Redesigning Framework Agreements in Chile Reduces Government Spending

Marcelo Olivares[1,2], Daniela Saban[3], Gabriel Weintraub[3], Eduardo Lara[1,2], Piero Zanocco[1], and Paula Moreno[4]

[1]Universidad de Chile
[2]Instituto Sistemas Complejos de Ingeniería
[3]Stanford Graduate School of Business
[4]ChileCompra

Abstract

Framework agreements (FAs) are procurement mechanisms used in private and public organizations by which a central procurement agency selects an assortment of products, typically through auctions, and then affiliated organizations can purchase from the selected assortment as needs arise. During 2018 and 2019, the Chilean government central procurement agency (ChileCompra) purchased an average of US$2.8 billion worth of products and services per year (23% of its total budget) using FAs. Analyzing the data from the 2014 Food FA, used between 2014 and 2017 to purchase US$200 million worth of products annually, we observed that ChileCompra FAs exhibited low levels of competition in the auctions used to select the assortment, which could potentially result in larger government expenditures. With this motivation, we collaborated with ChileCompra to redesign FAs to enhance competition. We used the 2017 Food FA as a pilot study. Our main changes were twofold. First, using NLP algorithms, we standardized the catalogue of products that were required by ChileCompra to be part of the FA. Second, taking advantage of this standardization, we implemented an experimental design in the FA auction, such that different product categories faced different levels of competition to become part of the assortment. We analyzed the impact of the redesign through a difference-in-differences regression analysis and showed that inducing more intense competition in the auction stage reduced transaction prices by 8%. This pilot study ultimately led ChileCompra to implement a similar design in all of its FAs, and many of the improvements in the design of the FAs were included in the new regulation on government purchases. If we were to extrapolate the savings from our pilot re-design to all of these FAs, the total savings would amount to around US$64 million in 2022.

1 Introduction

When conducting their procurement processes, large organizations can choose among different mechanisms to select their supplier base. With a centralized approach, the organization may choose to run an auction to select a single supplier. By contrast, with a decentralized approach, the purchasing decisions are made by the different suborganizations or units, which may run their own auctions or buy the products they need in the open market. Although there is an

increasing trend to centralize purchases through central procurement bodies (see, e.g., OECD (2019b)) there are pros and cons to procurement centralization. On the one hand, centralization may allow for tighter control of expenditures and may be able to take advantage of the economies of scale resulting from aggregating the demand of the different purchasing units. On the other hand, decentralization may be advantageous when purchasing units have heterogeneous needs and local information on the supplier base and on the market dynamics, which may lead to more efficient procurement (Dimitri et al., 2006).

Framework agreements (FAs) provide an intermediate approach between centralization and decentralization of purchases. In a FA, a central procurement unit preselects an assortment of products and suppliers; then, each affiliated organization can purchase from this selected assortment as needed. Examples of such mechanisms include the purchase of computers in a university and health plans with lists of prescription drugs available to plan enrollees, among others. Due to their flexibility, FAs are nowadays recognized as a fundamental tool in public procurement (Albano and Nicholas, 2016). As an example, our collaborator, the Chilean government central procurement agency (Dirección ChileCompra, or ChileCompra for short) purchased US$3,000,000,000 worth of goods and services though FAs in 2018, which represented 22.6% of the value of all public procurement in Chile (ChileCompra, 2022).

In government FAs, the preselection of suppliers is usually conducted through a first-price auction where bids correspond to prices for those suppliers who are added to the FA. Importantly, once selected, the set of products and suppliers that constitute the FA assortment will typically be available for several years, and the government organizations can then buy from this set as needs arise. Moreover, the bids in the auction effectively act as ceiling prices: while the FA is in place, the suppliers are allowed to decrease their prices through promotions or price changes, but price increases are heavily regulated and are mostly governed by inflation.

Therefore, when designing the rules of the auction, the central procurement unit needs to carefully consider the level of price competition that these rules will induce in order to optimize over the following trade-off. On the one hand, rules for which only a few products end up being in the FA (more competitive auction rules) may increase suppliers' incentives to place aggressive bids in the auction, so that their products have a better chance of being part of the small selection of items. Thus, this will result in lower ceiling prices. On the other hand, rules

for which it is "easy" to be part of the FA (less competitive rules) may result in higher bids and hence higher ceiling prices. At the same time, they will result in more suppliers being added and thus, potentially, in more intense price competition and more promotions during the FA. As the government ultimately cares about spending (and hence the prices at which *transactions* occur), a natural question is whether more competition in the FA auction will effectively reduce government spending.

In theoretical work, Saban and Weintraub (2021) (see also the follow-up paper, Choi et al. (Forthcoming)) study these sources of competition in stylized models of FAs and show that increasing competition in the auction stage typically reduces prices and improves market outcomes. In this paper, we report on a market design intervention in which we redesign a large FA run by the Chilean government to understand the effect of increasing competition at the FA auction. We show that this indeed results in important savings for the government.

An initial descriptive analysis of ChileCompra's 2014 Food FA, used by the Chilean government between 2014 and 2017 to purchase US$200 million worth of products annually, reveals low levels of competition at the auction used to define the FA assortment. For instance, about half of the auctions in the FA received a single bid and 75% of all bids were awarded. The high number of auctions with a single bid may be explained by the fact that suppliers self-reported their product attributes using free unstructured text, which resulted in identical products being treated as different ones and thus not being part of the same auction. As stated above, the bids in these auctions act as ceiling prices; thus, this observed low competition in the auction may have resulted in higher ceiling prices and, potentially, in higher transaction prices.

With this motivation, together with ChileCompra, we redesigned the 2017 Food FA. The new design consisted of two main changes. First, we built algorithms using NLP to standardize the catalogue of products that would be required in the auction, and defined each product based on measurable product attributes. Second, taking advantage of this standardization, we implemented an experimental design in the FA auction, such that different product categories faced different levels of competition to become part of the FA, allowing us to measure the impact of increasing competition in the auction. Specifically, we set different thresholds to award the winning bids: the lowest 20% (competitive treatment) and the lowest 80% (non-competitive treatment) of all received bids for a product would be awarded and added to the FA, depending

on which treatment the product category was facing.

We measured the impact of this intervention on submitted bids, winning bids, the prices that were posted in the ChileCompra marketplace, and the transaction prices, using a difference-in-differences estimation that matched products between the old and new auction designs. The empirical results show that products that were in the 20% competition threshold treatment in the auction stage had 8.1% lower awarded median bids, and 8.2% lower transacted median prices. These savings in the procurement cost of the government amounted to US$3.6 million yearly in the treatment group and to US$11 million yearly if we extrapolate them to the entire Food FA.

After the successful implementation in the food FA, ChileCompra started to implement a similar design in all of its FAs, and many of these improvements were included in the new regulation on government purchases (ChileCompra, 2021a). By 2022, most of the operating FAs had adopted the new design (ChileCompra (2021b),ChileCompra (2022)). Similar to the Food FA, these new FAs incorporated a structured product catalogue (with the assistance of automated algorithms to structure product attributes) and enforced tighter competition to enter the FA using lower thresholds to award bids. ChileCompra executives reported successful implementations and if we extrapolate the savings from the Food FA to all the FAs operating in 2022, the total savings would amount to around US$64 million per year.

Overall, our market redesign and empirical study show the value of restricting competition in the FA auction, providing important guidelines on how to implement FAs in practice. Furthermore, our work contributes to the literature on applied market design in which theory together with a deep understanding of the institutional and operational details of the market are used to improve outcomes (Vulkan et al., 2013).

2 Background: Framework Agreements in Chile's Public Procurement System

In this section we give a brief description of FAs and the potential problems we have identified in FAs in Chile.

What is a FA? A framework agreement (FA) is a purchasing mechanism by which a purchasing organization (in our case, ChileCompra) selects a set of suppliers and agrees on the terms and conditions that will be applied to any subsequent transaction. At a high level, an FA can be thought of as a process consisting of two main stages: an auction stage and an operation stage.

In the auction stage, ChileCompra publishes the rules of the auction, including which products they are seeking to purchase and the rules to decide which suppliers and products will be offered as part of the FA. Suppliers can then choose which products to bid for; a bid includes the price for the product as well as additional information (we explain the bidding process in detail below). There is typically one auction per product, but some auctions aggregate more than one product based on similarity of product attributes. For instance, perishable products of different brands but identical remaining attributes competed in the same auction. Using the published rules, ChileCompra determines which supplier–product combination is added to the FA, i.e., added to ChileCompra's online marketplace. This concludes the auction stage.

