AI Enabled Performance Monitoring and Accountability in Public Procurement

PABLO A. CELHAY[1], PAUL J. GERTLER[2,3], MARCELO OLIVARES[4], RAIMUNDO UNDURRAGA[4], and PIERO ZANOCCO[5]

[1]PUC Chile
[2]UC Berkeley
[3]NBER
[4]Universidad de Chile
[5]UC Davis

September 5, 2026

ABSTRACT. Using a new e-procurement platform in Chile and natural language processing AI to identify the set of close substitutes for each product, we construct quality-adjusted measures of efficiency – the amount paid relative the lowest price available – at virtually zero marginal cost. We then build individual buyer performance reports based on their overspending on all purchases that are generated automatically monthly. We develop a principal-agent model that predicts efficiency gains and the mechanisms behind those gains from performance monitoring reports when individual buyer performance is observable to managers so that managers are able to hold buyers accountable. Performance information can also improve buyer performance even when it is not visible to managers if buyers are intrinsically motivated. We test these predictions using a field experiment that randomizes monthly performance reports across 184 procurement units with over 3500 buyers. Consistent with the model, the performance reports reduce overspending by 15%, equivalent to 0.1% of GDP, when individual performance is disclosed to managers, but have no effect when the information is not visible to managers. Consistent with the model, the biggest gains are for products where the lowest price supplier fluctuates over time, and for buyers who are the poorest performers prior to the intervention and are the most responsive to manger persuasion. Follow-up surveys confirm the mechanisms: public reports trigger manager oversight actions that shift buyer willingness to prioritize price.

JEL CODES. D23, D73, L20, H50, H57, M12, M14, D83

KEYWORDS: Public Procurement, Efficiency, Organizational Culture.

CONTACT. . **Acknowledgments**: We are greatful to Augustin Bergeron, Ernesto dal Bo, Claudio Ferraz, Asim Khwaja, Jonas Hjort, Ricardo Perez-Truglia, Dina Pomeranz, Damian Vergara, and Guo Xu for helpful comments and to Ignacio Torres for outstanding research assistance. We are grateful for support from the Experimental Policy Initiative of the Chilean Public Budget Office, including Antonio Garrido, Ryan Cooper, Slaven Razmilic, Nicolás Blanco and Gonzalo Gaete. We further thank Tania Perich, Alejandro Burr, and Dora Ruiz at *ChileCompra* for their support in accessing and processing the procurement data. We acknowledge funding from Fondo Evaluación de Impacto DIPRES 2020, FONDEF Project ID19I10303, MIPP Institute and Berkeley LIFT. Undurraga further acknowledges funding from the MIGRA Millenium Nucleus under grant ANID-MILENIO-NCS2025_69 as well as by Fondecyt Regular ID 32164 Folio 1260544. Undurraga and Olivares also acknowledge funding from Fondecyt Regular Project No. 1230416. The authors declare that they have no material or financial interests in the results of this paper.

1. Introduction

Governments worldwide spend approximately US$13 trillion annually on public procurement — around 12% of global GDP — yet lose an estimated quarter of this due to inefficiencies (World Bank, 2022; Bosio et al., 2021; OECD, 2023). A key source of inefficiency is that procurement officers often face limited accountability and few incentives to secure competitive prices. As a result, ensuring that public buyers get value for money remains a persistent challenge (Bandiera et al., 2021b; Best et al., 2023; Bosio et al., 2021). The difficulty arises in part because conventional monitoring tools— audits, oversight boards, and compliance reviews — are costly, can be counterproductive, and can increase bureaucratic frictions that can exacerbate passive waste (Kelman, 1990, 2005; OECD, 2013; Gerardino et al., 2024). This creates potential for large gains from identifying cost-effective ways to monitor procurement processes and improve price discipline (Bandiera et al., 2009; Best et al., 2023).

Two recent technological innovations offer a low-cost performance monitoring opportunity. The first is the development of *e-procurement* platforms built around Framework Agreements, which centralize purchasing through competitive digital marketplaces and generate rich transaction-level records on prices, products, suppliers, and individual buyers — precisely the inputs needed for performance monitoring at low cost (Lewis-Faupel et al., 2016; Bandiera et al., 2021a; Olivares et al., 2025). The second is the application of natural language processing AI to create product catalogs, which enables the automated construction of quality-adjusted groups of close substitutes at scale — a task previously requiring costly manual classification that inhibited systematic price comparisons across purchases (Khandelwal et al., 2023; Sun et al., 2014). Together, these innovations make it feasible to create a monitoring technology that automatically generates individual buyer performance information at virtually zero marginal costs, thereby avoiding the cost and bureaucratic burden that have historically limited the effectiveness of procurement oversight.

In collaboration with Chile's Public Budget Office (DIPRES) and Public Procurement Office (*ChileCompra*), we leverage these innovations to study performance monitoring in public procurement. Chile was an early adopter (2007-2008) of *e-procurement* system based on Framework Agreements that operates through an online platform in which pre-qualified suppliers compete to sell standardized products to government buyers. We use Natural Language Processing algorithms to group products into quality-adjusted reference groups of close substitutes, and measure *overspending* by comparing the price they paid to the lowest available price in the relevant reference group. We then aggregate transaction overspending to the buyer level and produce automated monthly performance reports for each individual buyer. Since the underlying data are already embedded in the platform, this monitoring system operates at virtually zero marginal cost.

We first study this monitoring technology through the lens of a principal-agent model in which a manager seeks to induce efficient search effort from buyers who are extrinsically motivated. The manager cannot observe buyer effort directly, but can use individual performance information to

better identify under-performing buyers and target her oversight actions accordingly. The model predicts that performance information improves efficiency by enabling managers to identify poorly performing buyers and hold them accountable. Managers are better able to use oversight actions to persuade buyers to put in more effort to find lower prices, and buyers have a greater incentive to respond. The model also predicts that there are larger gains from buyers that are more responsive to manager oversight and are among the poorest performers where there is the most to gain prior to the introduction of the monitoring system, and for products where there is more price dispersion.

We also consider intrinsically motivated buyers who get value from doing a better job, i.e. finding lower prices (Bénabou and Tirole, 2003). This is consistent with the notion that workers who enter the government are motivated by public service. In this case, performance information lets the buyer know how much they could gain with more effort and motivates poor performers to do better. Here, managers do not need to hold buyers accountable and their oversight actions have no effect on buyer performance. In this case, individual performance monitoring information does not need to be made available to managers for monitoring to be effective.

We tested these predictions using a cluster randomized field experiment with 184 procurement units and more than 3,500 buyers. Units were randomly assigned to one of three arms: (i) a control group receiving no reports; (ii) a *Public* arm in which both managers and buyers received individual performance reports and buyers were informed that their manager could see their data; and (iii) a *Private* arm in which only buyers received individual reports and were explicitly informed that their manager could not see their individual performance. In the *Public* arm, managers receive the individual buyer performance information necessary to hold buyers accountable and can use that information to take action to persuade them to put in more effort to find lower prices. In the *Private* arm, managers do not get the individual information, and performance improves only if the information motivates buyers to do so at their own initiative. As such, the design allows us to assess the importance of accountability in monitoring systems and to identify whether extrinsic versus intrinsic motivation drives the results.

The results are consistent with the predictions of the model for extrinsically motivated buyers. We find that Public reports reduced overspending by 1.6 percentage points, a 15% reduction relative to the control mean and equivalent to US$4.5 million over five months for the products under study. Extrapolating the treatment effect to all Chilean public procurement for a year yields approximately US$0.15 billion in savings, or 0.1% of GDP. Consistent with the model, we find larger effects of information on efficiency from the poorest performers where there is the most to gain and the buyers who are most responsive to manager persuasion oversight and persuasive actions. In addition, the effect of information has the greater impact in product categories where price rankings shift frequently over time, requiring buyers to continuously re-evaluate their options.

Private reports had no effect on overspending and rules out intrinsic motivation. Buyers in the private arm received the same individual performance information, yet did not reduce overspending, confirming that individual accountability and not information alone drives the efficiency gains. This implies that visibility of buyer performance information to the manager is necessary for

monitoring systems to be effective.

Our paper contributes to two strands of literature. First, we contribute to the literature on *inefficiency in public procurement* (Best et al., 2023; Dimitri et al., 2006; Bandiera et al., 2021b). A central debate in this literature concerns the optimal balance between rules and discretion in procurement (Kelman, 1990; OECD, 2009). Several recent studies document the benefits of discretion in procurement outcomes (Coviello et al., 2018; Carril, 2019; Decarolis et al., 2020; Bandiera et al., 2021a; Szucs, 2023), while others show that the capacity of bureaucracies to monitor and enforce contracts limits the gains from relaxing procurement rules (Bosio et al., 2021; Spagnolo, 2012; Palguta and Pertold, 2017; Liscow et al., 2023). Our paper complements this work by showing that performance monitoring can substantially reduce passive waste while keeping rules, organizational structure, and contracts fixed, a dimension largely unexplored in this literature. Moreover, we show that recent technological innovations enable a cost-effective scalable monitoring technology for purchases made in the context of framework agreements.

Second, we contribute to the literature on *performance monitoring and management practices in organizations*. A large body of work studies how monitoring technologies, management practices, and incentive schemes affect worker productivity.[1] Within this literature, Bandiera et al. (2021a) show in a public procurement context that financial incentives alone do not improve performance unless officers face an active monitor — highlighting the complementarity between information and oversight that our paper studies directly. We extend this literature by showing that AI-enabled e-procurement data make individual performance monitoring feasible at virtually zero marginal cost in a public sector setting (Lewis-Faupel et al., 2016), and that this monitoring technology can achieve efficiency gains without modifying financial incentives, civil service rules, or organizational hierarchies.

2. Public Procurement in Chile

Like most public procurement systems, Chile's combines some degree of centralization with input from local purchasing units (OECD, 2019; Carpineti et al., 2006). In 2003, Chile adopted new regulations that mandated the creation of a Central Procurement Body (CPB), *Chilecompra*, located in the Ministry of Finance. *Chilecompra* creates the rules and mechanisms for purchases, but procurement itself is decentralized to procurement units located in the ministries.[2]

[1] Some studies in this literature include Hubbard (2000, 2003); Baker and Hubbard (2003, 2004); Halac and Prat (2016); de Rochambeau (2021); Kelley et al. (2021); Mattsson (2022); Bloom and Van Reenen (2007); Bloom et al. (2019); Rasul and Rogers (2018); Bandiera and Rasul (2011, 2013); Khan et al. (2016, 2019); Dal Bó et al. (2021); Balán et al. (2022)

[2] As emphasized by OECD (2019), CPBs like *Chilecompra* have a number of advantages: (a) tighter control of expenditures; (b) economies of scale with fewer suppliers; (c) more bargaining power; (d) lower administrative costs; and (e) higher productivity. In contrast, non-centralized, local bodies can be more effective when they have an information advantage in selecting local suppliers and there is heterogeneity in local purchaser preferences.

2.1. Organizational Structure and Priorities

Procurement involves three layers of authority. At the top, Chile's Public Budget Office (DIPRES, part of the Ministry of Finance) is responsible for allocating resources annually to each local procurement unit. Each local procurement unit is managed by a Director and a Financial Manager ("Manager"), who lead a team of procurement officers ("Buyers"). In 2019, there were 224 public procurement units with 9,300 registered buyers in the central government, with a median of 20 procurement officers per unit.[3] The Manager communicates purchasing needs to buyers who are in charge of implementing the solicited purchases through purchase orders.

Although managers have formal authority over buyers, their ability to discipline or reward purchasing performance is constrained by civil service regulations that fix salaries and limit discretionary personnel decisions. In practice, the instruments available to managers are non-pecuniary: setting purchasing goals, holding periodic coordination meetings, supervising individual purchase orders, and providing guidance to buyers.

The budget allocation process creates a system in which efficiency is not a primary concern. Each unit receives an annual budget, and the next year's allocation is driven by the extent to which the current year's budget allocation is fully spent. In fact, the highest proportion of purchases is concentrated at the end of the fiscal year, suggesting wasteful spending (Engel et al., 2021).[4] Indeed, our baseline survey of managers provides evidence that managerial attention is predominantly oriented toward budget execution rather than purchasing efficiency; 87% agreed with the statement "*there is pressure from DIPRES to fully spend this year's budget*," and 44% recognize that "*sometimes purchases need to be made at high prices to comply with budget execution by the end of the year*."[5] When asked what would happen if the unit failed to fully spend its allocation but could demonstrate considerable savings, fewer than one in five managers expects savings to be rewarded (Figure 1 Panel A). Managers also exhibit limited concern that DIPRES actively monitors procurement behavior or assigns significant weight to purchasing efficiency in budget decisions; 43% of managers report that savings are a low priority for DIPRES in budget allocation decisions, and only 24% report that DIPRES closely monitors its procurement expenditures (Panel B of Figure 1). This is consistent with a system where the value of generating savings is low and possibly negative, providing little basis for managers to internalize efficiency as an important performance criterion.

In fact, consistent with these priorities, managers provide little oversight of spending efficiency. Generic oversight actions are present but far from universal. About 30% of buyers report that their manager sets explicit efficiency goals, holds regular meetings about purchasing performance, or

[3]This excludes municipalities, public universities, and public enterprises, which collectively contribute an additional 6,000 buyers to the procurement system, but these public entities operate independently of the Ministry of Finance's direct oversight.

[4]b Similar findings have been documented in other contexts: For example, Liebman and Mahoney (2017) show that spending by the U.S. Federal Government in the last week of the year is 4.9 times higher than the rest of the year.

[5]The baseline survey was conducted with $N = 258$ managers (Directors and Finance Executives) prior to the intervention. A full description of the survey instrument and sample is provided in Section **??**.

monitors and supervises individual purchases (Figure 2 Panel A).[6] More targeted and information-intensive actions are essentially absent (Figure 2 Panel B). Only 6% of buyers report that their manager keeps records that track overspending by individual buyer, and only 3% report having received training specifically aimed at reducing overspending. Performance-based consequences for individual purchasing behavior are nonexistent, reported by less than 1% of buyers in the control group.

Together, these patterns suggest that the organizational environment is characterized by two compounding features. First, managers faced institutional incentives that discouraged prioritizing procurement efficiency over budget execution and limited perceived consequences for overspending. Second, the absence of systematic buyer-level performance information meant that management oversight of buyer behavior was generic and untargeted. The result was a system where buyers operated with substantial discretion and little visibility: managers could not easily distinguish efficient from inefficient purchasers, and had limited institutional reason to try.

2.2. Procurement Channels

The 2003 reform established three main purchasing channels based on the size of purchase. The first is *auctions* for large purchases above USD $74,000. Auctions were used for more than 57% of government expenditures and roughly 33% of transactions in 2019.[7] The second is *direct purchases* from a supplier chosen by the buyer for small expenditures less than USD $740. Direct purchases accounted for 17% of transactions and 20% of expenditures. The third is *Framework Agreements* (FAs) for expenditures between USD $740 and USD $74,000. FA purchases were used for 22% of government expenditures and 50% of transactions.

FAs use competitive bidding to select a set of products and services to be listed on an electronic platform called "Mercado Público" from which local buyers make purchases. It seeks a balance between a CBP competitive bidding process to lower costs and maintain quality while allowing the buyer to have flexibility in the choice of specific products and suppliers (Albano and Nicholas, 2016). FAs are implemented in two stages. First, Chilecompra posts a catalog of desired products and auctions the rights to list these products on the online marketplace to suppliers. Bids contain the highest price that sellers can charge, including delivery. Chilecompra sets other minimum service quality requirements, such as maximum delivery time, so that suppliers compete in price only. In the second stage, buyers make purchases on the online marketplace and have complete flexibility to choose among the listed products at the posted prices.

Chilecompra generates competition in the online market by (i) awarding multiple suppliers for similar products the right to list their products, and (ii) allowing suppliers to discount prices

[6] Figure 2 uses the survey of control group buyers to measure what normal managerial behavior looks like absent any external stimulus. Under random assignment, the control group provides a valid counterfactual for what treated units would have looked like absent the intervention, so we interpret control group endline means as representative of pre-intervention managerial behavior across all units.

[7] https://datos-abiertos.chilecompra.cl/

below their first stage bid at anytime. FAs use competition to enter the market and during market operation to ensure competitive prices, including delivery costs and time, and reduce buyers' search effort to find products and services that satisfy the preferences of their local units (Saban and Weintraub, 2021).[8] By increasing competition through the auction stage, the implementation of FAs reduces the chances of active waste, such as corruption in supplier selection. However, passive waste may still remain in the second stage, as buyers have discretion in choosing products in the online catalog.

2.3. Purchasing in Framework Agreements

Purchasing efficiently requires effort to search the marketplace for suitable products that meet their needs and compare their prices (Dinerstein et al., 2018). However, this can be quite complicated because most products are not easily comparable. Consider a buyer seeking for "Grated Cheese" (*Queso Rallado*) to be delivered to municipal schools. Figure 3 provides an example of the search process. Each product shown is a Product ID and the price displayed is the minimum price across all the suppliers that sell that product. The price variation between the different products is large, ranging from $659 to $3,433 due to differences in brands, sizes, and quality. The red boxes highlight the two products that have the same brand and size, whereas the other products are different brands or sizes.

Supplier prices, including delivery costs, also vary by location (see Figure 4). Because there are no standardized attributes associated with each product, the platform does not provide a clean, unified view to compare alternatives on all relevant dimensions at once. Identifying a cheaper acceptable alternative often requires additional effort: clicking more product pages, checking descriptions, making calculations, and selecting suppliers from the listed options.

