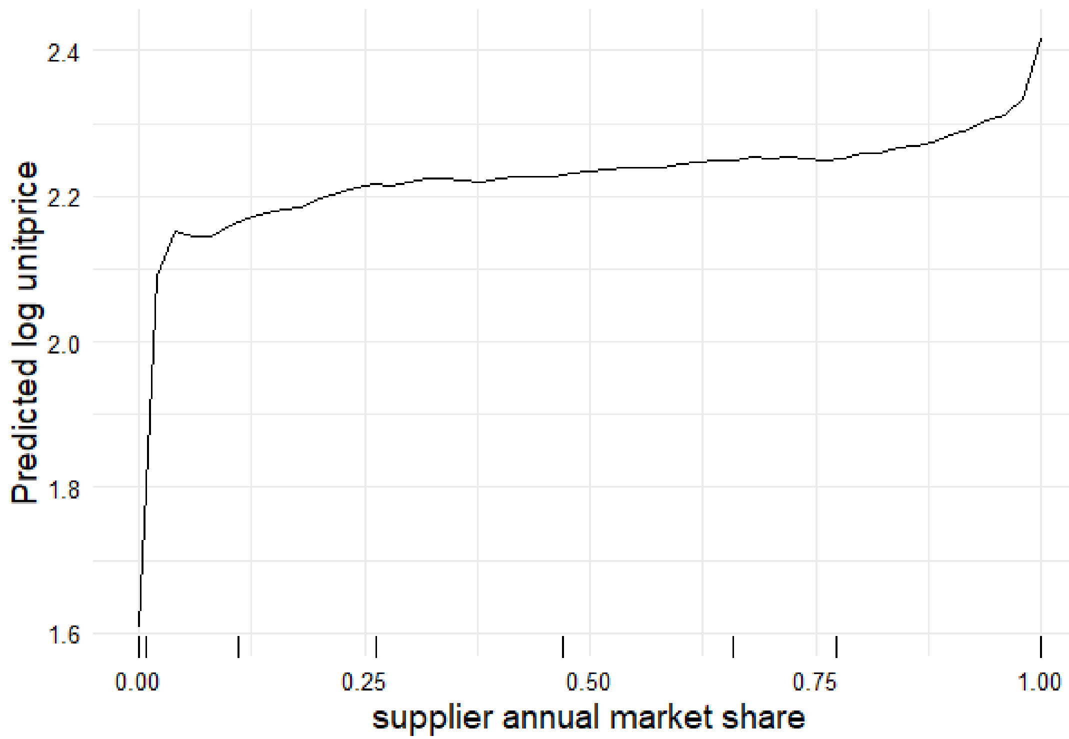




Results III: Random forest



90% of the data falls in the 1-10 range



Model selection

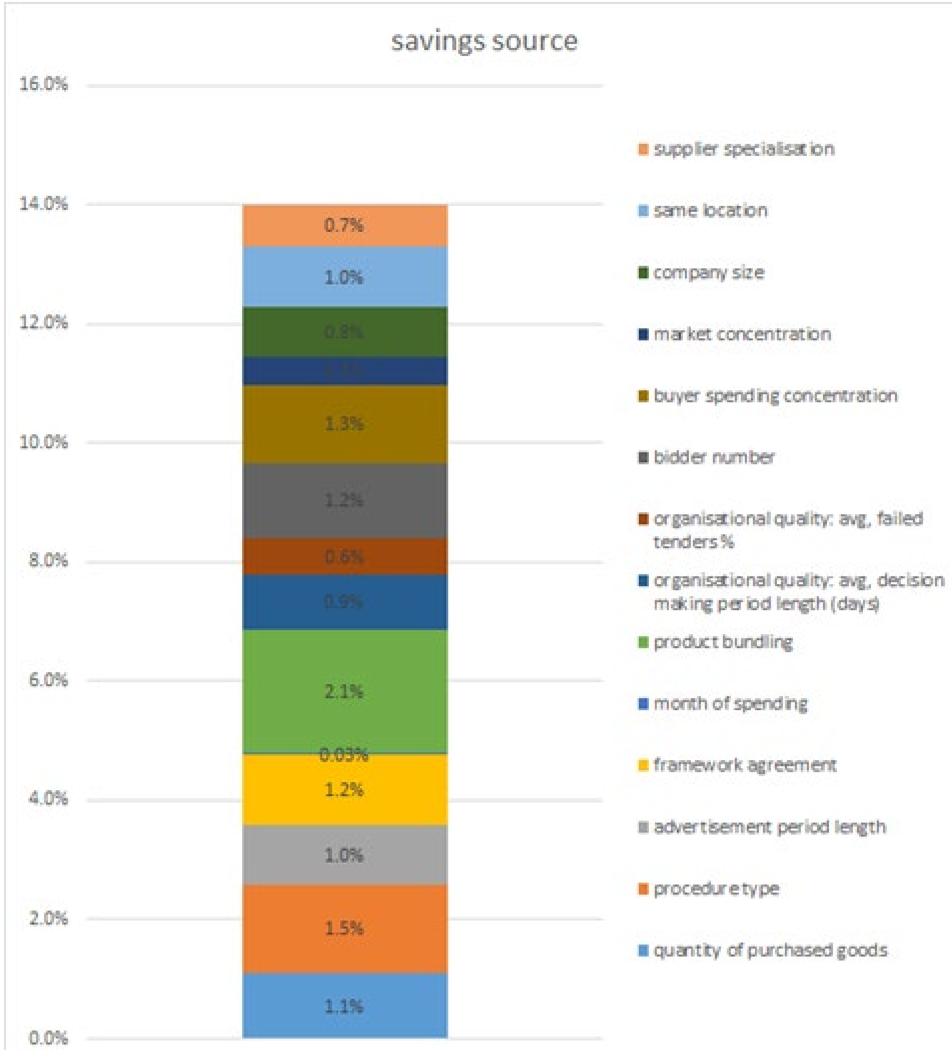
Model performance, test dataset (25%)

Model	R2	MSE
OLS (Table 5, model 3)	0.886	2.612
Interacted OLS (Table B1, Model 3)	0.887	2.401
Random Forest	0.847	1.451

Savings analysis: policy scenarios

directly policy influenceable	quantity of purchased goods	moving 30% of the lower deciles (582< units bought) to the next higher decile (583-3000 items bought)
	procedure type	move 5% of items of competitive procedures to restricte procedure types
	advertisement period length	move 20% of shorter advertisement periods (1-12 days to longer advertisement (13-183 days)
	month of spending	smooth spending across for Dec, Jan and Febr by reallocating 30% of items to a nearby cheaper month (March)
	product bundling	moving 5% of items in the 1st decile to the 2nd decile
	framework agreement	moving 2% of contracts without framework agreement t with framework agreement
	organisational quality: avg. decision making period length (days)	moving 7,5% of tenders in the longest quintiles to the 3 quintile
	organisational quality: avg. failed tenders %	move 50% of items from the lowest half success rate organisations (< 67.5%) to highest success rate organisations (>67.5%)