In the operation stage—which for the FAs implemented before 2018 lasts for about four years—the public organizations can purchase products from the assortment in the FA auction stage through ChileCompra's online marketplace. In the case of the Food FAs, an organization can buy *any* product available in the selected assortment without the need of running additional tendering. This allows public organizations to purchase goods in a more dynamic and agile way, speeding up the buying process. Importantly, during the operation stage, suppliers can make price and stock adjustments to the products they offer.

The following subsections describe in detail the auction stage of a FA run by ChileCompra in 2014, which sourced food products to the public organizations in Chile, and operated from November 2014 to August 2018. Hereafter, we refer to that FA as the 2014 Food FA or simply FA 2014.

2.1 The rules for the 2014 Food FA

FA definition. In 2014, ChileCompra carried out an auction for the FA of perishable and non perishable foods. Within food items, ChileCompra defined a preliminary list of products of interest, and potential suppliers were invited to submit bids for these products. Each product

comprised three attributes defined by ChileCompra: category (e.g., drinks), type of product (e.g., juice), and brand (e.g., Andina). Potential suppliers were allowed to submit bids for any product matching these attributes. We describe the rules that decided the winning bids in detail below. After the winning bids were selected, the FA was expected to be in operation for three years.

Bidding process. To submit a bid for a product, potential suppliers had to provide two additional attributes: model (e.g., orange 250 ml.) and units (e.g., 4 bottles). Once a product was defined, suppliers had to provide a price; this price will effectively act as a ceiling price during the operation of the FA.[1] Suppliers also had to provide general purchasing information, including compliance with regulations, shipping rates for the different regions, and more.

In addition, each seller could choose to submit bids for up to 60 new products (after defining all five attributes of these products), which were not included in the original catalogue provided by ChileCompra.

Auctions were defined differently depending on whether the product was perishable or non-perishable. For nonperishable products, all bids submitted for the same product competed in the same auction. (Recall that a product is uniquely defined by five attributes.) For perishable products, suppliers competed only with other suppliers located in their same region. That is, the auctions were defined locally based on the location of the supplier.

Awarding the set of suppliers. A supplier's bid was assessed according to a weighted sum of scores associated with five dimensions, including price, technical requirements, sustainability, shipping costs, and volume discounts. Each dimension score was normalized to a 0–100 scale. The score for price was computed relative to the minimum bid price among all bidders in the auction (excluding extreme outliers), and received a 60% weight in the final score. Scores for the shipping costs and volume discounts were also computed relative to minimum bids among all suppliers, whereas sustainability and technical requirements were calculated based on specific criteria that do not depend on the others' bids. In terms of the weight for the final score, regional shipping costs accounted for 15%, volume discounts for 20%, sustainability for 3%, and technical

[1] As we later discuss, suppliers are allowed to decrease the price during the operation phase of an FA; by contrast, price increases are heavily regulated and can, for the most part, only be adjusted for inflation as dictated by the national inflation index.

Table 1: Summary of the 2014 Food FA auction stage

	# Products	# Auctions	Avg. # bids	# Single-bid Auctions
All products	6,563 (100%)	11,676	3.4	5,977 (51.2%)
Defined by ChileCompra	4,187 (63.8%)	9,510	4.0	3,932 (41.3%)
Defined by suppliers	2,376 (36.2%)	2,166	1.1	2,045 (94.4%)

Notes: This table shows the results of the auction for the 2014 Food FA for products defined by ChileCompra and the suppliers, respectively. We label a product as being defined by ChileCompra if the first three attributes were among those specified by ChileCompra. Otherwise, a product is said to be supplier-defined. We say that a product was added to the FA if at least one supplier selling that product was added to the FA. A single-bid auction is an auction in which only one bid is submitted. Note that there are more auctions than products because bids can be regional for perishable products.

requirements for 2%. For more details on the score calculations, see Levy (2017). All bids with a total score above 70 points were awarded. In addition, if a supplier was awarded 80% or more of the bids, all of its bids were awarded.

2.2 Potential issues with the 2014 Food FA auction: Noncompetitive awarded bids

We collected bid data for the 2014 Food FA auction, which is summarized in Table 1. Of the 11,676 auctions, 51% received only one bid, with no competition upon entry. Moreover, auctions receiving more than one bid had on average 3.4 suppliers bidding and more than 75% of the bids awarded, which suggests limited competition. The lack of competition was particularly severe for products where suppliers defined the product attributes, leading to 94.4% of these auctions receiving a single bid and 1.1 bids per auction on average. Similar findings were reported by the OECD (2019a).

At first glance, this lack of competition may appear to be worrisome. For instance, a common finding in the auction design literature (see, e.g., Myerson (1981); Klemperer (2004)) is that, in general, attracting more bidders to the auction will decrease the bids and hence the purchasing prices. However, this finding applies to the standard auctions, where the auctioneer commits to buying the item from the winner but may not apply to framework agreements. To see why, we next explain the different dimensions of price competition in FAs.

Price competition in the FAs. Broadly speaking, there are two different ways in which suppliers may compete in prices during a FA (see also Demsetz (1968)).

1. *Price competition in the auction stage to enter the FA*. In general, whether a supplier is included in the FA depends on the rules of the auction and the bids placed by it and the other suppliers participating in the same auction. By placing a lower bid, a supplier typically increases its chances of being part of the FA. Obtaining low bids may be important as bids effectively act as ceiling prices. Therefore, more competitive auctions may lead to lower bids and, potentially, to lower transaction prices.

2. *Price competition in the operation stage for demand.* An important characteristic of an FA relative to other more commonly used auction mechanisms is that, even when a supplier is added to the FA, it is not guaranteed any fixed number of orders. In fact, there is competition within the FA assortment between all the suppliers offering similar products to capture the demand for the products. *Suppliers are allowed to lower the prices of their products once in the FA*, either temporarily via promotions or permanently by requesting a price change. Naturally, one would expect that by lowering their prices, suppliers may be able to increase their market share. As a result, suppliers may also compete in prices during the operation stage of the FA. In fact, as the operation stage typically lasts for several years, these promotions are not uncommon, as suppliers may experience temporary idiosyncratic changes in their costs (e.g., they may want to sell excess inventory).

Should FA auctions be more competitive? Based on the above discussion, it is unclear whether making the rules of the auction more competitive translates to lower government spending. On the one hand, it is true that by making the auctions more competitive, the auctioneer may be able to decrease the awarded bids and hence the ceiling prices. On the other hand, if fewer bids are awarded, then there will be fewer suppliers in the FA, which in turn may decrease the price competition during the operation of the FA. A stylized model of FAs analyzed in Saban and Weintraub (2021) (see also the follow-up paper Choi et al. (Forthcoming)) suggests that, in fact, making the auctions more competitive should decrease the purchasing prices even in a setting where suppliers are strategic and account for the fact that, even if they do not face competition to enter the FA, they will face competition in the FA assortment to capture demand.

To shed light on this market design question in a real-world setting, we worked in collaboration with ChileCompra in the design of a new Food FA, where we tested the impact of increasing

competition in the auction stage of the FA on public spending. We describe this initiative in the next section.

3 Designing Framework Agreements to Increase Competition

The data describing the 2014 Food FA shows that there is not much competition to enter the FA in the auction stage. This may result in high ceiling prices, which in turn could result in higher transaction prices during the operation of the FA.

To tackle this potential issue, we worked with Chilecompra to test a new design of the framework agreement in the 2017 FA Food. The new design focused on the following changes:

1. *Implementing a product catalogue.* As a first step, we built algorithms that enabled the standardization of the product catalogue based on measurable product attributes. As we explain below, building the catalogue allowed us to implement changes to the auction rules, to measure their impact, and also served as a key tool for the operation of the FA.

2. *New auction rules.* We then defined new auction rules that take advantage of this product standardization in order to measure the impact of generating more competition in the auction stage.

In addition, we streamlined the new FA into a two-step evaluation process, dividing it into a technical evaluation of suppliers, followed by an economic evaluation of the bids. This is the general format recommended by the World Bank to increase transparency in public procurement (US Agency for International Development (2015)).

Next, we describe these and other improvements that were implemented in further detail.