Buyers within procurement units can be heterogeneous in terms of their preferences for efficiency. Qualitative research suggests that buyers purchasing differs along two dimensions: (i) how much they value price vs. quality; and (ii) the amount of effort they are willing to spend to find options in the marketplace. "Efficient" purchasers care about price and actively search for the lowest price options. "Picky" buyers engage in an active search but prioritize top-of-the-line quality regardless of price. "Lazy" purchasers are reluctant to search and choose the first option available, paying little attention to prices or quality. This heterogeneity in search effort and price-quality tradeoffs is not directly observable to managers, creating an information asymmetry that is central to understanding why performance monitoring may matter — and whether managers have the incentives to act on it.

Search costs in repeated purchases may require reevaluating options to identify the lowest price option each time. In this case, search costs depend on the number of eligible products, price dispersion, and how prices are adjusted dynamically. Some buyers may display "stickiness" in

[8]While competitive bidding can lower procurement prices (Bulow and Klemperer, 1996), there is a risk that under-qualified contractors win bids (Carril et al., 2022).

repeated purchases: rather than re-evaluating each time, buyers may stick with the same Product ID or supplier purchased in prior orders to reduce search costs (Wathne et al., 2001; MacKay, 2022). This strategy is optimal if the lowest price supplier does not change over time but is suboptimal when the lowest price supplier changes over time. Hence, the potential gains from the buyers' search effort depend on two empirical features of the FA marketplace: (1) *Cross-sectional price dispersion* across alternatives within a group of comparable products, where greater dispersion increases the savings from switching to a cheaper option; and (2) *Changes in relative price rankings over time* require buyers to re-evaluate which product and supplier on every purchase, as the cheaper options may vary dynamically (more on this is presented in Section **??**).

3. Measuring Buyer Performance

We developed a system to monitor purchase efficiency in the FA channel, which has two key advantages. First, products offered in the FA digital marketplace are classified with a standardized catalog using a set of objective attributes that capture product differentiation (OECD, 2019). In addition to observable product characteristics, FA contracts establish service delivery time and location standards, warranties, and other quality dimensions. Consequently, FA products and contractual terms are sufficiently standardized to facilitate price comparisons. Second, the FA marketplace provides a digital searchable database for all products and services each day to be able to find the prices of close substitute products for comparison purposes.

3.1. Overprice

In the FA channel, buyers purchase by issuing *purchase orders*. A *purchase order* may contain multiple *products*, and for each *product*, the buyer may purchase several *units*. *ChileCompra* provided line-item-level transaction records between February 2020 and January 2021, where each line item corresponds to a distinct product regardless of the number of units purchased per product. For each line item, the data report the quantity ordered, the per-unit price paid, the product ID, the product category, the buyer identifier, the procurement unit, the delivery region, and the transaction date. We measure purchase efficiency through an *overprice* index at the *per-unit of product*, which is defined as the difference between the per-unit price paid for the product and the lowest per-unit alternative price among comparable products available on the same date, scaled by the latter.

More formally, consider a transaction on day t in which the buyer purchases a product i from supplier s at a *per-unit price* p_{ist}. Let R_i be the *reference group* for this product, i.e., a set of products that are sufficiently similar in functionality and quality to be plausible close substitutes for product i. A reference group forms a partition among the set of products available in the marketplace, so that each product has a unique reference group. Denote a *per-unit reference price* $r_{it} = \min\{p_{jst}\}_{j \in R_i}$ as the minimum per-unit price in the reference group of product i on day t. The difference between the per-unit price paid and the per-unit reference price, $p_{ist} - r_{it}$, is the potential savings that buyers would generate by purchasing the lowest-price substitute. Scaled by

the latter, the per-unit *overprice* (op_{ist}) is defined as:

$$op_{ist} = \frac{p_{ist} - r_{it}}{r_{it}}. \tag{1}$$

This metric has a simple interpretation: the proportional savings *per-unit of purchased product* that could have been achieved on date t if the buyer had selected the lowest-priced alternative from the reference group.

3.2. Building reference groups

The validity of op_{it} depends critically on how the reference group is constructed. Products grouped into a reference group must be sufficiently comparable in functionality and quality so that they constitute reasonable substitutes for the product purchased. Grouping based solely on identical products (*Product ID*) is too narrow a definition in most cases and may omit many meaningful substitutes, limiting the opportunity to reduce spending. For example, considering only `ballpoint pens` when `felt tip` might work just as well would miss a relevant close substitute. Groupings of *close substitutes* that include multiple Product IDs sharing the same quality-relevant attributes would be broader. This approach provides a reasonable balance by constructing substitute sets tight enough to preserve comparability while large enough to make reasonable price comparisons, but it requires identifying detailed product characteristics for thousands of products, which may be difficult to scale.

Constructing such groups requires two things: structured information on what each product is, and a rule for how finely to group products on that information. We take these in turn, and then characterize the price variation within the resulting groups, which determines how much a buyer stands to gain from search.

From free text to product attributes. Constructing attribute-based reference groups requires converting open-text product descriptions into structured information. *ChileCompra* developed a new catalog design in the auction stage of the FA, with the objective of generating price competition to enter the market (Olivares et al., 2025). Each FA catalog maintains a tagged data system that maps a product's free-text description to standardized categorical and numerical attributes, and we use these attributes to define reference groups. Tagging is overseen by *ChileCompra* managers, assisted by automated tools based on an NLP algorithm adapted from Sun et al. (2014), itself built on *Word2Vec* (Mikolov et al., 2013).[9]

Catalogs differ substantially in how many attributes they tag and how consistently they do so, and not all support reference groups of comparable quality. Working with *ChileCompra*, we identified three catalogs in which attributes are tagged densely and consistently enough to construct reference groups: Computers, Office Supplies, and Food. We restrict the analysis to these categories. Table 1 lists examples of product types in each category together with the attributes

[9]This work was conducted in 2019, prior to the massification of Large Language Models.

reliably identified for each. [10]

For example, consider `Grated Cheese` as described in Table 2, which shows the extracted attributes for the two products highlighted in red in Figure 3.[11] Each row corresponds to a unique `Grated Cheese` product in the catalog. The two share the same brand and weight and are sold by the unit. Because comparing prices is only meaningful across products of the same weight and quantity, exact matches on those attributes are required for products to enter the same reference group; the remaining products on the search page in Figure 3 are excluded on this basis alone. The two products do differ in one attribute, flavor: one is *parmesano*, the other *reggianito*. Whether this difference warrants separating them is a judgment about quality, and the price data are informative. The last column of the table reports each product's average price across all suppliers offering it during the baseline period, January to May 2020; the two differ by about 4 percent. A gap this small is unlikely to reflect a difference in quality that would invalidate a price comparison, and *ChileCompra* accordingly treats flavor as non-essential here and keeps the two products in one reference group.

Defining groups. Any reference group definition faces a trade-off. Narrow groups guarantee that the products being compared are genuinely close substitutes, but may leave few alternatives against which to benchmark a purchase, so that a buyer paying the only available price registers no overprice even when cheaper substitutes exist. Broad groups expand the set of alternatives, but risk attributing to buyer inefficiency a price gap that in fact reflects unobserved quality. We assess three candidate definitions against this trade-off: (i) identical products (same SKU), (ii) products sharing the same tagged attributes, and (iii) products of the same Product Type.

Two statistics summarize the price variation within a reference group, and we use them both to compare candidate definitions here and to characterize the returns to search in Section **??**. The first measures *cross-sectional price dispersion* between products. Let $\bar{p}_i = \frac{1}{T}\sum_t p_{it}$ denote the average price of each supplier-product, indexed i, over a period T, which removes the variation associated with dynamic pricing. For each reference group R, the *Coefficient of Variation* is $CV_R = \sigma_R/\bar{p}_R$, where $\bar{p}_R$ and σ_R are the mean and standard deviation of the average prices $\bar{p}_i$ across all products $i \in R$.

The second captures *time variation in price rankings*. Define $P_t = \{p_{it}\}_{i \in R}$ as the set of prices in reference group R on date t, and $w_{it} = \text{rank}_{P_t}(p_{it})$ as the position of the product's price in this ranking. The standard deviation of this ranking for each product, $SDRank_i = \text{StDev}(\{w_{it}\}_t)$, captures how frequently that product changes its relative price position over time. The average of these standard deviations across all products $i \in R$ constitutes the *Price Ranking StdDev* for the reference group.[12]

[10] Some products cannot be tagged informatively enough to support a quality assessment — most notably several fresh-food categories. *ChileCompra* excluded these from efficiency reporting. The exclusions were made by *ChileCompra* in the course of building the monitoring system, prior to and independently of the experiment.

[11] Appendix Table 14 provides additional examples for other product categories, showing product descriptions along with the extracted attribute values.

[12] An alternative is to measure the longitudinal price variation of each supplier; however, this does not necessarily

Using baseline data for Computers, Office Supplies, and Food catalog purchases during the period January to May 2020, Table 3 reports, for each candidate definition, the number of suppliers per group, the Coefficient of Variation, the Price Ranking StdDev, and Overprice. First, “Same Product Type" group fail the comparability requirement decisively: they average 393 suppliers, a coefficient of variation of 0.64, and a mean overprice above 1,000 percent. Dispersion of this magnitude reflects quality heterogeneity rather than buyer inefficiency, and these groups are unusable for benchmarking.

Second, “Identical-product" group sit at the opposite end. They are comparable by construction, and their low dispersion (CV of 0.08) confirms it, but they are also the smallest, with 8.2 suppliers on average. “Same Attributes" groups occupy the middle: 12.2 suppliers on average, with a CV of 0.11 and ranking variability comparable to “Identical-product" group, so the broader group admits more alternatives without importing the dispersion that signals quality heterogeneity. “Same Attributes" groups nevertheless yield a mean overprice of 0.21, roughly double that of “Identical-product" groups. This gap could reflect either the larger choice set or residual unobserved quality differences, and the data do not distinguish between the two.

We therefore adopt a Mixed definition, combining identical and same attributes. Products enter the same reference group when they share the same Product Type and identical values of all identified attributes, including brand, and when they are sold in the same geographic region. The regional dimension matters because Framework Agreement contracts regulate delivery standards — delivery fees, shipping costs, fulfillment performance, product availability, shipping arrangements, and delivery schemes — at the regional level, and suppliers face different delivery costs and conditions across areas. Constructing groups separately by region therefore holds service and logistics conditions fixed within any price comparison.

The Mixed definition further isolates outlier products into single-product groups. *ChileCompra* continuously revises its reference group definitions, comparing the price spread across brands and other attributes to evaluate whether these capture quality differences that merit separating products out. Part of this process identifies prices exhibiting abnormally high overprice, which are typically unique transactions associated with unusual product characteristics. To be conservative, product IDs with an overprice greater than 100 percent are removed from their original reference group and reassigned to single-product groups. This preserves nearly the full choice set (11.7 suppliers, against 12.2) while bringing the mean overprice from 0.21 to 0.11, close to the identical-product benchmark. The isolated observations are precisely those where an unobserved quality difference is most plausible, so the restriction removes the cases in which attribute-based grouping is least defensible, at little cost in comparison scope.

Two caveats bear noting. First, the choice is a judgment about where to sit on the comparability–coverage trade-off rather than a validated optimal approach, since we observe no independent measure of true substitutability. Second, the 100 percent threshold is conservative by construction,

imply changes in the price order of suppliers if price shocks are common to all suppliers in a reference group (i.e. these could be “parallel” shifts of the price time series).

following *ChileCompra*'s own practice of avoiding comparisons across products of mixed quality; it is not derived from the data. When presenting results, we therefore report estimates using the strictest alternative (identical-product groups only) as a robustness check on how far the findings depend on this choice.

Under the Mixed definition, 514,398 of the 867,916 products transacted between February 2020 and January 2021 have a well-identified reference group and hence a computable per-unit overprice, representing almost 60 percent of demand. Of these, 206,704 were purchased by buyers working in procurement units subject to the experiment, and constitute our sample of analysis (more in Section 5).

Price dispersion. The potential gains from the buyers' search effort depend on the two sources of price dispersion defined above. Greater cross-sectional dispersion increases the savings available from switching to a cheaper alternative within the group. Changes in relative price rankings over time require buyers to re-evaluate which product and supplier to choose on every purchase, as the cheaper option may vary dynamically. The potential returns to search are therefore larger when comparable products exhibit both higher price dispersion and higher ranking variation.

Figure 5 illustrates the distinction, presenting posted prices for two product groups between January and May 2020. The top graph shows the `Grated Cheese` example from Section 3.2, which includes the two Product IDs of similar characteristics but different flavors. Each Product ID is represented by a different color, and the time series represent the prices of distinct suppliers. Most of the price variation is cross-sectional, with a Coefficient of Variation of 5 percent. Interestingly, there is negligible variation in prices over time: the ranking of suppliers by price remains essentially the same throughout. For this type of product, a buyer can invest once in searching among the available options to find the cheapest one and then stick to it for future purchases.

In contrast, the bottom panel of Figure 5 illustrates the price dynamics of a `Cookie Snack` (Galleton Nutrabien); this product group includes four Product IDs with similar characteristics (each one represented with a distinct color) offered by multiple suppliers. The price dispersion across suppliers is similar to the `Grated Cheese` example, with a Coefficient of Variation around 5 percent. However, the dynamics over time differ, showing frequent changes in the relative price rankings across the purchasing options. While for `Grated Cheese` (top graph) finding a low-price supplier and sticking to it provides a reasonable purchase strategy, this "stickiness" performs worse for the `Cookie Snack` (bottom), which requires a new search on every purchase to identify the lower-priced supplier at each point in time.

4. A Model of Procurement with Search Frictions

4.1. Set Up

Consider a procurement unit composed of a Manager and a Buyer. The Buyer is tasked with purchasing a standardized product at the lowest possible price. In the absence of corruption or

active waste, passive waste (inefficiency) may still arise if the Buyer's search effort to seek out the lowest price is insufficient. The Manager can try to persuade the Buyer to put in more search effort through a series of managerial actions such as setting efficiency goals, holding regular meetings, supervising purchases, creating purchasing manuals, training programs, and performance-based rewards, to name a few. However, the manager only observes the price paid and not the Buyer's search effort.

Suppose that the Buyer is extrinsically motivated and derives no utility from low prices *per se*. In this case, he only supplies search effort in response to rewards for finding lower prices and managerial pressure. By default, he exerts minimal search effort and obtains a purchase price μ for the product[13]. The Manager can try to persuade the buyer to provide a higher search effort through a set of actions. Doing so can potentially lower the price down to a minimum of $p^m \leq \mu$.[14] Let $a \in [0, A]$ represent the intensity of managerial actions, with $c(a)$ being the cost of a, which we assume strictly increasing and convex.[15]

The effectiveness of the Manager's actions depends on the Buyer's responsiveness, captured by the Buyer's type $\theta_B \in \Theta \subseteq [0, 1]$. A Buyer with a higher θ_B is *more receptive to managerial persuasion*, making it more likely that the Manager's actions successfully induce the price p^m. The Manager persuades the Buyer to buy at price p^m with probability $P(\theta_B, a) = \theta_B \phi(a)$, with $\phi(a)$ strictly increasing and concave function of a.

We assume that θ_B is unobservable to the Manager, i.e., she does not know how responsive the Buyer is to her persuasive actions. However, the Manager can learn about θ_B through information $I \in [0, \infty)$ accumulated from the Buyer's purchase history and his efficiency performance. The more information the Manager accumulates, the more accurate is her estimate of θ_B, allowing her to better adjust actions. In the absence of any information, the Manager's best guess is the population average $\bar{\theta}$. Using information I, the Manager's belief about θ_B is given by:

$$\hat{\theta}_B(I) = \alpha(I)\bar{\theta} + (1 - \alpha(I))\theta_B \tag{2}$$

with $\alpha(I)$ strictly decreasing in I, $\alpha(0) = 1$ and $\alpha(I) \xrightarrow{I \to \infty} 0$. Thus, if the Manager has no information, the Manager's belief is $\bar{\theta}$. As information goes to infinity, $\hat{\theta}(I)$ converges to θ_B.[16]

[13] One can think of minimal effort as a single draw in a sequential search process: the Buyer samples one price from the distribution of available prices and accepts it, so the expected purchase price is the mean μ.

[14] In principle, the buyer can face a distribution of multiple prices for the sought product, yet for simplicity we assume the buyer decides between two potential prices, either μ or the minium price.

[15] The intensity of managerial actions could refer to the number of actions taken or the quality of those actions, yet for tractability we do not make a specific distinction of attributes of the actions.

[16] How information is used by the Manager to update her belief can be rationalized as follows. Suppose the Manager holds a Normal prior $\theta_B \sim \mathcal{N}(\bar{\theta}, \sigma_0^2)$ and each period observes a noisy signal $s_t = \theta_B + \varepsilon_t$, with $\varepsilon_t \sim \mathcal{N}(0, \sigma_\varepsilon^2)$ i.i.d. After accumulating I signals with sample mean $\bar{s}_I$, the posterior yields $\hat{\theta}_B(I) = \alpha(I)\bar{\theta} + (1 - \alpha(I))\bar{s}_I$, where $\alpha(I) = \sigma_\varepsilon^2/(\sigma_\varepsilon^2 + I\sigma_0^2)$. Notice that the posterior mean depends on the realized signals $\bar{s}_I$, which are stochastic. As $I \to \infty$, $\bar{s}_I \to \theta_B$ almost surely. Equation B.1 is a simplification that keeps the model tractable and is sufficient to capture the key dynamic: as the Manager accumulates more information, her belief transitions from the prior $\bar{\theta}$ toward the Buyer's true type θ_B, allowing her to better tailor her actions over time.