indirectly policy influenceable	bidder number	moving 13% weakly competitive items (1-2 bidders) to more competitive items (2-7 bidders)
	buyer spending concentration	move 5% of items from high spending concentration buyers (the top 4 highest deciles of buyers) to average spending concentration buyers (5th decile)
	market concentration	move 50% of items in the highest concentration market (the top highest decile market) to lower concentration markets (8th highest decile)
	same location	increase market share of local suppliers (i.e. same state by 10%
	company size	decrease share of small- and medium-sized companies by 20% (to the advantage of large companies)
	supplier specialisation	moving 3% of items supplied by highly specialised suppliers (1-3 lowest deciles) to average specialised suppliers (4th decile)

Savings estimates



Policy uses

- Embedding evidence in day-to-day policy making
  - Large volumes of data, real time monitoring, short feedback loops
- Main uses
  - Budgeting: predicting expected unit prices (e.g. Brazil)
  - Planning purchases strategically, e.g. buying in bundle (e.g. Ecuador)

Further background information

Mihály Fazekas, Alexandre Borges de Oliveira, and Nóra Regös (2021), *Lowering prices of pharmaceuticals, medical supplies and equipment: Insights from Big Data for better procurement strategies in Latin America*. Policy Research Working Paper, World Bank, under peer review

Contact

Mihaly Fazekas

FazekasM@spp.ceu.edu