3.1 Implementing the product catalogue

We start this section by explaining the need to create a standardized product catalogue. Then, we explain the algorithms we used to create and maintain such a catalogue.

3.1.1 The need for a standardized product catalogue

In the previous section, we noted that there appeared to be low competition to enter the FA in the auction stage. One potential cause for the low level of competition, may be the lack of

product standardization established by ChileCompra, which allowed suppliers to differentiate their products by self-reporting the attributes. Moreover, suppliers used free unstructured text to describe these attributes, which exacerbated the issue. In practice, this lack of product standardization generated several operational issues.

First, the catalogue of the products needed was not fully specified. This resulted in many products being added, auctioned, and awarded for which the government organizations did not have a real need. For example, 40% of awarded products were not sold during the operation of FA 2014. Moreover, the 50% of products with lower sells represented only 1% of total expenses.

Second, products that in practice were very similar in terms of their attributes were awarded using different auctions, inducing little competition. For example, the product *grapeseed oil Chef 500 ml.* was offered in bundles of 1, 6, and 30 units, generating three different auctions for the same product. This problem was even more severe for those products where suppliers could define all the product's attributes, essentially resulting in those products being awarded through single-bid auctions.

Third, suppliers were allowed to add products during the operation of the FA. Ideally, suppliers would only be allowed to introduce products that were not part of the current assortment. However, in practice, due to the lack of standardized product attributes, this rule was hard to implement. In fact, a visual inspection of the products added during the operation of FA 2014 revealed that many of them corresponded to products that were already part of the assortment, with only small variations to differentiate them.

Therefore, our objective in creating and implementing a product catalogue was threefold:

1. Specify in a precise manner which products need to be purchased.

2. Define the auctions in a more precise and systematic way.

3. Evaluate the need for new products during the operation of the FA.

Moreover, as we later describe in Section 4, having such a product catalogue in place would allow us to implement our experiment and measure its impact.

3.1.2 Creating and implementing a standardized product catalogue

Given the large number of products (see Table 1), it is challenging to construct a detailed product catalogue that would cover the procurement needs of a diverse set of institutions. This is the main reason why ChileCompra opted to allow the suppliers to determine the attributes of the products for which they would submit bids in the first place. However, as argued above, the lack of standardization in these product attributes made it hard to determine whether two bids corresponded to the same product (in the auction stage) or whether a new product was already offered by another supplier in the FA (in the operation stage).

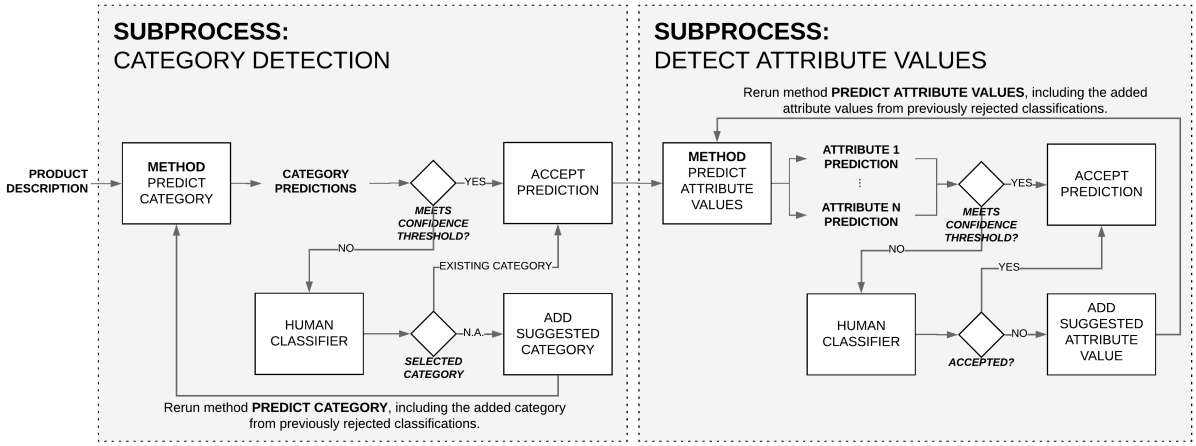
The solution proposed was to (i) define the attributes of each product that needs to be procured in the catalogue in the auction stage, and (ii) limit the entry of new products not contained in this catalogue. It was important for these attributes to be verifiable in order to ensure that the winning bids actually supplied the product that was specified in the catalogue. Moreover, well-defined and verifiable product attributes would enable a rigorous audit to evaluate whether new products added during the operation of the FA were indeed not contained in the original catalogue, thereby limiting the artificial differentiation that suppliers were introducing.

In order to create the catalogue, we used natural language processing (NLP) techniques to process the unstructured text with the product description in order to generate standardized attributes for each SKU. The process we followed is described next.

First, we collected data from external retail catalogues published online and defined a basic product hierarchy based on this catalogue. We used external catalogues because product descriptions and their categorization were more precise and consistent compared to the product descriptions of products sold in the Food FA used by ChileCompra.

Second, we processed the unstructured text of product descriptions to identify relevant attributes that could be used to standardize products. Identified attributes were stored in a dictionary, which was expanded as the algorithm was trained. The algorithm also identified products for which there was low confidence in the attribute identification, separating them out for manual inspection. This approach consisting of hybrid classification using NLP and manual inspection was based on the work of Sun et al. (2014) and adapted to the context of our data (see Guerra López (2019) for details). Figure 1 provides a visual scheme of the algorithm used to identify product attributes and to standardize products in the catalogue.

Figure 1: Summary of the NLP algorithm used to build a structured catalogue based on external data.



Notes: External catalogues (from private retailers) were used for the initial training of the classification algorithm. Unsupervised algorithms were first used to discover an initial set of product categories and attributes. Then, new products with their unstructured text description were processed sequentially using the Category Detection Method to identify the product category. When the prediction did not meet a minimum confidence threshold, the product was revised manually so it could be assigned to a category. If the suggested category did not exist, a new category was added and the method was rerun. Once the category was detected, the Detect Attribute Values method attempted to identify the value of each attribute assigned to that category. When predicted attributes did not meet a minimum confidence threshold, they were processed manually so that they could be assigned an existing or new attribute value. (Source: Translated from Guerra López (2019)).

The set of products with identified attributes was used to build the basis for a structured catalogue. ChileCompra could add additional products not included in this base catalogue that were considered within the procurement needs of some purchase units in the government; these new products were analyzed through the semiautomated process depicted in Figure 1 to identify attributes and generate a standardized product description. As a result, we obtained an expanded catalogue that includes products with specific attributes, which we refer to as *standardized SKUs*. This final list was revised by ChileCompra to define the products that would be included in the new FA auctions.

As a result of this standardization, each SKU in FA 2017 is uniquely defined by a set of verifiable attributes, which include *category*, *type*, *brand*, *format*, and *units*. An example of this catalogue with product attributes is shown in Table 2. *Category* is a generic classification for the SKU, such as water, canned food, or yogurt. *Type* refers to a more detailed product specification, a refinement of the category, which, in turn, further refined by a separate *brand* attribute. *Format* refers to the unit of measurement of a single unit of product, that is, the number of grams (gr.) or milliliters (ml.) in the product container. Finally, products may

Table 2: Example of standardized products for the 2017 Food FA generated using the NLP algorithm.

Product I.D.	Full Product Description	Subcategory	Type	Brand	Format	Units
1477739	sparkling mineral water Next bottle 500 ml. 1 unit	water	sparkling mineral water	Next	500 ml.	1
1445299	water-based canned salmon Robinson Crusoe bundle of 3 units 80 gr.	canned-food	water-based canned salmon	Robinson Crusoe	80 gr.	3
1449854	ravioli Carozzi meat bag of 400 gr. 1 unit	pasta	ravioli	Carozzi	400 gr.	1
1450212	instant coffee Nescafe can 420 gr. 1 unit	tea, coffee	instant coffee	Nescafe	420 gr.	1
1479364	probiotic yogurt Uno multifruit bottle 90 ml. 12 units	yogurt	probiotic yogurt	Uno	90 ml.	12
1443780	juice Watt's nectar peach bottle 1,5 lt. 1 unit	juice	nectar	Watt's	1,5 lt.	1

Notes: Each product description is decomposed into five attributes: subcategory, type, brand, format, and number of units. Actual product descriptions were originally in Spanish and were translated to English in this example to facilitate the interpretation of the attribute classification.

contain a number of *units* per package. For instance, water might be offered in individual bottles or six packs. Providing this level of detail in the units measurement was important for defining bid prices in the FA auctions.