4.2. The Manager's Problem

The Manager's objective is to minimize overspending net of the cost of persuasion effort. Let $v(p)$ be the Manager's utility from purchasing the item at price $p \in (p^m, \mu)$. We assume that $v(\cdot)$ is strictly decreasing and concave. Then, the Manager's payoff is given by:

$$u_m(a, \hat{\theta}_B(I)) = P(\hat{\theta}_B(I), a)\bar{v} + (1 - P(\hat{\theta}_B(I), a))\underline{v} - c(a), \tag{3}$$

with $\bar{v} = v(p^m)$, and $\underline{v} = v(\mu)$.

Buyers are required to follow a purchasing protocol, and at the minimum, Managers are required to oversee that the Buyer correctly follows this protocol. Thus, we will assume that the Manager must exert at least a minimum intensity of managerial actions $a_0 = \arg\max\{u_m(a, \bar{\theta}) \mid a \geq 0\}$. Thus, the Manager's problem is defined as:

$$a^*(I) \in \arg\max\{u_m(a, \hat{\theta}_B(I)) \mid a \geq a_0\}, \tag{4}$$

The Manager's optimal action intensity depends on her belief about the Buyer's type, $\hat{\theta}_B(I)$. If the Buyer's true type is lower than $\bar{\theta}$, then she will realize that the Buyer is less responsive to her actions. Therefore, her optimal action profile will not change, $a^*(I) = a_0$. Conversely, if $\theta_B > \bar{\theta}$, new information reveals that the Buyer is more responsive than the Manager initially believed, and hence will optimally increase the intensity of her actions above the minimum a_0.[17]

Once the action profile $a^*(I)$ is chosen, the probability of achieving the lowest price depends on the true Buyer's type θ_B. Let $f : \mathbb{R}_+ \to \mathbb{R}$ be an exogenous efficiency measure such that $f(p)$ is the efficiency of a purchase of the item at price $p \in (p^m, \mu)$, with f strictly decreasing in p. (For instance, $f(p)$ could be defined as the overprice of the purchase: $f(p) = -(p - p^m)/p^m$.) Then, the purchase efficiency from Manager's actions profile is given by:

$$\pi(I) = P(\theta_B, a^*(I))f(p^m) + (1 - P(\theta_B, a^*(I)))f(\mu). \tag{5}$$

It follows that the effect of information in purchase efficiency is driven solely by Manager's actions. Purchase efficiency $\pi(I)$ is constant in I for low types ($\theta_B < \bar{\theta}$) and strictly increasing in I for high types ($\theta_B > \bar{\theta}$). Furthermore, the magnitude of these effects depend on how close is Buyer's purchase choice relative to what the Manager aims, $\bar{v} - \underline{v}$. More formally, the following testable predictions emerge:

Prediction E.1. *Purchase efficiency $\pi(I)$ is weakly increasing in $I > 0$. If $\theta_B > \bar{\theta}$, then purchase efficiency $\pi(I)$ is strictly increasing in $I > 0$.*

Prediction E.2. *Intensity of Manager's actions $a^*(I)$ is weakly increasing in $I > 0$.*

[17] To avoid uninteresting cases, we assume that $\lim_{a\to 0} \phi'(a) = +\infty$, $\lim_{a\to\infty} \phi'(a) = 0$, $\lim_{a\to 0} c'(a) = 0$, and $\lim_{a\to+\infty} c'(a) = +\infty$.

Prediction E.3. *Persuasion probability $P(\theta_B, a)$ of buying at p^m is weakly increasing in $I > 0$.*

4.3. Price Dispersion

Are there bigger gains to information in reference groups with more price dispersion? We start this discussion by building price dispersion into the model in the context of a Buyer sequential search process (Weitzman, 1979). Each search is a random sample of a price at cost $\kappa > 0$ from a distribution $G(\sigma)$ with support $\Gamma = [p^m, \bar{p}]$. Let μ and σ^2 the mean and variance of $G(\sigma)$. By default, absent any managerial action, the Buyer exerts a minimal effort and samples only once, yielding the expected price μ. The larger is the number of searches the lower is the expected price obtained from searching. The Manager's actions persuade the buyer to increase the number of searches such that minimum price p^m is obtained.

The return to searching depends on the price dispersion, captured by the variance of the price distribution σ^2. When the distribution G widens, the best available price p^m falls. Thus, the Manager has more incentives to persuade the Buyer to purchase at p^m, implying the intensity of her managerial actions is increasing in price dispersion σ. As both p^m and a^* can change with σ, then efficiency $\pi(I)$ indirectly depends on σ.

The two mechanisms work as follows. Define efficiency $\pi(I)$ using the overprice measure $f(p) = -(p - p^m)/p^m$. First, in the absence of manager persuasion actions, the minimum price p^m falls if price dispersion σ increases. Thus, $f(\mu)$ decreases with larger dispersion, generating efficiency losses. Second, as discussed above, greater dispersion increases $a^*(I)$ and the probability of purchasing at price p^m, generating efficiency gains. Further, the gains from information can be strengthened or weakened by price dispersion. Given that the Manager increases $a^*(I)$ when she faces bigger dispersion, there might be decreasing marginal returns from information. [18] In general, the direction of these effects is ambiguous. Therefore, the following predictions hold:

Prediction E.4. *Purchase efficiency $\pi(I)$ increases with σ only if the efficiency gains from the Manager's persuasion actions exceed the efficiency loses from lowering the minimum price p^m.*

Prediction E.5. *The effect of information on purchase efficiency increases with price dispersion σ only if the Manager's incentive to increase persuasion actions when prices are more disperse dominates the diminishing returns to further managerial actions.*

4.4. Buyer Performance Gap

Does information get the most from buyers where there is the most to gain, i.e., those who are the worst performers? Purchase efficiency depends directly on the potential reduction in price from persuasion, $\Delta p = \mu - p^m$. The gap Δp increases the Manager's incentive to persuade the Buyer, but this comes at the cost of decreasing $f(\mu)$ if persuasion fails. Thus, the net effect on efficiency

[18] Determining the direction of these effects requires further assumptions on the functional form of ϕ. See Appendix C.

depends on whether the gains from more intense managerial actions outweigh the losses from a less favorable counterfactual outcome.

A similar dynamic shapes how effective information is: since both dispersion and Δp widen the potential gains from persuasion, they raise the Manager's incentive to respond to information, but diminishing marginal returns to managerial actions mean this incentive may not translate into much additional persuasion once the Manager is already acting intensely. Whether the effect of information on purchase efficiency is strengthened or weakened by price dispersion or by Δp therefore depends on which of these forces dominates (see Appendix C for derivations).

Prediction E.6. *Purchase efficiency $\pi(I)$ increases with Δp only if the efficiency gains from the Manager's persuasion actions exceed the efficiency losses from a less favorable counterfactual outcome $f(\mu)$.*

Prediction E.7. *The effect of information on purchase efficiency increases with Δp only if the Manager's incentive increase actions when the potential reduction in prices is large dominates the diminishing returns to further managerial actions.*

4.5. Intrinsically motivated Buyers[19]

Thus far, the Buyer is extrinsically motivated, as he only searches for low prices if the Manager undertakes managerial actions. Suppose now that the Buyer has no extrinsic motivation but intrinsically values low prices by $u_b(p) = -p$. The Buyer follows an optimal stopping rule characterized by a reservation price r, accepting the first sampled price below r [20]. Importantly, when the Buyer is only intrinsically motivated, Manager's persuasion actions have no effect on the Buyer's search decision, implying information I acquired by the manager has zero effect on the purchased price through Manager's actions. Therefore, Predictions **??** and **??** do not hold in this case.

However, information still plays a role: rather than inducing search, it helps the Buyer to find lower prices and adjust his reservation price accordingly. The Buyer has incomplete information about the true price distribution G. The Manager can share information $I \geq 0$ that informs the Buyer about the true price dispersion, expanding the range of prices the Buyer perceives as accessible. Knowing the availability of lower prices allows the Buyer to optimally decrease his reservation price. The expected purchase price $p^*(I)$ is therefore a strictly decreasing function of the information I, which in turn determines purchase efficiency $\pi(I) = f(p^*(I))$. Hence, predictions **??** also hold when the Buyer is intrinsically motivated. The following predictions differ for intrinsically motivated buyers:

Prediction N.1. *Purchase efficiency $\pi(I)$ is strictly increasing in $I > 0$.*

[19] A buyer can be intrinsically motivated and still respond to extrinsic motivation. See Appendix Section **??** for details. This case offers mixing predictions that we cannot test through our experiment.

[20] See Appendix **??** for a full derivation of the search model

Prediction N.2. *Intensity of Manager's actions $a^*(I)$ is constant in $I > 0$.*

Prediction N.3. *Persuasion probability $P(\theta_B, a)$ of buying at p^m is constant in $I > 0$.*

Analogously to the potential price reduction from persuasion in the extrinsic model, define Δp as the potential reduction in price available to the Buyer from search. Specifically, Δp is the gap between the expected price he would obtain if he knew the full extent of price dispersion in the market, and the expected price when he has no information, $I = 0$. This gap captures how much room the Buyer has to improve using the information provided by the Manager. Since the expected purchase price increases as the Buyer accesses to a narrower range of prices, a wider Δp means the Buyer starts further from the best attainable outcome. At the same time, a greater gap Δp leaves more scope for search to lower his purchase price. It follows that purchase efficiency increases with higher price dispersion σ, but decreases with a wider price gap Δp. However, whether price dispersion and Δp amplify the gains from information is not clear in general. Information and price dispersion increases the returns from search, but the reservation price can become less sensitive to information as σ grows, to further increments in his knowledge about prices (see Appendix C for a detailed explanation). The relationship between σ, Δp and information cannot be determined without further assumptions about the underlying price distribution G. We summarize the discussion above in the following predictions:

Prediction N.4. *Purchase efficiency $\pi(I)$ is increasing in price dispersion σ.*

Prediction N.5. *The effect of information on purchase efficiency is reduced by price dispersion σ when the reservation price becomes insensitive to information as sigma grows.*

Prediction N.6. *Purchase efficiency $\pi(I)$ is decreasing in the price gap Δp.*

Prediction N.7. *The effect of information on purchase efficiency is strengthen by price gap Δp when the reservation price responds sharply to information as the Buyer has less knowledge about G at baseline.*

5. Intervention and Experimental Design

We test the predictions of the model by generating experimental variation in buyer performance information I. In this section, we describe the information intervention and the experimental design used to test the predictions.

5.1. Buyer Performance Reports

In collaboration with *ChileCompra* and Chile's Public Budget Office (DIPRES), we constructed monthly overspending reports for each buyer for the period September 2020 to January 2021. The report is signed by DIPRES and sent by email in the first week of each month, with information covering the previous month. The emails carried login credentials to a centralized platform where

each buyer had an individual homepage from which reports could be downloaded at any time. Reports were confidential to each buyer: individual performance information was not shared with other buyers in the purchasing unit.

Appendix Figure B.1 shows an example of the reports. Each monthly report tells the buyer their personal overspending rate and ranks it against those of peer buyers in the purchasing unit.[21] It also breaks down the share of buyers in the unit by overspending band, showing where each buyer falls relative to their peers, and displays the buyer's distribution of purchases across procurement mechanisms (Framework Agreement, Direct Purchases, Auctions, etc.). The report also provides the unit-level overspending rate, highlighting the unit's position relative to other public services within the state. Finally, the report provides recommendations encouraging buyers to prioritize competitive procurement mechanisms and conduct thorough price searches to reduce overspending.

Buyers were prepared for the reports in advance. In June 2020, three months before the first report was issued, treated buyers received a link by email with a 10-minute training video explaining how *ChileCompra* measures overspending and giving step-by-step examples of efficient search on the Mercado Público platform.

5.2. Study Arms and Random Assignment

The Buyer Performance Reports generate variation in the performance information reaching the buyer, call it I_B. Remember that this is confidential and thus not shared with other buyers in the purchasing unit. However, we also vary I_M, i.e., whether the same information reaches his manager, for which we specify two versions of the intervention: *private* and *public*. Both interventions raise I_B identically (every treated buyer receives the Buyer Performance Reports), but differ in whether I_M is active or not. In the private information intervention, I_B is not shared with the manager and the buyer knows that I_B is private information. In the public intervention I_B is shared with the manager and the buyer knows I_B is no private information, thus he is aware that is exposed to individual accountability. Comparing the two interventions holds the intrinsic channel fixed and switches on the extrinsic channel.

The study design employs a cluster randomization at the procurement unit level where each unit and their buyers were randomly assigned into one of three following arms:

1. ***Private Intervention.*** The buyer receives a 10-minute training video and the Buyer Performance Report, which carries a non-disclosure message "Your individual report has NOT been sent to the unit's manager." On the manager's side, she receives a performance report aggregated at the unit-level (Unit-level Performance Report), which does not allow managers to identify individual buyer performance.[22]

[21] The *ChileCompra* data are confidential and not available to the public or other government institutions, meaning that neither buyers nor the public services they work for can calculate overspending on their own, yet the report provides a detailed explanation of how overspending is calculated (Appendix Figure B.2).

[22] See Appendix Figure B.3 for an example of the Unit-level Performance Report.

2. ***Public Intervention.*** The buyer receives a 10-minute training video and the Buyer Performance Report, but with a disclosure message "Your individual report has been sent to your manager." The manager receives the Unit-level Performance Report as well as individual Buyer Performance Reports plus a list ranking each buyer by overspending and contribution to the unit total, allowing her to identify individual buyer performance.[23]

3. ***Control.*** Buyers and managers receive no information.

Table 4 summarizes the the details of the specific activities in each of the intervention arms. In addition, within procurement units assigned to the Private or Public intervention arms, 75% of buyers were randomly assigned to the main study sample and the remaining 25% to a placebo sample who just receive a simple message indicating that their overspending is being monitored, but do not provide the 10-minutes training video nor the Buyer Performance Reports. We use the placebo sample to test whether the information reports raise efficiency by simply drawing attention to overspending (Hawthorne effect), rather than through conveying information.

5.3. Hypotheses and Identification

Table 6 maps the study design to the predictions of the model. Both the Public and Private intervention raise the buyer information I_B, but only the Public information additionally raises the manager information I_M. Thus, only in units exposed to the Public intervention buyers can be held accountable for their individual purchase efficiency, which allows the manager to implement oversight and persuasive actions effectively. Hence, we test the predictions for buyers whose are extrinsically motivated by comparing the public information treatment group to the control group.

If buyers are extrinsically motivated, purchase efficiency should increase under the Public intervention, but not in the Private intervention. (Prediction E.1). The effect of performance information in the pubic intervention is strictly positive only for buyers who are responsive to managerial actions (Prediction E.1). We distinguish buyers who are responsive from those who are not based on their answers to whether they prioritize getting a lower price over other criteria under different probabilities of being audited by their managers collected at baseline.

The effect of the Public intervention operates through two mechanisms: managerial oversight and persuasive actions, and buyers responsiveness to those actions. If buyers are extrinsically motivated, then the Public intervention should increase the intensity of managerial actions a^* (Prediction E.2). We measure a^* using manager oversight and communication activities collected from the buyer's follow-up survey. Second, because buyers know that they can be held accountable by their manager in the public intervention, we expect information to increase their responsiveness, θ_B, to increase (Prediction E.3). We measure θ_B at end line from buyers' stated willingness to prioritize purchase efficiency when their purchases are audited.

[23] See Appendix Figure B.4 for an example of the ranking list.

In contrast, if buyers are intrinsically motivated, purchase efficiency should increase under both the Public and Private interventions, since both share individual performance information with the buyer (Prediction N.1). A buyer who intrinsically values low prices lowers his reservation price once he learns his level of overspending and what prices are attainable, so purchase efficiency π should rise. The latter implies that a null in this arm is evidence against the intrinsic channel. Further, since buyers can only be held accountable under the Public intervention, the Private intervention should have a null effect on managerial actions a^* and on buyer responsiveness θ_B, regardless of whether buyers are extrinsically or intrinsically motivated (Predictions N.2 and N.3).

This gives us two testable patterns to identify buyers' motivation using the difference in treatment effects across arms: if both the Public and Private interventions raise purchase efficiency, buyers are intrinsically motivated; if only the Public intervention raises purchase efficiency while the Private intervention has no effect, buyers are purely extrinsically motivated. Comparing buyers exposed to Private versus Public interventions holds I_B fixed and moves I_M and θ_B together, and so isolates the extrinsic channel: any additional efficiency gain must operate through managerial targeting and changes in individual receptiveness to managerial persuasion rather than through information about individual performance and prices. Note, however, that the public arm moves I_M and θ_B jointly and cannot separate them, i.e., the design identifies the extrinsic channel as a whole, not the manager's actions and the buyer's responsiveness as distinct forces.