To further understand the value of constructing a structured catalogue, we used the classification algorithm to identify attributes for the products auctioned in the 2014 Food FA. Table 3 shows some examples, where each row represents a different auction. The unstructured text in the product description was analyzed to identify the attributes used in the new catalogue for the 2017 Food FA: category, type, brand, format, and number of units. These examples reveal different ways in which suppliers were able to differentiate products to compete on separate auctions, thereby reducing competition to enter the FA.

Table 3: Examples of auctions in the 2014 FA including the full description of the auctioned product, which was partially entered by the suppliers using unstructured text.

Auction	Full Product description	Subcat	Type	Brand	Format	Units
967288	almidon de maiz Dropa Maizena 250 gr 60 unidades	cereal	corn starch	Dropa Maizena	250 gr.	60
967290	almidon de maiz Dropa Maizena 250 gr. 10 unidades	cereal	corn starch	Dropa Maizena	250 gr.	10
967291	almidon de maiz Dropa Maizena 250 gr. 12 unidades	cereal	corn starch	Dropa Maizena	250 gr.	12
965302	cerdo al vacio trozo pulpa deshuesada de 100 gr. caja de 20 k	meat	pork loin	NA	100 gr.	20 kg.
965303	cerdo al vacio trozo pulpa deshuesada de 150 gr. caja de 20 k	meat	pork loin	NA	150 gr.	20 kg.
965304	cerdo al vacio trozo pulpa deshuesada de 200 gr. caja de 20 k	meat	pork loin	NA	200 gr.	20 kg.
965305	cerdo al vacio trozo pulpa deshuesada de 250 gr. caja de 20 k	meat	pork loin	NA	250 gr.	20 kg.
965306	cerdo al vacio trozo pulpa deshuesada de 300 gr. caja de 20 k	meat	pork loin	NA	300 gr.	20 kg.
965307	cerdo al vacio trozo pulpa deshuesada de 350 gr. caja de 20 k	meat	pork loin	NA	350 gr.	20 kg.
969730	jugo nectar watt´s durazno 1.5 litro botella 6 unidades	juice	nectar	Watt's	1.5 lt.	6
969731	jugo nectar watt´s durazno botella 1.5 l 6 unidades	juice	nectar	Watt's	1.5 lt.	6
969738	jugo nectar watt´s naranja 1.5 ls botella 6 unidades	juice	nectar	Watt's	1.5 lt.	6
969739	jugo nectar Watt´s naranja botella 1.5 l 6 unidades	juice	nectar	Watt's	1.5 lt.	6

Notes: The columns subcat(egory), type, brand, format, and units were not specified in the 2014 Food FA. This information was generated by the NLP algorithms used to construct the catalogue in the 2017 Food FA, in order to compare similarity across the products.

Table 3 groups the listed products based on their similarity in attributes. The first group of products corresponds to corn starch (of the same brand) where the only difference across auctions is the number of units per package. The second group corresponds to pork loin (with no brand), where differentiation occurs in the the individual size of each cut but the total weight of each package is the same (20 kg.). The third group corresponds to juice concentrate (nectar), all of the same brand and format but with differences in the text describing liters and the flavor. These examples illustrate how a small variation in the offered products allowed suppliers to

avoid competition in the auction stage, leading to a large fraction of single-bid auctions. Hence, a structured catalogue based on standardized product attributes could be used to better design the auction stage of the FA and increase competition. We describe this process in the next section.

3.2 Implementing changes to the FA auction rules

The new auction design incorporated several rules that were focused on measuring the effect of enhancing competition. The changes were focused ion the rules of the auction stage, and were grouped into five categories (a summary of these changes is also presented in Table 4):

1. *Separation of technical and economic evaluations.* In the new FA, an initial stage required suppliers to submit information that included their history of sales, infrastructure, and financial statements. This information is used by ChileCompra to perform a technical evaluation of the suppliers. Those suppliers that pass the evaluation, continue on to the second stage—an economic evaluation—which focuses on price competition to enter the FA, implemented through competitive bidding.

2. *Request for products based on the implemented catalogue.* The structured catalogue generated through the NLP techniques described in the previous section defined the actual catalogue of products for the 2017 Food FA auction. This selection of products broadly covered the procurement needs of the government. In contrast to the FAs used previously (including the 2014 Food FA), the new design of the 2017 Food FA prevented suppliers from adding new products so as to enhance competition for those products already in the catalogue.

3. *New auction definitions.* As before, the auction stage was designed as a set of independent auctions. However, each auction was now defined to typically comprise bids for one standardized SKU to be awarded in that auction. Some auctions aggregated more than one standardized SKU based on the similarity of their product attributes.

4. *Inclusion of delivery costs in bids.* The bidding rules were set so that bid prices included the product and delivery costs, with predefined minimum order sizes to reduce transportation costs for the suppliers. As distribution costs can vary significantly across different

Table 4: Main innovations in the design of the new food FA.

Design attribute	Previous FA	New FA
Initial product catalogue	Few products, low standardization. Allows suppliers to define products and add new products during the operation stage.	Many products in the initial catalogue. High standardization based on specific attributes, limited options to add new products during the operation stage.
Winner allocation	Low competition to enter the FA due to low product standardization. Many auctions with single bids.	High standardization increases competition to enter the FA. Experimental design to test different competition levels.

regions of Chile (urban vs. rural areas and proximity to centralized warehouses), the country was divided into a set of geographic regions and auctions were run separately for each region for all products. This allowed suppliers to submit different bid prices for the same product across different regions.

5. *Changes in the rules that decide the winning bids.* We changed the way in which winning bids are decided for each auction. Awarded bidders were set as a percentage of the bids submitted to each auction, with a minimum of 3 awarded bids to ensure an adequate supply. Since this was a major change in the auction rules used historically by ChileCompra, it was implemented through an experimental design assigning different awarding thresholds for different products categories. Recall from the discussion in Section 2.1 that, a priori, it was not clear whether setting a more competitive award rule would result in lower transaction prices. Providing empirical evidence through an experimental design was critical to measure the impact of increasing competition and inform the design of future auctions. The implementation of this experiment is described in detail in Section 4.

It is worth highlighting that *all* auctions were equally affected by changes 1–4 listed above.

In addition to these changes in the auction design, the new Food FA included some additional innovations to facilitate the operation of the FA. A price index mechanism was introduced to reduce suppliers' exposure risk to market fluctuations. Price indexes were adjusted every six months based on external price indexes published by the government statistics bureau (Instituto Nacional de Estadísticas).[2] Second, suppliers would face restrictions to add new products during the operation of the FA. Only a limited number of new products were allowed per year,

[2]See Gur et al. (2017) for a theoretical motivation of this intervention.

which had to be previously authorized by ChileCompra to ensure that these product were not already offered in the existing assortment. To reduce ChileCompra's administrative costs of checking new product additions, the product descriptions provided by the supplier were processed through the NLP algorithms used to catalogue products, which could efficiently validate whether the product was already present in the market.

The auction rules were published in August 2017, followed by one week of consultations by suppliers. Technical evaluations began in November 2017; technically qualified suppliers began submitting bids starting in March 2018. Winners were awarded in June 2018 and the FA operation started in August 2018.

The auction stage included 52,844 auctions, which received on average 5.3 bids per auction, higher than the 3.4 bids per auction that were received in the 2014 Food FA auction. Recall that the 2017 Food FA auctions were run locally on each region whereas in the 2014 Food FA there were many auctions that covered the whole country; therefore, it was notable that the average number of bids per auction increased in the new FA. Moreover, the percentage of single-bid auctions decreased from 51.2% to 28.7%, suggesting that the new auction design was effective in inducing more competition to enter the market. The next section describes the methodology used to measure the impact of this new auction design on prices and procurement expenses.

4 Measuring the Impact of Increasing Competition in the FA Auction

One of the most critical changes in the new Food FA was the idea to intensify competition to enter the FA in the auction stage, by awarding entry to a smaller group of suppliers based on their bid price. Specifically, increasing competition in the auction stage was aimed at reducing prices and thereby expenditures of buyers. To evaluate its actual impact, we implemented the new Food FA through a field experiment focused on measuring the effect of different degrees of competition to enter the FA on government spending. This approach consisted in comparing prices of similar products between the previous and the new Food FA, which required identifying product attributes in order to match products with similar attributes across the two FAs. This

sample of matched products was used to conduct a difference-in-differences regression to measure the impact of inducing more competition to enter the market in the auction stage. To provide a clean estimation of this causal effect, we randomized product categories by applying different awarding thresholds that determined the winners in the auction stage. This section describes the design and implementation of this field experiment, and the econometric models used to measure its impact on bids and prices.

4.1 Experimental design

The implementation of the new Food FA was designed to measure the implications of increasing competition in the auctions used to select the suppliers that enter the FA. This was conducted through an experimental design where each product auctioned in the new Food FA was assigned to one of two possible awarding thresholds:

1. **Noncompetitive (baseline) group**, where the lowest 80% of all received bids in the auction were awarded and included in the FA assortment.

2. **Competitive (treatment) group**, where only the lowest 20% of all received bids in the auction were awarded and included in the FA assortment.

In both cases, the rules imposed a minimum of three bids awarded so as to reduce potential risks in the supply (due to limited availability, delays in the distribution, among etc.). For example, if an auction in the competitive treatment received 10 bids, then the three lowest-bid products would be added to the FA, effectively awarding 30% of the bids (not 20%). All the other changes in the rules of the auction described in Section 3.2 were applied to *all* auctions regardless of whether they were assigned to the baseline or the competitive treatment. Hence, the randomized assignment of this treatment allowed us to isolate the impact of this design variable on the outcomes of interest.

Product categories were assigned randomly to the competitive and noncompetitive treatments. We chose to make the assignment at the product category level so as to avoid having two close-substitute products being assigned to different competitive treatments. Of the 45 proposed assignments, ChileCompra changed one assignment from competitive to noncompetitive and one from noncompetitive to competitive.[3]

[3]Rice was moved from the competitive to the noncompetitive treatment in order to increase the likelihood of

4.2 Matching products across FAs

We sought to compare the bids and prices of products in the old and new FA, as well as of the two competitive treatments to measure the impact of the new design.

To do so, we relied on the product characterization provided by the new FA, described in Section 3.1. More specifically, each product in the old FA was standardized by identifying product attributes (the same as those used in the new FA) from the information contained in the product description, using the NLP algorithms trained with the new FA's catalogue. Because some of the product attributes changed, we generated groups of products that were "close" in the space of attribute values. For example, bottles of 950 ml. and 1 liter were considered to be similar in that attribute. In general, two products were associated with the same standardized SKU if: (i) they shared the same category, product type, brand, and format (e.g., grams); and (ii) the percentage difference between the format values was less than 20%.

While in principle we could apply this methodology to any SKU in the FA catalogue, some categories consist mostly of products that cannot be meaningfully compared through standardized attributes. For example, meat products and fresh produce are usually defined by nonobservable attributes that buyers may obtain through experience or local knowledge. Consequently, we excluded from our analysis fresh bread, vegetables, fruits, seafood, meat products, and others perishable product categories for which the observable attributes were not sufficient to describe the product's quality. Product categories that are reasonably well described by their attributes are in general packaged food products, referred to hereafter as *pantry* products, of which there are 35 categories.

Table 5 shows the results of this attribute identification for pantry products across the two FAs. For FA 2014, products with identified attributes—7,642 out of 8,937 products—accounted for more than 90% of the sales in that FA, which shows the high degree of effectiveness of the product standardization algorithms. Using these attributes to identify similar products, we found that these 7,642 products actually corresponded to 2,703 *unique* products, revealing a significant redundancy of products in the 2014 Food FA. Of the 2,703 unique products, 928 products had a match with a similar product in the 2017 Food FA; these matched products account for about 68% of the sales of pantry in 2014 Food FA, and 74.1% in FA 2017. These

allocating to local suppliers. Coffee and Tea was moved from the noncompetitive to the competitive treatment because it was considered an important category to enhance competition, given the large volume of purchases.

matches defined the sample we used to evaluate the impact of the changes in the FA design, by comparing the prices of the same standardized SKU across the old and new FAs.

	No. of Products (% sales)	
	FA 2014	FA 2017
Total number of *pantry* products	8,937 (100%)	3,789 (100%)
Products with identified attributes	7,642 (92.4%)	3,704 (97.3%)
Unique products based on selected attributes	2,703 (92.4%)	2,027 (97.3%)
Matched products across FAs	928 (68.2%)	928 (74.1%)

Table 5: Product identification summary for all *pantry* products in the 2014 (old) and 2017 (new) FAs. Identified attributes include: subcategory, type, brand, and format. Products that have the same values in all four attributes are considered identical products; otherwise, they are considered different products. The last row counts unique products that were found in both FAs.

4.3 Hypotheses

The main goal of inducing more competition in the auction stage is to generate lower transaction prices and thereby reduce procurement costs. There are alternative mechanisms that can lead to this price reduction. First, awarding bids to a smaller fraction of bidders—the competitive treatment condition in our experimental design—should lower the cutoff bid price, thereby leading to lower *awarded bids*. Moreover, this increased competition in the auction stage might also lead suppliers to bid more aggressively, leading to lower *submitted bids*, which should further reduce the awarded bid prices. Overall, we expect that auctions in the competitive treatment should have lower awarded bids relative to the noncompetitive treatment.

Moreover, recall that the auction stage of the FA sets the ceiling prices that suppliers can charge, but that suppliers can further lower the prices during the operation stage through price promotions. As discussed in the Introduction, while awarding a smaller number of bids may decreases ceiling prices, it is not clear that it also decrease posted and transaction prices. In

fact, with fewer participants in the FA there may be less price competition in the marketplace during the operation stage, resulting in higher posted and transaction prices. This may be more apparent when one considers the long horizon of a FA and that the most cost-efficient suppliers may change over time in a way that cannot be captured by fixing the initial auction ceiling prices. In this situation, more competition may induce cheaper suppliers to decrease mark-ups through promotions throughout the FA in order to price out competitors.

Based on this discussion, we postulate the following hypotheses:

Hypothesis 1. *Lower ceiling prices*: Awarding a smaller number of bidders leads to lower awarded bids.

Hypothesis 2. *Prices for the competitive treatment are lower*: Awarding a smaller number of bidders, leads to lower posted and transaction prices.

Hypothesis 2 alt. *Prices for the competitive treatment are not lower*: Posted and transaction prices are not lower in the competitive treatment. Price competition in the FA operation may compensate for less competition in the auction.

We test these hypotheses in the next section.

4.4 Econometric model

We conducted a difference-in-differences regression to measure the effect of increasing competition in the auction stage, distinguishing the alternative mechanisms discussed in the hypotheses formulation. Specifically, we sought to estimate the causal effect of the competitive treatment condition on submitted and awarded bid prices in the auction stage and posted and transaction prices during the operation stage.

Auction stage: Analysis of bids. First, we focus on studying the effect of the new FA design and the competitive treatment on the auction bids. As we explained above, bid prices correspond to the ceiling prices that the suppliers can charge during the operation of the FA.

Let the index $t \in \{\text{Old}, \text{New}\}$ represent the old (2014) and the new (2017) Food FAs, respectively. Let b_{ijrt} represent the bid offered in FA t by supplier j to provide product i in region r,

including shipping.[4] Define B_{irt}^{med} as the *median* bid across all supplier bids submitted in the FA for that product-region-FA combination. In calculating this median bid price, we considered alternative methods to for discarding extreme prices that could be generated by bidding mistakes or unrealistically aggressive bidding. For the main results we used a modified Tukey rule. To make the bids for the 2014 and 2017 Food FAs comparable, all prices for 2014 were adjusted using the CPI dood price index. In addition, we normalized bid prices so that they represented bid prices per unit. See Appendix A for further details on all of this data preprocessing.

The following difference-in-differences specification is used to estimate the effect of the competitive treatment condition —denoted by the indicator variable $Comp_i$— on bid prices:

$$\log(B_{irt}^{med}) = \delta_r + \gamma_i + \alpha New_t + \beta New_t \times Comp_i + \varepsilon_{irt} \ , \tag{1}$$

where New_t is an indicator equal to one for the new FA and δ_r is a region fixed effect. The sample is comprised of all products i that were matched across the old and new FAs based on similar attributes, and the regression includes a product fixed effect γ_i. The coefficient α measures the average differences in bid prices between the old and new FAs for the noncompetitive baseline group. The key parameter of interest is β, the coefficient capturing the incremental change in bid prices for the auctions in the competitive treatment condition in the new FA.