In addition, we test two sources of heterogeneity: the potential reduction in price Δp, and the underlying price dispersion. Price dispersion reflects how much variation exists in the market a buyer faces, independently of what he already knows about it at baseline. Δp is the gap between the buyer's baseline purchase price and the best attainable price, capturing how much room a buyer has to improve given what he already knows about the market. We measure Δp by ranking buyers within their organization according to their average overprice at baseline and identify, for instance, whether a buyer falls within the bottom quartile of the organization's overprice distribution. We measure price dispersion in two ways, following the reference-group statistics defined in Section **??**. The Coefficient of Variation captures cross-sectional price dispersion between products within a reference group. The Price Ranking StdDev captures time variation in a product's relative price position within its reference group. Both measures relate to the buyer's room for improvement, but capture different sources of inefficiency, so testing both lets us separate the two channels.

Price dispersion and Δp operate similarly whether buyers are extrinsically or intrinsically motivated. In the Public intervention, managers observe individual buyer performance and can target their oversight efforts toward buyers with the largest overspending, so the gains from managerial action are larger when a bigger price reduction is attainable. Similarly, individual performance information may be particularly useful to buyers who are less efficient at baseline: intrinsically motivated buyers who perform poorly have larger returns to search, whether because they face less accurate baseline information or because they face greater price dispersion. The direct effect of these two variables on efficiency, however, differs across motivations: for extrinsically motivated buyers, the effect of price dispersion and Δp on efficiency is ambiguous, since both

raise the Manager's incentive to act but also widen the gap she needs to close (Prediction E.6). For intrinsically motivated buyers, price dispersion unambiguously improve performance (Prediction N.4), since a wider gap simply means more room for the Buyer's own search to matter. In contrast, larger Δp decreases efficiency, as buyers baseline knowledge about the price distribution is narrower (N.6) However, regardless of the nature of the buyer's motivation, whether these two variables strengthen or weaken the gains from information is ambiguous, and is one of the questions we aim to test [24] (Predictions E.5, N.5, E.7 and N.7).

We expect the direct effect of price dispersion and Δp on purchase efficiency to differ across treatment arms. If buyers are extrinsically motivated, the Private intervention does not allow the Manager to target inefficient buyers, so we expect efficiency to be lower in more dispersed markets and among buyers who were performing poorly at baseline, since neither the wider market nor the larger potential gain is exploited without managerial action. If buyers are intrinsically motivated, we instead expect a positive direct effect of the Private intervention, since search alone can turn a wider market or a larger baseline gap into a lower purchase price. However, as stated above, the direct effect of price dispersion and Δp in the Public intervention is ambiguous.

Regarding the effect of price dispersion and Δp on the gains from information, if buyers are extrinsically motivated, the Private intervention should have a null effect. If buyers are intrinsically motivated, the direction of this effect depends on whether the reservation price's response to price dispersion and Δp is reinforced or offset as information accumulates. Finally, the sign of these interactions under the Public intervention is in general ambiguous.

Finally, the reports could in principle raise efficiency simply by drawing attention to overspending rather than by conveying information. We use the placebo arm to test for this potential channel directly.

5.4. Timeline

The study period runs from February 2020 to January 2021 and covers 184 procurement units, each with a single manager and multiple buyers. We divide the timeline into two periods: pre-intervention from January 2020 to May 2020 and post-treatment period from July 2020 to January 2021. Between February and April 2020, a baseline survey is administered to managers and buyers. During June 2020, we randomized the procurement units into the experimental groups. Randomization is at the level of procurement unit, and we stratify by Ministry and by whether unit overspending was above or below the median in the previous fiscal year.[25] Immediately after randomization we sent the training video link by email to buyers in the Public and Private intervention groups. Then, three months later, starting in September 2020 through January 2021, managers and buyers received the monthly overspending performance reports. After sending the

[24] Appendix C provides the full derivation of these predictions.

[25] The distribution of experimental procurement units per ministry is detailed in Appendix Table B.1. Procurement units whose yearly budget assigned by DIPRES is unaffected by their level of overspending are excluded from the experiment, e.g., state autonomous institutions like the Central Bank, courts, or security agencies, to name a few.

last report (January 2021), a follow up survey was collected from managers and buyers.

5.5. Analysis Sample

Table 5 summarizes the experimental design and sample sizes. We restrict the analysis to 3,270 buyers who used the FA platform and specialized in Food, Office Supplies, or Computers, the categories for which *ChileCompra* can calculate overspending from transaction data and thus produce performance reports. These buyers collectively issued 192,528 transactions during the study period. Our main analysis sample consists of the 2,076 experimental buyers who made FA purchases during the post-treatment period, for whom we can measure the impact of the interventions on overspending. Among these buyers, 89% clicked on the training video. For the reports, we tracked whether users logged into the centralized reporting platform at least once during the five-month intervention period. The resulting login rate was 40% among buyers and 99% among managers. Both buyer and manager take-up rates are well balanced across treatment arms.[26]

5.6. Data

Our objective is to test the predictions of the model. The key variables of interest are purchase efficiency π, buyer's efficiency gap $\bar{v} - \underline{v}$, price dispersion σ^2 in the reference group, information I, managerial persuasive actions a, and buyer's persuasion probability $P(a, \theta_B)$. We measure them through two sources of data: *ChileCompra* administrative data and surveys of buyers and managers.

ChileCompra Administrative Data. *ChileCompra* maintains records of the universe of purchase orders across all three channels (FA, Auctions, Direct Purchases). For each purchase order, we know which purchase channel was used and the buyer's ID. For purchases made through the FA channel, the data contain detailed product-level information, including product ID, product category, quantity ordered, unit price paid, buyer ID, procurement unit, delivery region, and transaction date. Together with the catalog data used to construct reference groups, we measure overprice op_{it} (our measure of purchase efficiency π), the number of units purchased per product in a purchase order and the number of purchase orders made by the buyer. These data are also used to calculate the buyer-level efficiency gap $\bar{v} - \underline{v}$ and the product-level price dispersion σ^2 in the reference group from pre-intervention information.

Baseline Survey. The survey applied to procurement units managers covered beliefs about DIPRES monitoring, perceived oversight of procurement, and attitudes toward purchasing efficiency. The buyer survey included analogous information and belief measures as well as a stated preference module in which buyers were asked whether they would prioritize price over quality under different hypothetical audit scenarios. The response to these audit questions are used as proxy of buyer's

[26]The p-values for the null hypothesis of equal take-up rates across the Private and Public arms are 0.872 for buyers and 0.263 for managers.

persuasion probability $P(a, \theta_B)$.

Follow-up Survey. The manager survey covered modules similar to those in the baseline survey, as well as targeted information outcomes to test whether access to the reports changed managers' knowledge of their unit's overspending. The buyer follow-up survey covered modules similar to those in the baseline survey, plus (i) targeted information outcomes, including knowledge of buyers' own and their unit's overspending; (ii) a stated preference module analogous to the baseline module, to examine changes in $P(a, \theta_B)$; and (iii) a module in which buyers list the managerial actions to which they were exposed during the intervention period, which we use to measure a^*..

5.7. Baseline Balance

Purchasing Behavior. The analysis sample consists of the 2,076 buyers in experimental procurement units who issued FA purchase orders in Food, Office Supplies, or Computers (FOSC) during the post-treatment period (July 2020 – January 2021), the buyers for whom the effect of the interventions on overspending can be measured. Because the sample is defined on post-treatment purchasing, internal validity requires two things: that entry into the sample does not differ across experimental arms, and that among those who enter, pre-treatment purchasing behavior is comparable. Appendix Table B.2 addresses both, reporting pre-treatment means and tests of balance for these 2,076 buyers at different levels of aggregation. The extensive-margin rows of Panel A speak to the first requirement. Among the 2,076 buyers, 59 percent of those in control units had also purchased in FOSC through FA during the pre-treatment period (February to May 2020), and 92 percent had purchased through some channel. Neither share differs significantly from its counterpart in the Public or Private arms. Entry into the analysis sample is therefore not differential by treatment status.

Then, in the intensive-margin rows, we restrict the analysis to buyers who purchased in FOSC through FA before treatment and examines their purchasing intensity. Control buyers issued 115 purchase orders on average, with an average overprice of 8.7 percent per unit purchased, and we generally do not reject the null hypothesis of no differences across experimental groups.[27] In Figure 6 we further plot the pre-treatment distribution of average overprice across arms; whether averaged at the buyer or the procurement-unit level, the distributions do not differ statistically as we do not reject the Kolmogorov–Smirnov test.

Because the model predicts that treatment effects vary with the price dispersion buyers face, we also verify that the purchasing environment itself is comparable across arms. The average between-price coefficient of variation in the reference groups faced by control buyers is 10.7 percent, statistically indistinguishable from the Public and Private arms, as is the average price ranking standard deviation. A final check asks whether the analysis sample draws disproportionately from

[27] Panel B in the same table tests for statistical balance on overprice using purchase-level data. The per-unit overprice across purchases made by buyers in the control, Public, and Private groups are well balanced, consistent with the balance found for the same outcome when aggregated at the buyer or procurement-unit level.

any part of the within-unit performance distribution. Ranking each buyer by average pre-treatment overprice within their procurement unit, control buyers sit at the 54th percentile on average, with 46 percent below the median and 21 percent in the bottom quartile; all three are balanced across arms.

Finally, Panel C aggregates pre-treatment outcomes to the procurement unit, our unit of randomization. Control units average 13.7 buyers purchasing in FOSC through FA. Units also contain buyers who purchased outside FOSC or through other channels (29 and 37.8 per unit on average, respectively), and these counts are likewise balanced. Among those who purchased in FOSC through FA, they collectively issued 2,273 purchase orders. The average overprice per unit purchased is 10.1 percent, and no difference across experimental groups is statistically significant.

Overall, individual comparisons are well balanced at every level of aggregation, and the omnibus tests of joint orthogonality generally work well. At the procurement-unit level the evidence is more limited, as the number of restrictions tested is large relative to the small sample of experimental units, a configuration in which the asymptotic F-test is known to over-reject. Indeed, the asymptotic joint test is rejected in two of the three cases at this level. Thus, to be on the safe side we estimate standard errors using the wild cluster bootstrap procedure proposed by Cameron, C., Gelbach, J., and Miller, D. (2008), which is better suited to this setting. Under this test, we never reject the joint null.

Beliefs. Because the mechanism we test operates through how persuadable buyers are, which in turn depends on what buyers and their managers believe about efficiency, internal validity also requires that these beliefs be comparable across experimental groups in the pre-treatment period. We report balance tests on beliefs elicited through online baseline surveys administered to buyers and managers in the 184 experimental procurement units between February and April 2020 (before the intervention began). From buyers we use the stated-preference items that later serve as mechanism outcomes: whether they would prioritize price over quality under hypothetical DIPRES audit rates of 1, 5, and 10 percent. From managers we use their perceptions of DIPRES (whether it values savings and monitors purchases) and their beliefs about how budget execution determines next year's allocation. Appendix Table B.7 gives a precise definition of each variable.

Buyers. Appendix Table B.3, Panel A, tests whether survey participation itself is balanced, i.e., who responded at baseline, and who responded at follow-up. Overall, 2,649 buyers across the 184 experimental procurement units responded at least one of the two rounds of surveys, who collectively represent 32% of the total population, i.e., a fairly representative sample of buyers in these units. Of them, 84% responded the baseline survey, 45% responded the follow-up survey, and 28% responded both, i.e., close to two-thirds of those who responded the follow up survey responded the baseline survey as well. For each subgroups of respondents, differences in response rates are small and generally well balanced across experimental arms, with joint orthogonality test being never rejected. The latter suggests no selective attrition by treatment status.

Panel B reports baseline stated preferences among buyers who responded at baseline, regardless of whether they responded the follow up survey or not. Half of control group buyers say they

would prioritize price over quality if DIPRES audited 1 percent of their purchases, augmenting to 55 percent at a 5 percent audit rate, and to 59 percent when the rate is set at 10 percent. The increasing gradient reflects greater concern about scrutiny as monitoring becomes stronger. The three items are mostly balanced across arms: we do not reject the null hypothesis of global balance, and this is the case either using asymptotic or wild cluster bootstrapped standard errors. Panel C repeats the exercise on the subsample who also responded at follow-up. The average responses are qualitatively very similar to those shown in Panel B, and they also follow a positive gradient as the audit rate increases, suggesting that those who self-selected to answer the follow up survey are not different from those who decided not to answer it. Moreover, among the two-rounds respondents, the statistical balance across experimental groups on the baseline stated preferences is, again, well preserved.

Managers. Appendix Table B.4 repeats the exercise at the manager level, where each of the 184 procurement units contributes with one respondent. Panel A reports baseline and follow-up response rates of 92 and 90 percent respectively in control units, and 82 percent answered both surveys. For each subgroups of respondents, we find no statistically meaningful differences across arms, suggesting attrition is not attributable to treatment assignment. Panel B reports baseline beliefs among managers who responded at baseline, regardless of whether they responded the follow up survey or not. Control group managers place DIPRES's valuation of savings at 4.02 on a 1–7 scale; they report that budget execution is a source of pressure (4.20 on a 1–5 scale) and that next year's budget depends on this year's execution (3.32 on a 1–5 scale). These levels are consistent with the organizational environment described in Section 2, in which efficiency is not a salient managerial priority. We reject the null hypothesis of no differences across experimental groups in just one out of 21 contrasts at the 5% of statistical level, which is what we would expect to occur by chance. Second, the joint orthogonality test suggests global balance across groups. Finally, Panel C restricts the analysis to the subsample of managers who responded to both waves. Again, the levels of responses are quite similar compared to those from Panel B sample, suggesting respondent and non-respondent of follow up survey are likely comparable. Moreover, baseline beliefs among managers who responded to both waves are generally well balanced as well, and this is reflected in that we generally do not reject the joint orthogonality tests at conventional levels of statistical significance.

6. Efficiency Results

We estimate the effect of the information interventions on purchasing efficiency using the following regression specification:

$$op_{icbjt} = \alpha + \delta_1 T_j^{Public} + \delta_2 T_j^{Private} + \phi_c + \eta_t + X_j'\beta + \varepsilon_{icbjt} \quad (6)$$

where op_{icbjt} is overprice paid per-unit for transaction i of product type c by buyer b who works in procurement unit j in calendar week t. Only products from FA purchase orders in Food, Office Supplies, or Computers (FOSC) are considered, as those are catalogs for which we can measure overprice. The T_j^{Public} and $T_j^{Private}$ are treatment indicators, which equal one if the buyer's procurement unit was assigned to either the Public or Private interventions, respectively, and zero otherwise. X_j includes the average overprice per unit purchased in unit j during the pre-intervention period and ε_{icbjt} is the error term. The models are estimated by OLS and the standard errors are clustered at the procurement unit level, the unit of randomization.

6.1. Overprice

Table 7 presents the results. Column (1) shows that public reports reduce overprice by 1.6 percentage points (*p-val*$<$0.05). The magnitude is economically significant: a 15% reduction relative to the control mean, translating into US$4.5 million over five months for the products under study. Extending the annualized effect to all Chilean public procurement yields approximately US$0.15 billion in savings, or 0.1% of GDP. This result is consistent with model prediction E.1 in that purchasing efficiency is increasing in performance information when buyers are extrinsically motivated.

In contrast, the private reports have no effect on overprice. The null effect of private reports rules out intrinsic motivation as the operative channel, rejecting model prediction N.1. Performance information alone is insufficient to change purchasing behavior among buyers. Performance information needs to be visible to managers so that buyers are accountable for performance monitoring to improve efficiency.

6.2. Alternative explanations

Are lower prices achieved by purchasing lower quality? Prior to listing on *Mercado Público*, all products and suppliers must pass a competitive bidding process administered by *ChileCompra* that guarantees minimum quality standards including delivery times across all items available on the Framework Agreement platform. Nonetheless, it is still possible that buyers respond to an intervention that promotes efficiency through purchasing price reduction by buying products that were cheaper because of lower quality. Reference groups combine products that are potential close substitutes based on their observable attributes. However, the reference groups are not made up of identical products offered by different suppliers. Rather they include all products with the same attributes and attributes that can be considers good substitutes. For example, `felt tip blue pens` might be in the same reference group as `ball point black pens`. In principle, then, the lower priced substitutes could be products with different (lower quality) attributes.

We test this by restricting the reference groups to be composed only of identical products (same SKU) offered by different suppliers, i.e., products for which we would expect quality differences within the reference group to be negligible. As shown in column (2) of Table 7, the results are

robust, suggesting that unobserved quality differences do not play a role in the effectiveness of the intervention.

Are lower prices achieved through bulk discounts? Buyers might achieve lower prices by purchasing in larger quantities to exploit bulk discounts rather than by searching more carefully. Column (3) of Table 7 rules this out: there is no significant effect on the number of units per transaction for either arm.

Are efficiency gains from more effort per transaction offset by buyers processing fewer purchase orders? Searching more carefully takes time, and buyers facing a fixed time budget might respond by issuing fewer orders. Since overprice is measured per unit purchased, this would leave our estimates intact while imposing a real cost elsewhere, e.g., delayed or unmet purchasing needs that our outcome does not capture. Likewise, buyers might issue fewer orders in the monitored categories, changing the composition of transactions on which overprice is measured rather than the prices paid within them. Columns (3) and (4) show no effect on the number of purchase orders executed per buyer for either treatment arm, and this is the case considering FOCs products only (for which DIPRES can calculate overprice) or any product offered in the FA platform.