Recall that B_{irt}^{med}, which is the median across all *submitted* bids, measures whether the competitive treatment affects the bidding behavior of the suppliers. Any changes in the submitted bids should carry over into the winning bids. Furthermore, the competitive treatment condition also lowers the cutoff point of the winning bids, which can lead to lower awarded bids even when submitted bids remain constant. Hence, we also estimated regression 1 using the median among *winning bids* as the dependent variable. Taken together, these models allow us to test Hypothesis 1, by measuring the effect of less competition in the auction stage on the winning bids and identifying what part of this reduction is due to changes in bidding behavior.

Operation stage: Analysis of prices. While the bids constitute a ceiling on the price that suppliers can charge during the operation of the FA, the ultimate goal is to reduce posted and transaction prices, which translate directly into procurement savings. To conduct this analysis,

[4]Recall that some auctions were run at the national level in FA 2014. However, suppliers must specify a shipping rate for all regions they seek to supply, which is used to calculate the bid price for each region.

we assembled a weekly panel dataset including prices posted in the online platform and the actual transacted prices from purchasing orders. Let p_{ijrtw} denote the average posted prices by supplier j for product i in region r during calendar week w in FA t. As with the bids, we calculated the median posted price for a product across all suppliers during that week, denoted by P^{med}_{irtw}. That is, P^{med}_{irtw} denotes the median price at which product i could be purchased in region r during calendar week w under the operation of FA t in the period listed above.

We estimated the panel regression:

$$\log(P^{med}_{irtw}) = \delta_r + \gamma_i + \tau_{wc(i)} + \alpha New_t + \beta New_t \times Comp_i + \varepsilon_{irt} \ , \tag{2}$$

where δ_r and γ_i are region and product fixed effects respectively, $\tau_{wc(i)}$ is a product category-specific calendar week dummy variable to capture potential seasonality ($c(i)$ indicates the category of product i).

To analyze actual expenditures we used each product as the cross-sectional unit and calculated the median price across all units that were bought during a given week. To calculate the median, we considered each unit bought as one observation. For example, an order with n units for the same product is considered as n different observations in the median calculation. In this way, the median considers the quantity of products sold at each price level. Notice that this panel is unbalanced because some products were not sold every week. We estimated a regression similar to (2) using the logarithm of the median transaction price as the dependent variable.

Both panel regressions were estimated with data from 13 months prior to the operation of the new FA (August 2017 to August 2018) and the first 13 months of operation (September 2018 to September 2019).[5] Prices are adjusted for inflation according to the monthly CPI index for Food and Non-alcoholic Beverages. In addition, we eliminated outliers by applying the same procedure used for bid prices.

4.5 Estimation results

Table 6 shows the main estimation results showing the impact of the change in the level of competition in the auction stage on bids and prices. Columns (1) and (2) show the effect on submitted and awarded bids (regression 1); columns (3) and (4) show the effect on posted

[5] Using a window of 13 months ensured that all 52 calendar weeks are considered.

and transaction prices (regression 2). Appendix B contains robustness analyses and alternative regression specifications. The results obtained from these alternative regression specifications are qualitatively similar to the ones obtained with the main specifications described in the previous subsection.

	Bids		*Prices*	
	Submitted	Awarded	Posted	Transaction
	(1)	(2)	(3)	(4)
New	-0.141^{***}	-0.055^{***}	-0.025^{***}	0.004^{***}
	(0.006)	(0.007)	(0.001)	(0.002)
New×Comp	-0.002	-0.081^{***}	-0.092^{***}	-0.082^{***}
	(0.010)	(0.013)	(0.001)	(0.002)
Observations	12,349	11,382	973,195	180,421
R^2	0.923	0.892	0.961	0.974
Adjusted R^2	0.920	0.887	0.961	0.974
Note:		*p<0.1; **p<0.05; ***p<0.01		

Table 6: Main estimation results measuring the effect of the new FA design and the level of competition in the auction stage on bids and prices. All the specifications include product and region fixed effects. Regressions with posted and transaction prices include product category-specific calendar week dummies to control for seasonality. Robust standard errors for each estimated coefficient are reported in parentheses for all the models.

The estimation results in columns (1) and (2) suggest that the baseline groups in FA 2017—which were awarded 80% of the bids—exhibited lower bids relative to FA 2014: the *New* coefficient estimates that submitted bids were reduced by 14.1% and awarded bids were reduced by 5.5%. Although part of this effect may be attributed to changes in the auction design, it is not possible to disentangle this effect from other external factors, such as changes in food prices or other fluctuations in the open market. Hence, we cannot interpret α as a causal effect of the new design on prices.

The effect of the competitive treatment —which was induced through a randomized experimental design— can be interpreted as a causal effect of inducing more competition in the auction stage of the FA. The interaction term $New \times Comp$ shows that auctions that had lower thresholds for assigning winners led to lower awarded bids —on average 8.1% lower (column (2))— providing support for Hypothesis 1. However, there appears to be no effect on the bids submitted by the bidders: in column (1) the coefficient of $New \times Comp$ is small and not statistically significant. This suggests that the reduction in the price of the awarded bids was not induced by changes in bidding behavior across the different auction rules (i.e., competitive or noncompetitive) but rather due to the lower awarding threshold applied in the competitive treatment group.

In terms of prices, the results suggest that posted prices were lower across all products in the new FA design relative to FA 2014, showing a 2.5% reduction in the price of the products in the noncompetitive treatment and an additional 9.2% reduction of those in the competitive treatment. Overall, the reduction in posted prices of products in the noncompetitive treatment is substantially smaller than the reduction in the ceiling prices (determined by the awarded bids); that is, the pass-through from ceiling to posted price is not one-to-one. For the noncompetitive products, posted prices decreased by 2.5% relative to FA 2014, about half the 5.5% decrease observed for the awarded bids.

Products in the competitive treatment had posted prices that were 11.7% lower (0.025 + 0.092) relative to the FA 2014, compared to a 13.6% (0.055 + 0.081) drop in the awarded bid prices. One potential explanation is that in the old FA, even though ceiling prices were higher, this lack of competition in the auction stage was partially compensated for by more competition in the operation stage of the FA, as showcased by the fact that suppliers were frequently running price promotions. Overall, the new FA design combined with a more competitive auction stage was effective in reducing posted prices, and the effect is both statistically and economically significant.

More importantly, the estimates on transaction prices (column (4) in Table 6) reveal that the reduction in posted prices for the competitive treatment also translates into a significant reduction in transaction prices, in the order of 8%, relative to FA 2014. However, the reduction on transaction prices for the noncompetitive group is negligible. Taken together, these results

provide support for Hypothesis 2: inducing more competition in the auction stage was effective in reducing posted and transaction prices, thereby lowering procurement spending for the government. The savings in the government procurement cost amounted to US$3.6 million yearly if we consider only the treatment group and to US$11 million yearly if we extrapolate to the entire Food FA.[6] It is therefore important to introduce a competitive entry stage to achieve reductions in procurement costs when using FA purchasing mechanisms.

5 Generalizing the new FA Design and Transferability

5.1 Transferability to other Framework Agreements

Through the pilot study just described, we identified important steps that were fundamental to induce competition upon entry and for the efficient operation of the FA: product standardization and defining the catalogue of products to be procured. These steps have been adopted as part of the standard operating procedure used by ChileCompra in the design and implementation of FAs, which were formally included in the new regulation on public procurement (ChileCompra (2021a)). In addition, the experimental design used in the Food FA provided evidence in support of inducing more competition in the auction stage of the FAs. Based on this, ChileCompra adopted a new strategy for the design of all the FAs that were implemented after the Food FA. With this motivation, we collaborated with ChileCompra during August 2019–February 2020 with the objective of transferring the tools and know-how developed in the Food FA to the design of several new FAs: for office supplies, computers, office furniture, household cleaning, and hardware. This transferability was supported by a government grant, aimed at facilitating the adoption of the best practices that were identified in the Food FA pilot study.[7]

Competitive auction rules. The new FAs had a two-stage design: a first, selection stage based on technical criteria followed by a second, competitive stage conducted through an auction based on prices.

The design of the auction stage included a structured product catalogue, developed by Chile-

[6]The extrapolation more than doubles the savings because, as we explained before, there were many non-pantry product categories not considered in the experiment.

[7]Fondef Project ID16I10122, "Diseño De Una Plataforma Para La Implementación De Convenios Marco Competitivos," directed by Marcelo Olivares.

Compra, specifying detailed product attributes that uniquely identified the products to be auctioned. Some of these catalogues were constructed based on catalogues from the open market and processed using NLP tools (which included some extensions of those used for the Food FA).

Moreover, the awarding rules were uniformly changed to induce more competition relative to previous FAs. In some of the new FAs, awarding rules were based on a fixed number of winners per auction. For example, for the Computers FA (2020), products were grouped into quality categories where the X lowest bids in each quality category were awarded (the number of awardees X varied between 10 and 20 across the Desktop and Laptop Computer categories, based on the potential number of suppliers). Office Furniture and Office Supplies FAs used a similar awarding criteria. For other FAs, awarding rules were defined as a percentage of the submitted bids. For example, for the Household Cleaning FA, the lowest 30% of the received bids were awarded and for Industrial Cleaning Products FA, where products are bought in larger volumes, the cutoff was set at 10%. The Hardware Products FA used similar awarding criteria.

Product standardization and product catalogue. As discussed in Section 3.2, to implement a competitive yet meaningful auction, it is useful to have a standardization of the products to eliminate the differentiation introduced by suppliers that would reduce the competition. The NLP methodology developed in the pilot study was further developed to improve its accuracy and packaged into a stand-alone application programming interface, funded by government grants aimed at technological transfers to the public sector. This interface was transferred to ChileCompra and made available as open source code that could be used in other public sector and non-profit applications.[8]

After the implementation of the Food FA, ChileCompra changed its procurement strategy and opted to focus FAs on categories for which it was possible to describe products through observable attributes. Categories for which it was not possible to describe products through a structured catalogue (e.g., catering services, training programs) were moved from FAs to other procurement mechanisms such as auctions or direct purchases. This lead to fewer products and services being available through FAs, but these FAs exploited product standardization to generate a more competitive awarding process and thereby achieve lower prices. As a consequence of this new strategy, the share of purchases conducted through FAs was reduced from 23% in

[8]The source code is available at the repository `https://github.com/rguerral/catalogador_standalone`.

2018 to 11.6% in 2021 (because there were fewer categories), but product categories operating with the new FA design were actively used by buyers.

Overall, the entire process of defining a structured product catalogue (using algorithmic NLP methods), made the design and evaluation process more efficient, thereby reducing administrative costs. The methodology for constructing standardized catalogues and introducing higher levels of competition at the FA auction stage was implemented in subsequent FAs, which in 2022 amounted to US$800 million worth of purchases (including the Food FA). If we were to extrapolate the treatment effect on transaction prices to these FAs, the total savings would amount to $8\% \times 800 =$ US$64 million per year.

5.2 Impact beyond the design of FAs

The adoption of catalogue standardization in the design of FAs aimed at generating competition in the auction also enabled innovations during the evaluation and operation stage of the FA.

First, reducing the amount of unstructured text product description enabled a more automated process for submitting and evaluating the bids. A web server was developed to receive bids from suppliers in a standard format, providing information on product attributes, prices, and other bid variables that could be used in more complex auction formats. These improvements significantly reduced the work required by suppliers to submit bids and minimize errors in the submission. This bid validation process was implemented as a stand-alone application available in open source to the public.[9]

Second, product standardization also led to more structured transaction data, which made it amenable to the data analysis conducted during the operation of the FA. In particular, product attributes could be used to identify groups of substitute products and compare their prices. This led to the development of a monitoring system to measure *overspending* for government purchase units, by defining metrics that captured the potential savings that could be achieved by switching to substitute products with lower prices. Savings due to this monitoring system were estimated in the order of 1.5% of expenditures, equivalent to US$12 million if we use the overall 2022 FA expenditures as a baseline (Celhay et al. (2022)).

[9]The source code for the bid validation interface is available in the following repository `https://github.com/rguerral/site_validador`.

6 Conclusions

In this work, we collaborated with the Chilean government procurement agency, ChileCompra, to redesign the 2017 Food FA. The new design consisted of two main changes: (1) the implementation of a standardized catalogue of products to be purchased in the auction, and (2) an experimental auction design, in which different product categories faced different levels of competition to become part of the FA, allowing us to measure the impact of increasing competition in the FA auction. We showed that products that were assigned to the competitive treatment (the 20% competition threshold treatment at the auction stage) had 8.2% lower transacted median prices. This reduction in prices amounts to US$3.6 millions in yearly savings related to the purchase of those products in the treatment group, and to US$11 million if we were to extrapolate to the entire Food FA. The design was implemented in other FAs, resulting in total savings of around US$64 million per year for the Chilean government.

Overall, our work shows the value of enhancing competition in FA auctions. Moreover, we provided important guidelines on how to implement FAs in practice, which have since been adopted by ChileCompra in the design of other large FAs used to purchase products and services worth US$800 million per year. Finally, as a by-product of our work, we generated decision-support tools that are now actively used by ChileCompra to improve their procurement processes beyond FAs.

References

Albano, G.L., Nicholas, C., 2016. The Law and Economics of Framework Agreements. Cambridge University Press.

Celhay, P., Gertler, P., Olivares, M., Undurraga, R., 2022. Diseño y evaluacion de "nudges" para incentivar el arhorro en las compras publicas.

ChileCompra, 2021a. Aprueba nuevas politicas y condiciones de uso del sistema de información de compras públicas y registro de proveedores. Resolución Exenta N° 585-B .

ChileCompra, 2021b. Cuenta Pública ChileCompra: Gestión 2021 y avances

2022 URL: https://www.chilecompra.cl/wp-content/uploads/2022/05/Cuenta-Publica-ChileCompra-2021.pdf.

ChileCompra, 2022. Datos Abiertos de las Compras Públicas en Chile. https://datos-abiertos.chilecompra.cl/. [Online; accessed 2-Dec-2022].

Choi, J.o., Saban, D., Weintraub, G.Y., Forthcoming. The design of optimal pay-as-bid procurement mechanisms. Manufacturing and Service Operations Management Available at SSRN: https://ssrn.com/abstract=3785023 or http://dx.doi.org/10.2139/ssrn.3785023.

Demsetz, H., 1968. Why regulate utilities? The Journal of Law Economics 11, 55–65. URL: http://www.jstor.org/stable/724970.

Dimitri, N., Piga, G., Spagnolo, G., 2006. Handbook of procurement. Cambridge University Press.

Guerra López, R.A., 2019. Diseño de un sistema de catalogación de productos de e-commerce utilizando procesamiento de lenguaje natural (nlp) y machine learning. https://repositorio.uchile.cl/handle/2250/170306.

Gur, Y., Lu, L., Weintraub, G.Y., 2017. Framework agreements in procurement: An auction model and design recommendations. Manufacturing & Service Operations Management 19, 586–603.

Klemperer, P., 2004. Auctions: theory and practice. Princeton University Press.

Levy, E.D., 2017. Rediseño de bases de licitación para convenios marco competitivos. caso de estudio basado en convenio marco de alimentos. https://repositorio.uchile.cl/handle/2250/147385.

Myerson, R.B., 1981. Optimal auction design. Mathematics of operations research 6, 58–73.

OECD, 2019a. Public Procurement in Chile: Policy Options for Efficient and Inclusive Framework Agreements. OECD Publishing.

OECD, 2019b. Reforming Public Procurement: Progress in Implementing the 2015 OECD Recommendation. OECD Publishing.

Saban, D., Weintraub, G.Y., 2021. Procurement mechanisms for assortments of differentiated products. Operations Research 69, 795–820.

Sun, C., Rampalli, N., Yang, F., Doan, A., 2014. Chimera: Large-scale classification using machine learning, rules, and crowdsourcing. Proceedings of the VLDB Endowment 7, 1529–1540.

US Agency for International Development, 2015. Systems for improved access to pharmaceuticals and services (siaps) program. Technical Report .

Vulkan, N., Roth, A.E., Neeman, Z., 2013. The handbook of market design. Oxford University Press.

A Data pre-processing

In order to compare prices between different FA and to rule out unreasonable values, we establish a procedure to process bids, transaction, and posted prices.