Is there selection bias from choice of purchase channel? Buyers may have responded to monitoring by shifting purchases away from the FA platform (where their behavior is tracked) toward less transparent mechanisms such as procurement auctions or direct purchases, which would introduce sample selection bias into our estimates. Column (6) of Table 7 shows no effect of either treatment on the share of purchase orders made through the FA channel, ruling out this possibility. The absence of channel switching also speaks to the nature of the efficiency gains: since active waste such as corruption is considerably less prevalent in FA purchases than in auctions or direct purchases (Bandiera et al., 2009), the fact that buyers did not migrate toward less regulated channels suggests that the intervention reduced passive waste without inadvertently encouraging active waste.

Hawthorne Effects. Buyers may improve efficiency just because they know they are being monitored, even if no evidence of purchasing behavior is shown to them. We find this is not sufficient to trigger a change in purchasing behavior. First, both treatment arms received the same 10-minute training video two months before the performance reports began. Restricting the analysis to the post-training pre-reports period yields no significant effect on overspending for either arm (Appendix Table B.5), confirming that the efficiency gains are driven by the performance reports rather than the training. Second, placebo buyers in treated units who received monthly monitoring notifications but no performance information show no reduction in overspending (Appendix Table B.6), ruling out Hawthorne effects as an explanation.

6.3. Heterogeneity

Buyer Responsiveness to Manager's Persuasive Actions. We first test the prediction that the effect of the information intervention on efficiency for extrinsically motivated buyers only impacts the buyers who are sufficiently responsive to manager oversight actions (prediction E.1). Specifically, the purchase efficiency $\pi(I)$ is increasing in $I > 0$ for buyers where $\theta_B > \bar{\theta}$ and zero for $\theta_B < \bar{\theta}$.[28] We test these predictions by adding to the above model buyer responsiveness measured at baseline and an interaction of the treatment dummies with buyer responsiveness. We measure buyers' responsiveness based on answers to hypothetical questions of their willingness to prioritize price over quality under three hypothetical audit scenarios: if DIPRES were to audit 10%, 5%, or 1% of their purchases.[29]

Table 12 shows the results. Column 1 reports the results assuming an audit rate of 1%, column 2 for an audit rate of 5% and column 3 for 10%. The interactions of treatment and persuadability are negative and statically significant for the public intervention in all 3 models. The size of the coefficient is roughly double the estimated treatment effect without the interactions. Moreover, the main effect for the public intervention becomes small and not statistically significant, implying that all of the effect comes from responsive buyers. This is consistent with (prediction E.6) that the purchase efficiency $\pi(I)$ is increasing in $I > 0$ for buyers where $\theta_B > \bar{\theta}$ and zero for $\theta_B < \bar{\theta}$. The interactions have no impact in the private intervention consistent with buyers being extrinsically and not intrinsically motivated.

Price Dispersion. We now test the predictions that there is less efficiency for product whose reference groups have with greater price dispersion (predictions E.4 N.4), and whether the effect of the information intervention is higher for products whose reference groups have more price dispersion (predictions E.5 and N.5). We test these predictions by adding to the above model price variation in the product's reference group and an interaction of the treatment dummies with price variation. We use two measures of price dispersion. The first is the standard deviation of daily price rankings over time (*Rank Std*). It captures how often the cheapest option changes. The second is the cross-sectional coefficient of variation of posted prices (*Between Price CV*). It captures how spread out prices are at any given point.[30] Both measures of price dispersion are constructed using data from the pre-intervention period.

Table 8 shows the results. Both measures are positively associated with overspending in the control group, consistent with the model's prediction that absent treatment, higher σ is associated with higher overprice as buyers do not exert the search effort needed to exploit available price

[28] In the case of intrinsically motivated buyers efficiency is predicted to be strictly increasing in information regardless buyer responsiveness (prediction N.1). However, we already has evidence that buyers are extrinsically and not intrinsically motivated, so we expect a null effect in the private information treatment.

[29] The exact question was: *"If DIPRES audits X% of your purchases, would you prioritize price over quality?"* Responses are coded as a binary dummy equal to 1 if the buyer answered yes.

[30] *Rank Std* is computed by ranking products by price on a daily basis and taking the standard deviation of each product's rank over time, averaged within reference groups. *Between Price CV* is the coefficient of variation of average prices across products within a reference group. Both measures are described in Section **??**.

differences. The treatment interactions discriminate sharply between the extrinsic and intrinsic channels. The public treatment effect is significantly larger in absolute value when *Rank Std* is higher (-0.007^{**} to -0.008^{***}), consistent with the prediction that higher σ amplifies the returns to manager-induced search: where price rankings shift frequently, efficiency requires sustained effort that only extrinsic accountability can reliably activate. The interaction with *Between Price CV* is precisely estimated at zero for both arms. This distinguishes the extrinsic from the intrinsic channel: static cross-sectional dispersion creates a one-time search opportunity that does not require ongoing managerial accountability to exploit, and accordingly does not amplify the treatment effect. Private reports interact with neither measure, ruling out the intrinsic channel for the σ prediction as well.

Buyer Performance Gap. Finally, we test the predictions that the information treatment has larger effects on efficiency for extrinsically buyers who have the most to gain from more search effort. These are the buyers who are the least efficient before the the performance monitoring system is introduced (predictions E.6 and E.7). We test these predictions by adding to the above model the buyer's baseline average overprice and an interaction of the treatment dummies with baseline overprice. We use two measures of the overprice. The first is an indicator of whether the buyer is in the bottom quartile of the baseline overprice distribution, the buyers with the smallest gaps. The second is whether the buyer is in the bottom half. We expect the signs on the interactions to be positive, i.e., the effect of the information treatment to be smaller for the better performing buyers with the smaller price gaps.

Table 12 reports the results. We find that the buyers with smaller baseline price gaps are more efficient (have lower overprice) than buyers who have larger price gaps during the experimental period. Specifically, the buyers in the bottom quartile of the baseline price gap distribution have a 2.7 pp lower price gap than those with higher baseline price gaps. Moreover, there is no effect of the public intervention on buyers in lowest 25% baseline price gap distribution compared to a -1.4 pp impact on buyers with higher baseline price gaps. Not surprisingly, there is no effect of the private information arm.

7. Mechanism Results

The model predicts that performance information improves efficiency by raising the manager's information I about individual buyer types, which increases the optimal intensity of persuasion actions $a^*(I, \theta_B)$. The experimental design moves I in two distinct ways. Private reports give managers aggregate unit-level performance only, leaving them unable to identify or target individual buyers, while buyers are explicitly told their report has not been shared, leaving their responsiveness to manager actions θ_B unchanged. Public reports raise I substantially by providing individual buyer rankings, and simultaneously increase effective θ_B by informing buyers that their manager can see their individual performance. The public treatment therefore operates on both sides of $P(\theta_B, a) = \theta_B \phi(a)$: it gives managers the information to act and gives buyers the incentive to respond. Yet all this is conditional on a first stage result: that performance information

I was effectively received by managers and buyers. The sections that follow trace the evidence for each step.

7.1. Information

Table 9 shows treatment effects on information outcomes for both managers and buyers measured in the follow-up survey of February 2021. Both treatment arms substantially increased manager awareness of their unit's overspending. Column (1) shows effects of 60.5 and 66.5 percentage points for Public and Private respectively, with no significant difference between the two arms. Column (2) shows that perceived DIPRES monitoring of purchases increased by approximately one point on the 1–7 scale for both arms equally, consistent with both groups having received reports signed by DIPRES.[31] Column (3) shows that Public managers were 24.2 percentage points more likely to believe DIPRES was specifically aware of their unit's overspending, while Private managers showed no significant change. This should not be surprising as the specificity of the information received by public managers was much grater compared to that of private counterparts, which naturally increased their credibility.

The reports also increased buyer awareness of overspending. Columns (4)–(5) show that public buyers were 23 percentage points more likely to know their own overspending and 27 percentage points more likely to know their unit's overall overspending, while private buyers show effects of 14 and 15 percentage points respectively. Differences in the order of magnitude of the effects across treatment groups are substantial: Public group buyers became significantly more aware than Private group buyers on both individual and unit-level outcomes, consistent with managers in the public arm actively discussing performance with individual buyers and their teams. Finally, both arms equally increased the share of buyers who believe DIPRES is aware of their unit's overspending (column 6). Crucially, private buyers became more informed about their inefficiency yet did not reduce overspending — ruling out a pure information channel. Awareness of inefficiency alone is not sufficient to change behavior without the individual accountability that only public reports create.

7.2. Manager Oversite and Persuasive Actions

The model predicts that the intensity of manager oversight actions $a^*(I, \theta_B)$ is increasing in I when $\theta_B > \bar{\theta}$. By design, public reports raised I at the individual buyer level while private reports did not. The testable prediction is that public managers should show higher a^* than private managers despite both arms receiving the same aggregate information. Table 10 tests this prediction.

The managerial action outcomes are measured through the follow-up survey of buyers, who were asked to report on the oversight and communication activities undertaken by their manager during the intervention period, including whether the manager set explicit efficiency goals,

[31] Managers were asked: *"From 1 to 7, what do you think is the level of monitoring exercised by DIPRES on public purchases made by your unit?"* where 1 means *No monitoring at all* and 7 means *High level of monitoring*.

held regular performance meetings, provided training on reducing overspending, and offered performance-based rewards, among other activities (ten in total). Column (1) reports the Anderson summary index of all ten outcomes.[32]

Public reports increased the summary index by 0.264 standard deviations, while private reports had essentially no effect, with the difference between arms highly significant. The five individual outcomes showing the most notable differences favoring the public arm are that the manager sets efficiency goals, has high levels of communication with buyers, he train buyers on reducing overspending, promotes coordinated purchases, and implement performance based rewards. In contrast, despite being equally informed about aggregate overspending and equally aware of DIPRES monitoring, Private managers had no individual-level information on which to act and thus behaved indistinguishably from the control group. The contrast maps directly onto the model: public reports raised I to a level where managers could estimate individual buyer types and optimally adjust $a^*(I, \theta_B)$, while private reports left I too low to enable individual targeting.

7.3. Buyer Responses

We now examine whether the increase in manager actions $a^*(I, \theta_B)$ translated into changes in Buyer's persuadability towards efficiency, $P(\theta_B, a)$. Table 11 shows treatment effects on buyers' stated willingness to prioritize price over quality under three hypothetical audit scenarios: if DIPRES were to audit 10%, 5%, or 1% of their purchases.[33]

The control group means follow an intuitive pattern: fewer buyers say they would prioritize price as the hypothetical audit rate falls, reflecting lower concern about external scrutiny when monitoring is sparse. At a 10% audit rate neither arm shows a significant effect. That is, when external enforcement is already high, performance reports add little to existing incentives. As the audit rate falls to 5% and then to 1%, effects emerge exclusively for public buyers, reaching 15% and nearly 30% of the control group mean respectively.

In contrast, Private buyers show no change at any audit rate. The fact that private buyers, who became equally informed about their inefficiency, show no change at any audit rate confirms that it is individual accountability, not information alone, that changes buyer's persuadability towards efficiency.

8. Conclusion

Passive waste in public procurement — buyers paying more than necessary not out of corruption but out of insufficient effort or misaligned incentives — is large, costly, and difficult to address without changing the fundamental architecture of civil service rules. This paper shows that a

[32] Buyers rather than managers are the source of these outcomes to avoid social desirability bias in manager self-reports. All outcomes are binary dummies. The Anderson index is an equally-weighted average of z-scores across the ten outcomes in columns (2)–(11), expressed in control-group standard deviation units (Anderson, 2008).

[33] The exact question was: *"If DIPRES audits X% of your purchases, would you prioritize price over quality?"* Responses are coded as a binary dummy equal to 1 if the buyer answered yes.

low-cost information intervention can achieve meaningful reductions in overspending without modifying compensation structures, organizational hierarchies, or procurement regulations.

In collaboration with Chile's Public Budget Office and Public Procurement Office, we randomized monthly individual performance reports across 184 procurement units and over 3,500 buyers. Public reports, which gave managers access to buyer-level overspending data and made buyers aware that their manager could see their individual performance, reduced overspending by approximately 15% of the control group mean. Private reports, which provided managers only with aggregate unit-level performance and kept individual buyer data confidential, had no effect. The contrast between arms points clearly to *extrinsic* rather than *intrinsic* motivation as the operative channel: buyers changed their behavior not because they internalized efficiency as a goal, but because their managers could now identify and act on individual performance.

The mechanism evidence supports this interpretation at every step. Managers in the Public arm became more likely to set efficiency goals, hold performance discussions with buyers, provide training, and promote coordinated purchases — none of which occurred in the Private arm despite managers in both arms becoming equally informed about aggregate unit overspending. On the buyer side, Public reports shifted stated preferences for price over quality, and did so most strongly at low hypothetical audit rates where external enforcement is weakest. At the product level, treatment effects on actual overspending are largest in categories where price rankings shift frequently, requiring sustained search effort rather than a one-time investment — precisely the conditions where manager oversight adds the most value.

The intervention is attractive from a policy standpoint for three reasons. First, the marginal cost of scaling it up is virtually zero: the underlying performance data are already embedded in the *Mercado Público* platform and report generation is fully automated. Second, it operates without modifying civil service rules, financial incentives, or procurement regulations, avoiding the political economy frictions that have limited the reach of more ambitious reform efforts. Third, pre-registered expert forecasts closely match the experimental results, suggesting that the mechanism is intuitive and the findings are unlikely to be specific to our setting. Extending the annualized treatment effect to all transactions in the Chilean public procurement system would yield approximately US$0.15 billion in annual savings — equivalent to 0.1% of GDP — at essentially no additional cost.

References

ALBANO, G. L. AND C. NICHOLAS (2016): *The Law and Economics of Framework Agreements*, Cambridge University Press.

ANDERSON, M. L. (2008): "Multiple Inference and Gender Differences in the Effects of Early Intervention: A Reevaluation of the Abecedarian, Perry Preschool, and Early Training Projects," *Journal of the American Statistical Association*, 103, 1481–1495.

BAKER, G. AND T. HUBBARD (2003): "Make Versus Buy in Trucking: Asset Ownership, Job Design and Information," *The American Economic Review*, 93, 551–572.

——— (2004): "Contractability and Asset Ownership: On-Board Computers and Governance in U.S. Trucking," *The Quarterly Journal of Economics*, 119, 1443–1479.

BALÁN, P., A. BERGERON, G. TOUREK, AND J. WEIGEL (2022): "Local Elites as State Capacity: How City Chiefs Use Local Information to Increase Tax Compliance in the Democratic Republic of the Congo," *American Economic Review*, 112, 762–97.

BANDIERA, O., B. I. AND I. RASUL (2011): "Field Experiments With Firms," *Journal of Economic Perspectives*, 25, 63–82.

——— (2013): "Team Incentives: Evidence From a Firm Level Experiment," *Journal of the European Economic Association*, 11, 1079–1114.

BANDIERA, O., M. C. BEST, A. Q. KHAN, AND A. PRAT (2021a): "The allocation of authority in organizations: A field experiment with bureaucrats," *The Quarterly Journal of Economics*, 136, 2195–2242.

BANDIERA, O., E. BOSIO, AND G. SPAGNOLO (2021b): "Procurement in Focus: Rules, Discretion, and Emergencies," *CEPR Press*.

BANDIERA, O., A. PRAT, AND T. VALLETTI (2009): "Active and Passive Waste in Government Spending: Evidence from a Policy Experiment," *American Economic Review*, 99, 1278–1308.

BÉNABOU, R. AND J. TIROLE (2003): "Intrinsic and Extrinsic Motivation," *The Review of Economic Studies*, 70, 489–520.

BEST, M. C., J. HJORT, AND D. SZAKONYI (2023): "Individuals and Organizations as Sources of State Effectiveness," *American Economic Review*, 113, 2121–2167.

BLOOM, N., E. BRYNJOLFSSON, L. FOSTER, R. JARMIN, M. PATNAIK, I. SAPORTA-EKSTEN, AND J. VAN REENEN (2019): "What Drives Differences in Management Practices?" *American Economic Review*, 109, 1648–1683.

BLOOM, N. AND J. VAN REENEN (2007): "Measuring and Explaining Management Practices Across Firms and Countries," *Quarterly Journal of Economics*, 122, 1351–1408.

BOSIO, E., S. DJANKOV, E. GLAESER, AND A. SHLEIFER (2021): "Public Procurement in Law and Practice," *American Economic Review*, Forthcoming.

BULOW, J. AND P. KLEMPERER (1996): "Auctions Versus Negotiations," *American Economic Review*, 86, 180–194.

CAMERON, C., GELBACH, J., AND MILLER, D. (2008): "Bootstrap-Based Improvements for Inference with Clustered Errors," *The Review of Economics and Statistics*, 90, 414–427.

CARPINETI, L., G. PIGA, AND M. ZANZA (2006): "The variety of procurement practice: evidence from public procurement," .

CARRIL, R. (2019): "Rules versus Discretion in Public Procurement," *Job Market Paper*.

CARRIL, R., A. GONZALEZ-LIRA, AND M. WALKER (2022): "Competition under Incomplete Contracts and the Design of Procurement Policies," *BSE Working Paper #1327*.

COVIELLO, D., A. GUGLIELMO, AND G. SPAGNOLO (2018): "The Effect of Discretion on Procurement Performance," *Management Science*, 64, 715–38.

DAL BÓ, E., F. FINAN, N. LI, AND L. SCHECHTER (2021): "Information technology and government decentralization: experimental evidence from Paraguay," *Econometrica*, 89, 677–701.

DE ROCHAMBEAU, G. (2021): "Monitoring and Intrinsic Motivation: Evidence from Liberia's Trucking Firms," *mimeo*.