Bids. In the case of bids, we identify outliers within each distribution of bids defined by a standardized product i and FA t. In the case of FA 2014, we use prices before adding the respective shipping rate. In the case of FA 2017, we normalize prices to the Metropolitan Region (where the capital city Santiago is) for outlier identification. Additionally, we use the CPI index to adjust prices from the 2014 FA for inflation. In Chile there are available *CPI indexes that are product specific* that we use to bring to present value (to the year 2017) the prices offered in 2014, so that each product category is corrected differently according to the corresponding specific CPI.

We identify four types of outliers, using four different approaches. For each FA t and standardized product i, we define a set of similar products $C_{i,t}$. (A standardized product was defined in Section 4.2.) First, we compute the mean price $\hat{p}_{i,t}$ of its distribution. For each seller s, product j, region r and FA t, with bid $p_{s,j,r,t}$, we compute the ratio $R_{s,j,r,t} = \frac{p_{s,j,r,t}}{\hat{p}_{i,t}}$, were i is the respective standardized product associated to j. We define outliers as follows:

Type 1: $p_{s,j,r,t}$ is an outlier if $R_{s,j,r,t} < 1/2$ or $R_{s,j,r,t} > 2$.

Type 2: $p_{s,j,r,t}$ is an outlier if $R_{s,j,r,t} < 1/3$ or $R_{s,j,r,t} > 3$.

Type 3: $p_{s,j,r,t}$ is identified as an outlier following the Tukey's rule, i.e. prices above and below the first and third quartile of the distribution, whose distance to the nearest quartile is larger than 1.5 times the inter-quartile range.

Type 4: If 20% of the products in the set of similar products $C_{i,t}$ meet the Type 3 outlier criteria, all products in the set are considered outliers. That is, we exclude all the standardize products that exhibit a large proportion of extreme values.

First, to compute normalized bids or prices, we divide prices by the relevant number of units in the bundle. In addition, as a robustness we estimate regression (1) correcting prices by volume to take into account potential scale effects such as volume discounts. See Tables 8 and 10. Thus, prices of products specifying larger quantities of items would be comparable to product with a single or few items. Specifically, in the 2014 FA there were products sold in large bundles. To account for potential volume discouonts, we estimate the regression:

$$\log(B_{irt}^{med}) = \delta_r + \gamma_i + \beta \log(units_{irt}) + \varepsilon_{irt} \ , \tag{3}$$

where $units_{irt}$ represents the amount of items per bundle for product i. Then, we compute the adjusted bid price $\hat{B}_{irt}^{med} = \log B_{irt}^{med} - \hat{\beta} \log(units_{irt})$ and use this as dependent variable in regression (1). The results of this estimation are reported in Table 8 for submitted bids and in Table 10 for awarded bids, which are similar to those reported in our main results.

Transaction and posted prices. Regarding transaction and posted prices, we correct by inflation according to the monthly CPI index for Food and Non-Alcoholic Beverages, using as reference month December 2013. Similarly to bids, we identify three types of outliers. For each transaction/post k and FA t of a standardized product i. Let $R_{k,t} = \frac{p_{k,t}}{\hat{p}_{i,t}}$, where $p_{k,t}$ is the unit price of transaction k, and $\hat{p}_{i,t}$ is the average price in $C_{i,t}$. Then, we define outliers as follows:

Type 1: If the proportion of transactions with $R_{k,t} < 1/2$ or $R_{k,t} > 2$ is larger than 20%.

Type 2: If the proportion of transactions with $R_{k,t} < 1/3$ or $R_{k,t} > 3$ is larger than 20%.

Type 3: If the proportion of outliers following the Tukey's rule in $C_{i,t}$ is greater than 20%.

As a robustness check, Appendix B we show the results under the different outlier elimination methods.

B Robustness analysis and alternative regression specifications

B.1 Regressions for submitted bids

Table 7: Estimation for submitted bids under different outlier elimination rules

	Type 1	Type 2	Type 3	Type 4
	(1)	(2)	(3)	(4)
New	-0.148***	-0.151***	-0.146***	-0.141***
	(0.005)	(0.005)	(0.005)	(0.006)
New × Comp	-0.029***	-0.034***	0.009	0.002
	(0.007)	(0.008)	(0.009)	(0.010)
Observations	13,111	13,195	13,163	12,349
R^2	0.94644	0.93650	0.92588	0.92289
Adjusted R^2	0.94433	0.93401	0.92297	0.91965

Note: This table shows estimates for submitted bids according to (1). Each Model is defined by the type of method used for outlier detection according to section A. Standardized product and region fixed effect are included in all Models. We compute *robust* standard errors.

Table 8: Estimation for submitted bids: Adjusting by volume

	Type 1	Type 2	Type 3	Type 4
	(1)	(2)	(3)	(4)
New	-0.125***	-0.128***	-0.176***	-0.177***
	(0.005)	(0.005)	(0.005)	(0.005)
New × Comp	-0.023***	-0.028***	0.001	-0.008
	(0.007)	(0.008)	(0.009)	(0.009)
Observations	13,111	13,195	13,163	12,349
R^2	0.94616	0.93629	0.92623	0.92344
Adjusted R^2	0.94403	0.93379	0.92333	0.92023

Note: This table shows estimates for submitted bids when prices are adjusted by volume following the specification in (3). Each model is defined by the type of method used for outlier detection according defined in Appendix A. Standardized product and region fixed effect are included in all models. We compute robust standard errors.

B.2 Regressions for awarded bids

Table 9: Estimation for awarded bids under different outlier elimination rules

	Type 1	Type 2	Type 3	Type 4
	(1)	(2)	(3)	(4)
New	-0.072***	-0.073***	-0.064***	-0.055***
	(0.005)	(0.005)	(0.006)	(0.007)
New × Comp	-0.162***	-0.173***	-0.074***	-0.081***
	(0.008)	(0.009)	(0.012)	(0.013)
Observations	11,760	11,933	11,901	11,382
R^2	0.94555	0.93189	0.89614	0.89202
Adjusted R^2	0.94314	0.92893	0.89160	0.88708

Note: This table shows estimates for awarded bids. Each model is defined by the type of method used for outlier detection according to Appendix A. Standardized product and region fixed effect are included in all models. We compute *robust* standard errors.

Table 10: Estimation for awarded bids: Adjusting by volume

	Type 1	Type 2	Type 3	Type 4
	(1)	(2)	(3)	(4)
New	-0.033***	-0.035***	-0.105***	-0.105***
	(0.005)	(0.005)	(0.006)	(0.007)
New × Comp	-0.156***	-0.167***	-0.082***	-0.092***
	(0.008)	(0.009)	(0.011)	(0.012)
Observations	11,760	11,933	11,901	11,382
R^2	0.94506	0.93146	0.89717	0.89362
Adjusted R^2	0.94263	0.92848	0.89268	0.88875

Note: This table shows estimates for awarded bids when prices are adjusted by volume following (3). Each Model is defined by the type of method used for outlier detection according to Appendix A. Standardized product and region fixed effect are included in all models. We compute *robust* standard errors.

B.3 Regressions for posted prices

Table 11: Estimation for posted prices under different outlier elimination rules

	Type 1	Type 2	Type 3
	(1)	(2)	(3)
New	-0.029***	-0.033***	-0.025***
	(0.0007)	(0.0008)	(0.0008)
New × Comp	-0.076***	-0.075***	-0.092***
	(0.0009)	(0.0009)	(0.001)
Observations	1,012,598	1,018,778	973,195
R^2	0.97025	0.96865	0.96132
Adjusted R^2	0.97018	0.96857	0.96121

Note: This table shows estimates for posted prices as in (2). Each Model is defined by the type of method used for outlier detection according to Appendix A. Standardized product, region, and week category-specific fixed effect are included in all models. We compute *robust* standard errors.

B.4 Regressions for transaction prices

Table 12: Estimation for transaction prices

	Type 1	Type 2	Type 3
	(1)	(2)	(3)
New	-0.007***	-0.006***	0.004**
	(0.002)	(0.002)	(0.002)
New × Comp	-0.068***	-0.071***	-0.082***
	(0.002)	(0.002)	(0.002)
Observations	193,702	194,312	180,421
R^2	0.97441	0.97328	0.97395
Adjusted R^2	0.97409	0.97294	0.97360

Note: This table shows estimates for transaction prices following the model in (2). Each model is defined by the type of method used for outlier detection according to Appendix A. Standardized product, region and week category-specific fixed effect are included in all models. We compute *robust* standard errors.