DECAROLIS, F., R. FISMAN, P. PINOTTI, AND S. VANNUTELLI (2020): "Rules, Discretion, and Corruption in Procurement: Evidence from Italian Government Contracting," *CEPR Discussion Paper No. 14794*.

DIMITRI, N., G. PIGA, AND G. SPAGNOLO (2006): *Handbook of procurement*, Cambridge University Press.

DINERSTEIN, M., L. EINAV, J. LEVIN, AND N. SUNDARESAN (2018): "Consumer price search and platform design in internet commerce," *American Economic Review*, 108, 1820–59.

ENGEL, E., F. JORDAN, T. RAU, AND A. REPETTO (2021): "Audit Threats and Wasteful Spending by Government Agencies: Experimental Evidence from Chile," *mimeo*.

GERARDINO, M., S. LITSCHIG, AND D. POMERANZ (2024): "Distortion by Audit: Evidence from Public Procurement," *American Economic Journal: Applied Economics*, 16, 71–108.

HALAC, M. AND A. PRAT (2016): "Managerial Attention and Worker Performance," *American Economic Review*, 106, 3104–32.

HUBBARD, T. (2000): "The Demand for Monitoring Technologies: The Case of Trucking," *The Quarterly Journal of Economics*, 115, 533–560.

——— (2003): "Information, Decisions, and Productivity: OnBoard Computers and Capacity Utilization in Trucking," *The American Economic Review*, 93, 1328–1353.

KELLEY, E., G. LANE, AND D. SCHÖNHOLZER (2021): "Monitoring in Small Firms: Experimental Evidence from Kenyan Public Transit," *mimeo*.

KELMAN, S. (1990): "Procurement and Public Management: The Fear of Discretion and the Quality of Government Performance," *AEI Press*.

——— (2005): "Unleashing change: A study of organizational renewal in government," *Washington, D.C.: Brookings Institution*.

KHAN, A., A. KHWAJA, AND B. OLKEN (2016): "Tax Farming Redux: Experimental Evidence on Performance Pay for Tax Collectors," *Quarterly Journal of Economics*, 131, 219–71.

——— (2019): "Making Moves Matter: Experimental Evidence on Incentivizing Bureaucrats through Performance-Based Postings," *American Economic Review*, 109, 237–70.

KHANDELWAL, A., H. MITTAL, S. KULKARNI, AND D. GUPTA (2023): "Large scale generative multimodal attribute extraction for e-commerce attributes," .

LEWIS-FAUPEL, S., Y. NEGGERS, B. OLKEN, AND R. PANDE (2016): "Can Electronic Procurement Improve Infrastructure Provision? Evidence from Public Works in India and Indonesia," *American Economic Journal: Economic Policy*, 8, 258–83.

LIEBMAN, J. AND N. MAHONEY (2017): "Do Expiring Budgets Lead to Wasteful Year-End Spending? Evidence from Federal Procurement," *American Economic Review*, 107, 3510–49.

LISCOW, Z., , W. NOBER, AND C. SLATTERY (2023): "Procurement and Infrastructure Costs," *Available at SSRN: https://papers.ssrn.com/sol3/papers.cfm?abstract$_i$d* $= 4522676$.

MACKAY, A. (2022): "Contract Duration and the Costs of Market Transactions," *American Economic Journal: Microeconomics*, 14, 164–212.

MATTSSON, M. (2022): "Information Systems, Service Delivery, and Corruption: Evidence from the Bangladesh Civil Service," *mimeo*.

MIKOLOV, T., I. SUTSKEVER, K. CHEN, G. S. CORRADO, AND J. DEAN (2013): "Distributed representations of words and phrases and their compositionality," *Advances in neural information processing systems*, 26.

OECD (2009): "Principles for Integrity in Public Procurement," *Organisation for Economic Co-operation and Development, Tech. Report*.

——— (2013): "Monitoring of Public Procurement," Tech. rep., Available at https://www.sigmaweb.org/publications/Brief27_MonitoringofPP_2013.pdf.

——— (2019): *Reforming Public Procurement: Progress in Implementing the 2015 OECD Recommendation*, OECD Publishing.

OECD (2023): "Public Procurement Performance: A Framework for Measuring Efficiency, Compliance and Strategic Goals," Tech. Rep. 36, OECD Publishing.

OLIVARES, M., D. SABAN, G. Y. WEINTRAUB, E. LARA, P. ZANOCCO, AND P. MORENO (2025): "Saving millions in government procurement through data science and market design," *INFORMS Journal on Applied Analytics*, 55, 101–120.

PALGUTA, J. AND F. PERTOLD (2017): "Manipulation of Procurement Contracts: Evidence from the Introduction of Discretionary Thresholds," *American Economic Journal: Economic Policy*, 9, 1–24.

RASUL, I. AND D. ROGERS (2018): "Management of Bureaucrats and Public Service Delivery: Evidence from the Nigerian Civil Service," *Economic Journal*, 128, 413–446.

SABAN, D. AND G. WEINTRAUB (2021): "Procurement mechanisms for assortments of differentiated products," *Operations Research*, 69, 795–820.

SPAGNOLO, G. (2012): "Reputation, Competition and Entry in Procurement," *International Journal of Industrial Organization*, 30, 291–296.

SUN, C., N. RAMPALLI, F. YANG, AND A. DOAN (2014): "Chimera: Large-scale classification using machine learning, rules, and crowdsourcing," *Proceedings of the VLDB Endowment*, 7, 1529–1540.

SZUCS, F. (2023): "Discretion and Favoritism in Public Procurement," *Journal of the European Economic Association*, 0, 1–44.

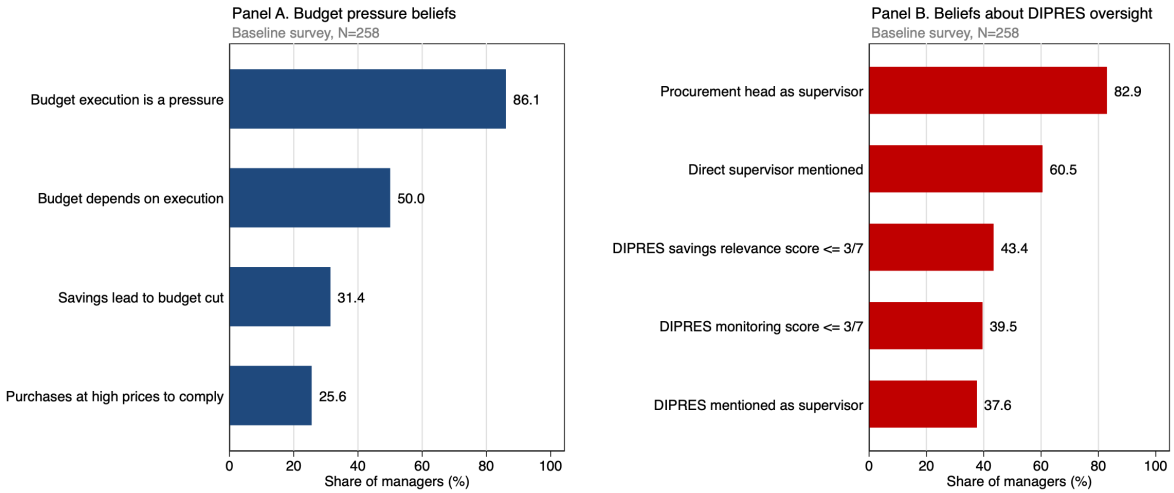
WATHNE, K. H., H. BIONG, AND J. B. HEIDE (2001): "Choice of Supplier in Embedded Markets: Relationship and Marketing Program Effects," *Journal of Marketing*, 65, 54–66.

WEITZMAN, M. L. (1979): "Optimal Search for the Best Alternative," *Econometrica*, 47, 641–654.

WORLD BANK (2022): "A Global Procurement Partnership for Sustainable Development: An International Stocktaking of Developments in Public Procurement," Tech. rep., World Bank.

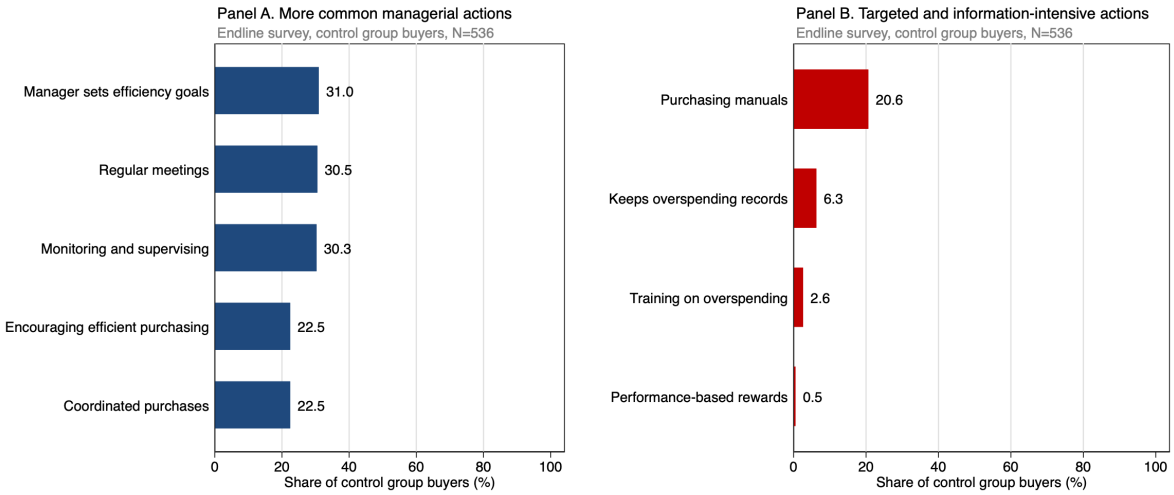
Figures and Tables

Figure 1: Managers' Beliefs About Budget Incentives and DIPRES Oversight



Notes: Panel A shows the share of managers agreeing or strongly agreeing (scores 4–5 on a 1–5 scale) with each statement, sorted in descending order. Panel B shows the share of managers placing their score at 3 or below on a 1–7 scale (low perceived relevance or monitoring) and the share citing each actor as a procurement supervisor (Q19, multiple choice), sorted in descending order. Both panels use the baseline survey of managers (Jefes de Servicio and DAFs), N=258.

Figure 2: Managerial Oversight of Buyer Purchasing Behavior



Notes: Both panels use the endline survey restricted to the control group (units that received no intervention), N=536 buyers across 58 services. Each bar shows the share of buyers reporting that their manager engaged in the corresponding action. Panel A shows generic oversight behaviors. Panel B shows targeted, information-intensive actions that require knowledge of individual buyer performance. Variables correspond to the outcomes in Table **??**.

Figure 3: Example of grated cheese products displayed in the online catalogue.

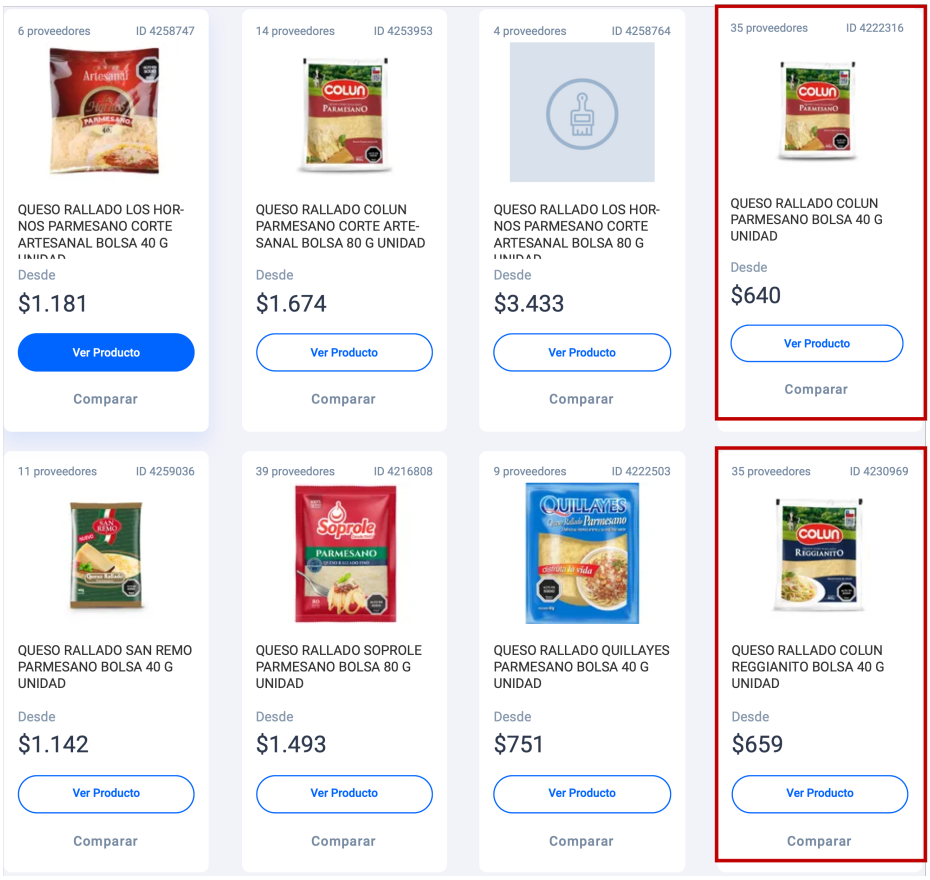


Figure 4: Example of supplier-level price comparison for a selected product and delivery region.



Table 1: Examples of product types and their attributes.

Product Type	Category	Categorical Attributes	Numerical Attributes
Desktop	Computers FA	Brand, Processor, Operating System, Hard Drive type	RAM, Hard Drive Capacity, Screen size
Pencil	Office Supplies FA	Brand, Size, Color	Units per package, Weight
Batteries	Office Supplies FA	Type, Size, Rechargeable	Units per pkg, Voltage
Instant Coffee	Food FA	Brand, Decaf, Format, Units per pkg	Weight
Grated Cheese	Food FA	Type, Brand, Flavor	Weight, Units

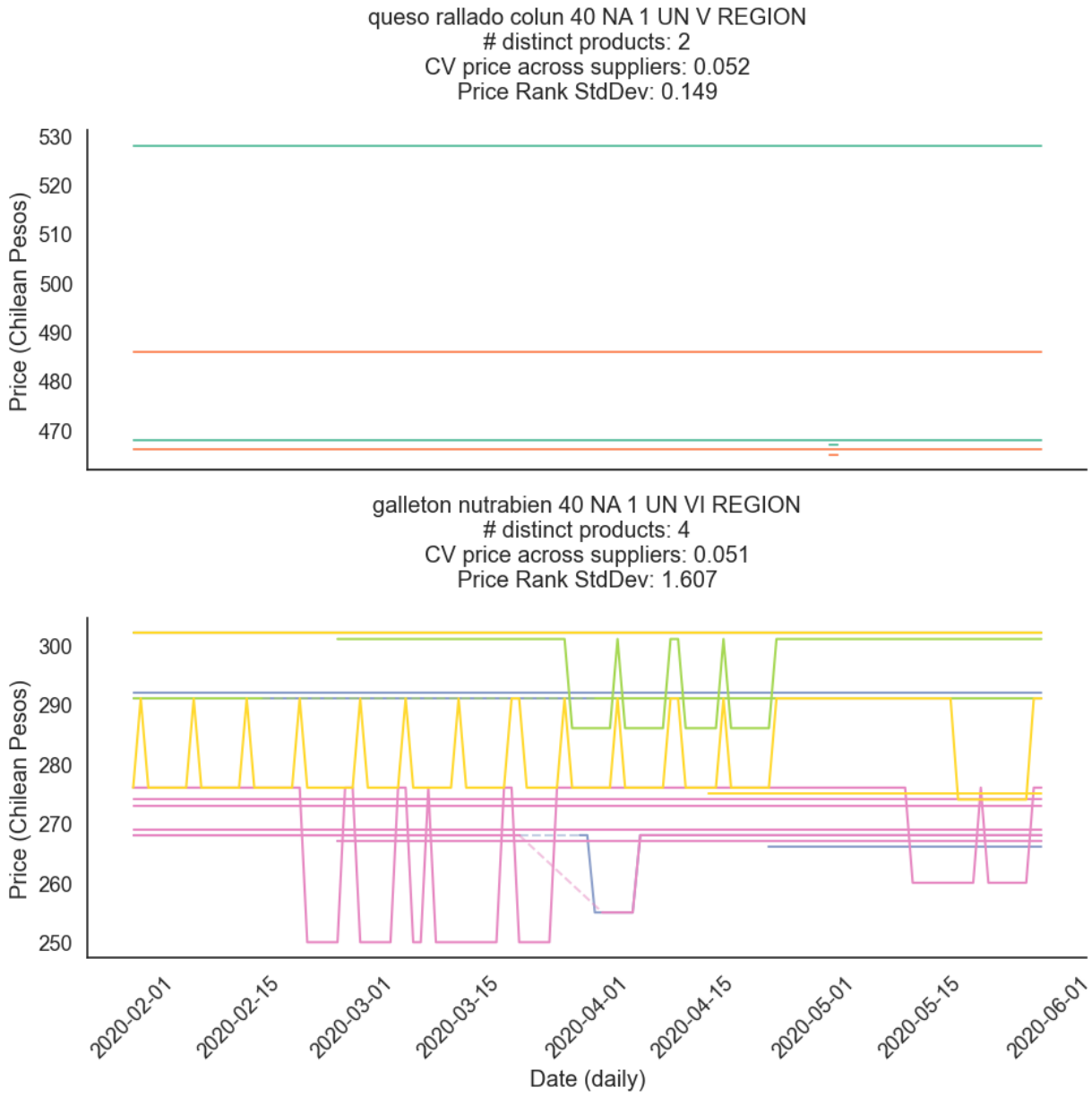
Notes: The table shows examples of products across the three FA categories under consideration: Computers, Office Supplies, and Food. Categorical attributes sort a product into one of a finite set of labels (e.g., brand, color, material), most of which carry no natural ordering. Numerical attributes express a characteristic as a quantity on a common scale (e.g., storage capacity, screen size, package weight) so that any two products can be ranked along it.

Table 2: Example of attribute extraction from free-text product descriptions in the Food Framework Agreement for Grated Cheese product type

Prod.Id	text description	prod.type	brand	flavor	weight	units	avg. price
1456879	QUESO RALLADO COLUN PARMESANO BOLSA 40 G UNIDAD	queso rallado	colun	parmesano	40gr	1	497.75
1456909	QUESO RALLADO COLUN REGGIANITO BOLSA 40 G UNIDAD	queso rallado	colun	reggianito	40gr	1	475.91

Notes: The table shows the raw description and the corresponding tagged attributes (brand, flavor, weight and number of units). The last column shows the average transaction price for each product during the baseline period.

Figure 5: Posted price dynamics within attribute-based reference groups (baseline period).



Notes: Each graph corresponds to one such group: a cookie snack sold in Region V (bottom) and the grated cheese example sold in Region V (top). Within each graph, lines plot the daily posted unit price for each product available in the group (offered by one supplier) during the baseline period (Jan. 2020–May 2020). Products that are identical (same Product ID) are displayed in the same color. If during some periods the product is not available, the series is connected with a dashed line. The subtitle in each graph displays: (i) the number of distinct products in the reference group, (ii) *CV Price across products*, defined as the coefficient of variation of posted prices across suppliers within the reference group (standard deviation divided by mean), and (iii) *Price Ranking StdDev*, defined as the standard deviation of each supplier's daily price rank within the group, taking the average of this metric across all suppliers in the group.

Table 3: Reference group properties under alternative definitions

	Num. Suppliers			CV Price			Price Rank. StdDev			Overprice		
Reference Group	Mean	p10	p90	Mean	p10	p90	Mean	p10	p90	Mean	p10	p90
Identical Product	8.2	1.0	17.0	0.08	0.03	0.13	0.59	0.00	1.33	0.08	0.00	0.21
Same Attributes	12.2	1.0	24.0	0.11	0.03	0.23	0.71	0.00	1.52	0.21	0.00	0.33
Mixed (Attrib./Identical)	11.7	1.0	23.0	0.11	0.03	0.20	0.70	0.00	1.49	0.11	0.00	0.27
Same Product Type	393.2	2.0	763.6	0.64	0.03	1.44	8.22	0.00	18.23	11.05	0.09	21.55

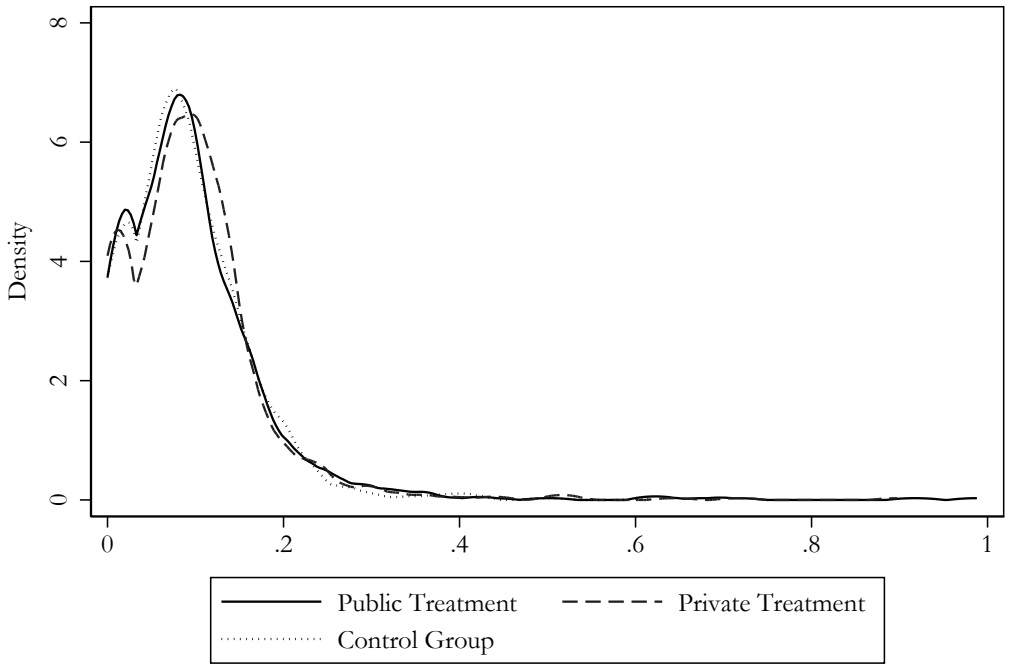
Notes: Sample of analysis are Framework Agreement product transactions from Computers, Office Supplies, and Food categories issues between January 2020 and May 2020. "Num. Suppliers" is the number of distinct suppliers offering a product in a reference group. "CV Price" summarizes cross-sectional dispersion in posted prices across products within group-day (aggregated over baseline days). "Price Rank. StdDev" summarizes within-product variability of the product's daily price rank within its reference group (aggregated within group). "Overprice" is calculated using the corresponding reference group definition for each transaction. For the statistics of Num. Suppliers, CV Price and SD Price Rank, the unit of observation is the reference group. For statistics of Overprice, the unit of observation is the product. Following equation 1, overprice is defined per unit of product, as the difference between the per-unit price paid for the product and the lowest per-unit price in that day's price distribution among products in the same reference group, scaled by the latter.

Table 4: Summary of Interventions

	Public Intervention Units	Private Intervention Units	Control Units
Buyers receive:	• 10-minutes Training Video • Buyer Performance Report • Non-disclosure message	• 10-minutes Training Video • Buyer Performance Report • Disclosure message	Nothing
Manager receives:	• Unit-level Performance Report • Buyer Performance Reports • Buyers' ranking list	• Unit-level Performance Report	Nothing
Placebo Buyers receive:	• Monitoring message	• Monitoring message	

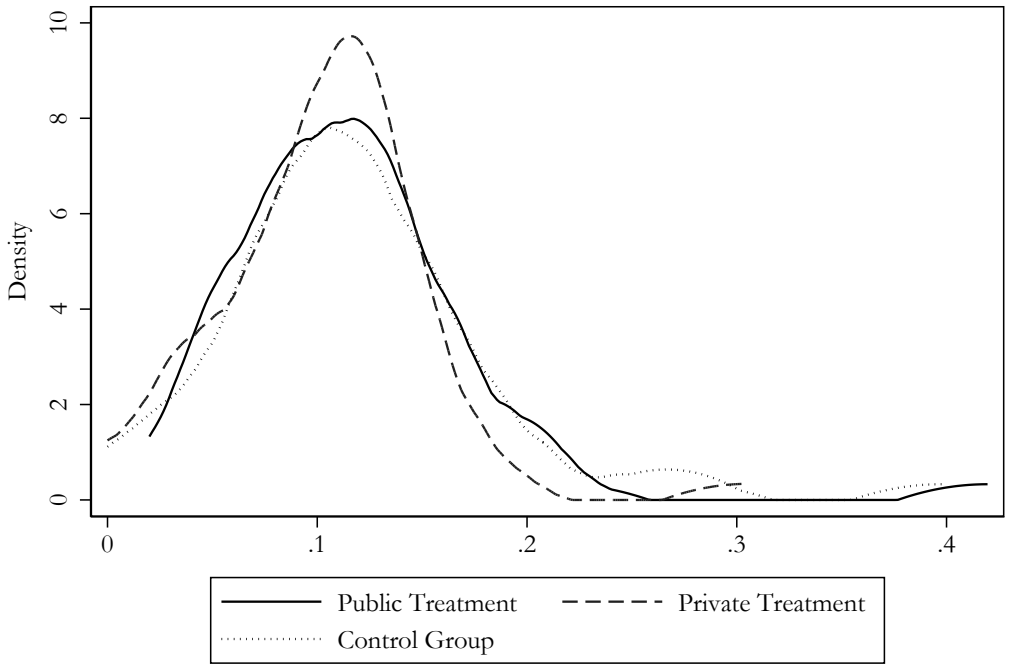
Figure 6: Baseline Distribution of Average Overprice per unit purchased, by Treatment Arm

(a) Buyer Level



Public - Private (p-value) = .211 Public - Control (p-value) = .956 Private - Control (p-value) = .225

(b) Unit Level



Public - Private (p-value) = .395 Public - Control (p-value) = .998 Private - Control (p-value) = .398

Note: Distribution of average overprice per unit purchased in the pre-intervention period (February–May 2020), by treatment group. The top panel shows the distribution for overprice averaged at the buyer level, while the bottom panel shows the distribution for overprice averaged at the unit level. Observations with overprice above the 99^{th} percentile are excluded.

Table 5: Experimental Design

	Public Intervention	Private Intervention	Control	Total
Randomization Unit: Procurement Units (Managers)	61	62	61	184
Full Sample — February 2020 through January 2021				
FA Buyers	1,052	1,127	1,091	3,270
FA Transactions	54,065	69,102	69,361	192,528
Experimental FA Buyers	778	829	1,091	2,698
Experimental FA Transactions	40,522	49,565	69,361	159,448
Placebo FA Buyers	274	298		572
Placebo FA Transactions	13,543	19,537		33,080
Pre-Treatment Sample — February 2020 through May 2021				
FA Buyers	776	817	739	2,332
FA Transactions	24,589	29,326	26,127	80,042
Experimental FA Buyers	570	594	739	1,903
Experimental FA Transactions	18,297	21,232	26,127	65,656
Placebo FA Buyers	196	223		419
Placebo FA Transactions	6,292	8,094		14,386
Post-Treatment Sample — July 2020 through January 2021				
FA Buyers	776	869	858	2,503
FA Transactions	29,476	39,776	43,234	112,486
Experimental FA Buyers	573	645	858	2,076
Experimental FA Transactions	22,225	28,333	43,234	93,792
Placebo FA Buyers	203	224		427
Placebo FA Transactions	7,251	11,443		18,694

Notes: Sample size by experimental group. Randomization is at the procurement unit level, stratified by Ministry and by whether the level of overspending of the unit was above or below the median in the previous fiscal year. Each procurement unit is randomly assigned to one of three experimental arms: Public, Private, or Control. Procurement units assigned to the Control arm received no intervention. Full sample cover February 2020 through January 2021, excluding June 2020 (video training month). Pre-treatment period cover February 2020 through May 2020. Post-treatment period cover July 2020 through January 2021. A procurement unit is directed by one Manager who lead a team of multiple buyers, the latter being in charge of issuing purchase orders. We count only buyers who specialize in Food, Office Supplies, or Computers FAs, the categories for which *ChileCompra* can calculate overspending from their transactions, and thus produce performance reports. Public Treatment: Experimental FA Buyers receive an individual report on his own overspending (Appendix Figure B.1) which carries the message "Your individual report *has been sent* to the Unit's Manager." The manager receives the Unit-level report (Appendix Figure B.3), the individual buyer-level reports (Appendix Figure B.1), and a list ranking each buyer by overspending and by his contribution to the unit total (Appendix Figure B.4). Private Treatment: Experimental FA Buyers receives an individual report on his own overspending (Appendix Figure B.1) which carries the message "Your individual report *has NOT been sent* to the Unit's Manager." The manager receives the Unit-level report (Appendix Figure B.3), and nothing more. Randomization of Placebo FA Buyers is at the buyer level within procurement units. FA Placebo Buyers in both public and control group units receive the same email message indicating that his overspending is being monitored, but do not receive Buyer- or Unit-level reports.

Figure 7: Experiment Timeline

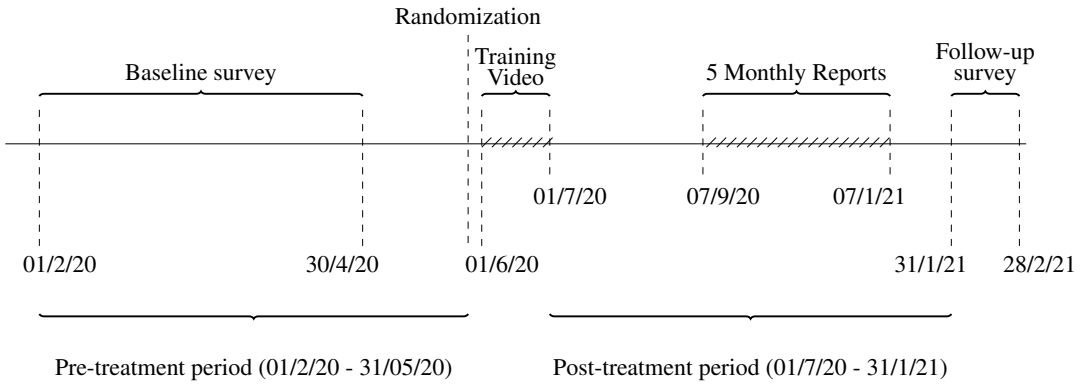


Table 6: Identification of Model Predictions

Object of Interest	Buyer's Motivation	Public Intervention	Private Intervention	Model Prediction	Test Results
Direct Effects on π					
$\partial\pi/\partial I$	Extrinsic	≥ 0	$= 0$	Predictions E.1	Table 7
$\partial\pi/\partial I$	Intrinsic	> 0	> 0	Predictions N.1	Table 7
$\partial\pi/\partial\sigma$	Extrinsic	$\lesseqgtr 0$	< 0	Prediction E.4	Table 8
$\partial\pi/\partial\sigma$	Intrinsic	$\lesseqgtr 0$	> 0	Prediction N.4	Table 8
$\partial\pi/\partial(\Delta p)$	Extrinsic	$\lesseqgtr 0$	< 0	Prediction E.6	Table 8
$\partial\pi/\partial(\Delta p)$	Intrinsic	$\lesseqgtr 0$	< 0	Prediction N.6	Table 8
Heterogeneous Effects on π					
$\partial\pi/\partial I$ when $\theta_B > \bar{\theta}$	Extrinsic	> 0	$= 0$	Prediction E.1	Table 12
$\partial\pi/\partial I$ when $\theta_B > \bar{\theta}$	Intrinsic	> 0	> 0	Prediction N.1	Table 12
$\partial^2\pi/\partial I \cdot \partial\sigma$	Extrinsic	$\lesseqgtr 0$	$= 0$	Prediction E.5	Table 8
$\partial^2\pi/\partial I \cdot \partial\sigma$	Intrinsic	$\lesseqgtr 0$	$\lesseqgtr 0$	Prediction N.5	Table 8
$\partial^2\pi/\partial I \cdot \partial(\Delta p)$	Extrinsic	$\lesseqgtr 0$	$= 0$	Prediction E.7	Table 13
$\partial^2\pi/\partial I \cdot \partial(\Delta p)$	Intrinsic	$\lesseqgtr 0$	$\lesseqgtr 0$	Prediction N.7	Table 13
Mechanisms of π					
$\partial a^*/\partial I$	Extrinsic	> 0	$= 0$	Prediction E.2	Table **??**
$\partial a^*/\partial I$	Intrinsic	$= 0$	$= 0$	Prediction N.2	Table **??**
$\partial P(\theta_B, a)/\partial I$	Extrinsic	> 0	$= 0$	Prediction E.3	Table 11
$\partial P(\theta_B, a)/\partial I$	Intrinsic	$= 0$	$= 0$	Prediction N.3	Table 11

Notes: First column presents the object of interest associated to model predictions. Second column presents the resulting implication for buyer's motivation (if extrinsic, intrinsic, or both). Next two columns present multiple combinations of the potential signs of the object under public and private interventions. Column 5 presents the associated prediction that is met. Last column indicates the corresponding Table providing the empirical test. Cases were Public and/or Private intervention effects are negative are out of the model predictions and thus are not considered.

Table 7: Treatment Effects on Purchase Efficiency

Outcome	Overprice per-unit purchased		Quantity Outcomes			
	Relative to all products in ref. group	Relative to identical products in ref. group	Log Number of units purchased per product	Number of F.A. purchase orders (FOCs products)	Number of F.A. purchase orders (Any product)	Share of purchase orders procured through F.A. (Any product)
Unit of Analysis	Product purchased in F.A. (FOCs products)	Product purchased in F.A. (Identical FOCs products)	Product purchased in F.A. (FOCs products)	F.A. Buyers	F.A. Buyers	F.A. Buyers
	OLS (1)	OLS (2)	OLS (3)	Poisson (4)	Poisson (5)	OLS (6)
Public Report	-0.016**	-0.024***	0.112	-0.299	4.827	-0.023
	(0.007)	(0.008)	(0.086)	(1.074)	(7.234)	(0.019)
Private Report	0.000	-0.002	-0.027	0.783	4.447	0.026
	(0.005)	(0.007)	(0.087)	(0.951)	(5.125)	(0.018)
Observations	93,769	25,246	93,769	2,076	2,076	2,076
No. Buyers	2,076	1,713	2,076	2,076	2,076	2,076
Control Mean	0.103	0.089	3.408	9.226	48.310	0.637
Baseline Outcome	Yes	Yes	Yes	Yes	Yes	Yes
Strata FEs	Yes	Yes	Yes	Yes	Yes	Yes
Product-type FEs	Yes	Yes	Yes	No	No	No
Week FEs	Yes	Yes	Yes	No	No	No
p-val. Public=Private	0.014	0.023	0.282	0.345	0.959	0.011

Notes: Intention-to-treat effects of exposure to the Public and Private performance reports over the full post-treatment period (July 2020 to January 2021). Regression columns (1)-(3) are at the product level for products purchased in F.A., and each product counts as one transaction regardless of the number of units purchased per product. Only purchases from the Food, Office Supplies, and Computers (FOCs) Framework Agreement catalogs are considered, for which a reference group can be identified. Following equation 1, overprice is defined per unit of product, as the difference between the per-unit price paid for the product and the lowest per-unit price in that day's price distribution among products in the same reference group, scaled by the latter. A reference group is a set of products that are sufficiently similar in functionality and quality to be plausible close substitutes for the product, following the mixed strategy detailed in Section **??**. Overprice regression (1) compares against all products in the reference group; column (2) restricts the comparison to identical products only (same SKU). Column (3) is the log number of units purchased per product. Regression columns (4)-(6) are at the buyer level and consider only buyers who purchased at least one product in F.A. channel. Column (4) is estimated by a Poisson regression model at the buyer level, and the outcome counts the number of purchase orders issued in the FOCs Framework Agreement catalogs. Column (5) replicates regression (4) but including any product. Column (6) is a dummy equal to one if the purchase order was issued through the Framework Agreement channel and zero if an alternative mechanism (Procurement Auctions, Direct Purchases) was used. All models control for the average outcome in the procurement unit during the pre-intervention period (Feb.-May 2020) and for strata fixed effects; additional fixed effects for transaction-level regressions are included as indicated in the table. A total of 23 overprice observations are dropped due to multicollinearity with control variables and fixed effects, and that explains why the number of observations reported in columns (1) and (3) is lower than the original experimental sample presented in Table 5. Standard errors clustered at the procurement unit level in parentheses. The bottom row reports the p-value for the null of no difference between the Public and Private treatment effects. $^{***}p < 0.01$, $^{**}p < 0.05$, $^{*}p < 0.10$.

Table 8: Treatment Effects on Overprice by Price Dispersion

	Overprice per-unit purchased			
	OLS (1)	OLS (2)	OLS (3)	OLS (4)
Public Report	-0.017***	-0.015***	-0.017**	-0.016***
	(0.006)	(0.005)	(0.007)	(0.006)
Private Report	0.000	-0.002	0.001	-0.001
	(0.005)	(0.005)	(0.006)	(0.005)
Price Ranking StdDev		0.030***		0.025***
		(0.003)		(0.003)
Public Report × Price Ranking StdDev		-0.007**		-0.006**
		(0.003)		(0.003)
Private Report × Price Ranking StdDev		-0.002		-0.001
		(0.003)		(0.003)
Between Price CV			0.002***	0.002***
			(0.000)	(0.000)
Public Report × Between Price CV			0.000	0.000
			(0.000)	(0.000)
Private Report × Between Price CV			0.000	0.000
			(0.000)	(0.000)
Observations	90,998	90,998	90,998	91,006
No. Buyers	2,065	2,065	2,065	2,065
Control Mean	0.103	0.103	0.103	0.103
Baseline Outcome	Yes	Yes	Yes	Yes
Strata FEs	Yes	Yes	Yes	Yes
Product-type FEs	Yes	Yes	Yes	Yes
Week FEs	Yes	Yes	Yes	Yes

Notes: Heterogeneous treatment effects of exposure to the Public and Private performance reports on overprice across proxies of search costs faced by buyers. Period of analysis is the full post-treatment period (July 2020 to January 2021). Buyers purchase by issuing purchase orders, each of which may contain multiple line items. Each line item is a distinct product, of which the buyer may purchase several units; each line item counts as one transaction regardless of units. The outcome is overprice. Following equation 1, overprice is defined per unit of product, as the difference between the per-unit price paid for the product and the lowest per-unit price in that day's price distribution among products in the same reference group, scaled by the latter. A reference group is a set of products that are sufficiently similar in functionality and quality to be plausible close substitutes for the product, following the mixed strategy detailed in Section **??**. All regressions compare against all products in the reference group, and are limited to purchases from the Food, Office Supplies, and Computers Framework Agreement catalogs, for which a reference group can be identified. All models control for the average outcome in the procurement unit during the pre-intervention period (February to May 2020), as well as for Strata FEs, product-type FEs, and week FEs. Rank Standard Deviation captures *time variation in price rankings* in the baseline period (February to May 2020). Define $P_t = \{p_{it}\}_{i \in R}$ as the set of prices in reference group R on date t, and $w_{it} = \text{rank}_{P_t}(p_{it})$ as the position of the price of the product in this ranking. The standard deviation of this ranking for each product, $SDRank_i = \text{StDev}(\{w_{it}\}_t)$, captures how frequently that product changes its relative price position over time. The average of these standard deviations across all products $i \in R$ constitutes the *Price Ranking StdDev* for the reference group. Between Price CV captures the *cross-sectional price dispersion* between suppliers. Define the average price of each supplier-product, indexed i, during the baseline period T as $\bar{p}_i = \frac{1}{T}\sum_t p_{it}$. Then, for each reference group R, products are aggregated to calculate the Between Price CV as $CV_R = \sigma_R / \bar{p}_R$, where $\bar{p}_R$ and σ_R are the mean and standard deviation of the average prices $\bar{p}_i$ between all products $i \in R$ in the reference group. Transactions for which Rank Standard Deviation or Between Price CV can not be computed are excluded. Standard errors clustered at the procurement unit level in parentheses. $^{***}p < 0.01$, $^{**}p < 0.05$, $^{*}p < 0.10$.

Table 9: First Stage for Predictions: Treatment Effects on Information

	Reported by Manager			Reported by Buyer		
	Informed about Unit's Overspending (=1)	DIPRES Monitors Purchases (1–7)	DIPRES knows Unit's Overspending (=1)	Informed about Own Overspending (=1)	Informed about Unit's Overspending (=1)	DIPRES knows Unit's Overspending (=1)
	OLS	OLS	OLS	OLS	OLS	OLS
	(1)	(2)	(3)	(4)	(5)	(6)
Public Report	0.605***	1.307***	0.242**	0.233***	0.274***	0.080***
	(0.084)	(0.331)	(0.091)	(0.043)	(0.053)	(0.029)
Private Report	0.665***	0.912***	0.063	0.142***	0.150***	0.084**
	(0.086)	(0.358)	(0.098)	(0.050)	(0.049)	(0.035)
Observations	140	140	140	1,098	1,100	884
Control Mean	0.347	4.408	0.735	0.282	0.160	0.827
Baseline Outcome	Yes	Yes	Yes	Yes	Yes	Yes
Strata FEs	Yes	Yes	Yes	Yes	Yes	Yes
p-val. Public=Private	0.430	0.229	0.065	0.053	0.015	0.899

Notes: Intention-to-treat effects of exposure to the Public and Private performance reports on information outcomes measured in the follow-up survey (February 2021). Regression columns (1)–(3) are at the manager level with robust standard errors in parenthesis. Columns (4)–(6) are at the buyer level with standard errors clustered at the unit level. Differences in number of observations in buyer-level regressions are due to not all buyers responded all the three questions. All models control for the individual baseline value of the outcome as well as strata fixed effects. The bottom row reports the p-value for the null hypothesis of no difference between Public and Private treatment effects. $^{***}p < 0.01$, $^{**}p < 0.05$, $^{*}p < 0.10$.

Table 10: Treatment Effects on Manager's Actions (reported by Buyer)

	Manager Actions Index (SD)	Manager Sets Efficiency Goals (=1)	Manager Has High-level of Communic. with Buyers (=1)	Manager Sets Regular Meetings (=1)	Manager Monitor and Supervise Purchases (=1)	Manager Keep Overspending Records (=1)	Manager Train Buyers on Reducing Overspending (=1)	Manager Encourage Buyers to Purchase Efficiently (=1)	Manager Set Purchasing Manuals (=1)	Manager Promote Coordinated Purchases (=1)	Manager Implement Perf.-based Rewards (=1)
	OLS (1)	OLS (2)	OLS (3)	OLS (4)	OLS (5)	OLS (6)	OLS (7)	OLS (8)	OLS (9)	OLS (10)	OLS (11)
Public Report	0.264***	0.076*	0.103**	0.024	−0.020	0.009	0.043**	0.069	0.005	0.074*	0.015***
	(0.085)	(0.040)	(0.042)	(0.034)	(0.045)	(0.035)	(0.018)	(0.042)	(0.043)	(0.042)	(0.006)
Private Report	−0.058	−0.030	−0.043	−0.066*	0.016	0.020	0.010	0.058	−0.006	−0.085*	0.006
	(0.093)	(0.036)	(0.049)	(0.035)	(0.041)	(0.033)	(0.017)	(0.039)	(0.045)	(0.048)	(0.006)
Observations	1,040	1,040	1,040	1,040	1,040	1,040	1,040	1,040	1,040	1,040	1,040
No. Services	153	153	153	153	153	153	153	153	153	153	153
Control Mean	0.000	0.302	0.690	0.304	0.304	0.062	0.027	0.228	0.212	0.226	0.005
Strata FEs	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
p-val. Public=Private	0.006	0.007	0.001	0.005	0.489	0.800	0.158	0.788	0.803	0.001	0.263

Notes: Intention-to-treat effects of exposure to the Public and Private performance reports on manager oversight and communication activities, measured in the buyers' follow-up survey (February 2021). Buyers' reports rather than managers' are the source of these outcomes to avoid social desirability bias in manager self-reports. The sample is restricted to buyers working in services with FA purchase data (=153). All regressions are at the buyer level, controlling for strata fixed effects. The dependent variable in column (1) is the Anderson index of the binary outcomes in columns (2)–(11), expressed in control-group standard deviation units (Anderson, 2008). Standard errors clustered at the unit level are shown in parenthesis. The bottom row reports the p-value for the null hypothesis of no difference between Public and Private treatment effects. $^{***}p < 0.01$, $^{**}p < 0.05$, $^{*}p < 0.10$.

Table 11: Treatment Effects on Buyers' Stated Persuadability toward Efficiency

	Buyer is More Persuadable toward Efficiency (10% Audit)	Buyer is More Persuadable toward Efficiency (5% Audit)	Buyer is More Persuadable toward Efficiency (1% Audit)
	OLS	OLS	OLS
	(1)	(2)	(3)
Public Report	0.061	0.081**	0.135***
	(0.038)	(0.036)	(0.038)
Private Report	0.015	0.007	0.065
	(0.043)	(0.040)	(0.040)
Observations	957	957	957
Control Mean	0.580	0.536	0.456
Baseline Outcome	Yes	Yes	Yes
Strata FEs	Yes	Yes	Yes
p-val. Public=Private	0.317	0.081	0.090

Notes: Intention-to-treat effects of exposure to the Public and Private performance reports on buyers' stated preferences about efficiency, measured in the buyers' follow-up survey (February 2021). Buyers were asked: *"If DIPRES audits X% of your purchases, would you prioritize price over quality?"*, with X set to 10%, 5%, and 1% across columns (1)–(3). Responses are coded as a binary dummy equal to 1 if the buyer answered yes, and zero otherwise. The lower the audit margin, the more informative the response is about the buyer's underlying commitment to efficiency: prioritizing price over quality despite a low audit probability is less likely to reflect compliance with audit risk and more likely to reflect a genuine preference for cost savings. Buyers who answer yes are classified as more persuadable by managers, since their stated preferences align more with the price-efficiency objective. All regressions are at the buyer level, controlling for the corresponding baseline audit outcome as well as strata fixed effects. Standard errors clustered at the unit level are shown in parentheses. The bottom row reports the p-value for the null hypothesis of no difference between Public and Private treatment effects. $^{***}p < 0.01$, $^{**}p < 0.05$, $^{*}p < 0.10$.

Table 12: Treatment Effects on Overprice by Buyers' Stated Persuadability toward Efficiency

	Overprice per-unit purchased		
	OLS (1)	OLS (2)	OLS (3)
Public Report	0.011	0.011	0.014
	(0.008)	(0.008)	(0.009)
Private Report	0.004	0.002	0.003
	(0.007)	(0.008)	(0.008)
Buyer is More Persuadable toward Efficiency (1% Audit)	0.001		
	(0.007)		
Public Report × Buyer is More Persuadable toward Efficiency (1% Audit)	-0.028***		
	(0.009)		
Private Report × Buyer is More Persuadable toward Efficiency (1% Audit)	-0.003		
	(0.009)		
Buyer is More Persuadable toward Efficiency (5% Audit)		0.001	
		(0.007)	
Public Report × Buyer is More Persuadable toward Efficiency (5% Audit)		-0.028***	
		(0.009)	
Private Report × Buyer is More Persuadable toward Efficiency (5% Audit)		0.001	
		(0.010)	
Buyer is More Persuadable toward Efficiency (10% Audit)			0.002
			(0.007)
Public Report × Buyer is More Persuadable toward Efficiency (10% Audit)			-0.029***
			(0.010)
Private Report × Buyer is More Persuadable toward Efficiency (10% Audit)			0.000
			(0.010)
Public + Public × Buyer is More Persuadable toward Efficiency (X% Audit)	−0.017*	−0.017*	−0.016*
	(0.009)	(0.009)	(0.009)
Private + Private × Buyer is More Persuadable toward Efficiency (X% Audit)	0.001	0.003	0.003
	(0.007)	(0.007)	(0.007)
Observations	38,423	38,423	38,423
No. Buyers	829	829	829
Control Mean	0.099	0.099	0.099
Baseline Outcome	Yes	Yes	Yes
Strata FEs	Yes	Yes	Yes
Product-type FEs	Yes	Yes	Yes
Week FEs	Yes	Yes	Yes

Notes: Intention-to-treat effects of exposure to the Public and Private performance reports on Overprice over the full post-treatment period (July 2020 to January 2021), by buyer's type. This is defined based on buyers' stated preferences for efficiency, measured in the buyers' baseline survey. Buyers were asked: *"If DIPRES audits X% of your purchases, would you prioritize price over quality?"*, with X set to 10%, 5%, and 1%. Responses are coded as a binary dummy equal to 1 if the buyer answered yes, and zero otherwise. Buyers who answer yes are classified as more persuadable toward efficiency, since their stated preferences align more with the price-efficiency objective. Regressions are at the product level for products purchased in F.A., and each product counts as one transaction regardless of the number of units purchased per product. Following equation 1, overprice is defined per unit of product, as the difference between the per-unit price paid for the product and the lowest per-unit price in that day's price distribution among products in the same reference group, scaled by the latter. A reference group is a set of products that are sufficiently similar in functionality and quality to be plausible close substitutes for the product, following the mixed strategy detailed in Section **??**. Overprice is calculated for purchases from the Food, Office Supplies, and Computers Framework Agreement catalogs, for which a reference group can be identified. All models control for the average outcome in the procurement unit during the pre-intervention period (Feb.-May 2020), strata FE, product-type FE, and week FE. Only purchases made by buyers that answered the stated preferences module in the baseline survey are considered. In all, standard errors clustered at the procurement unit level are shown in parentheses. The bottom rows report sum of coefficients for the corresponding subgroup. $^{***}p < 0.01$, $^{**}p < 0.05$, $^{*}p < 0.10$

Table 13: Treatment Effects on Overprice by Buyer's Level of Overspending at Baseline

	Overprice per-unit purchased	
	OLS (1)	OLS (2)
Public Report	-0.014**	-0.012*
	(0.006)	(0.007)
Private Report	-0.005	-0.003
	(0.006)	(0.006)
Buyer in Bottom 25% of Overprice within Service	-0.027***	
	(0.003)	
Public Report × Buyer in Bottom 25% of Overprice within Service	0.014*	
	(0.008)	
Private Report × Buyer in Bottom 25% of Overprice within Service	0.013**	
	(0.006)	
Buyer in Bottom 50% of Overprice within Service		-0.016***
		(0.003)
Public Report × Buyer in Bottom 50% of Overprice within Service		0.001
		(0.008)
Private Report × Buyer in Bottom 50% of Overprice within Service		-0.001
		(0.006)
Public Report + Public Report × Buyer in Bottom X% of Overprice within Service	-0.001	-0.011
	(0.010)	(0.008)
Private Report + Private Report × Buyer in Bottom X% of Overprice within Service	0.009	-0.003
	(0.007)	(0.007)
Observations	72,829	72,829
No. Buyers	1,277	1,277
Control Mean	0.101	0.101
Baseline Outcome	Yes	Yes
Strata FEs	Yes	Yes
Product-type FEs	Yes	Yes
Week FEs	Yes	Yes

Notes: Intention-to-treat effects of exposure to the Public and Private performance reports on overprice over the full post-treatment period (July 2020 to January 2021), by buyer's level of overspending in the pre-intervention period (February to May 2020). This is defined based on a dummy equal to 1 if the buyer's overspending is in the bottom X% of the distribution of overspending across buyers within the service, and zero otherwise. All regressions are at the product level for products purchased in F.A., and each product counts as one transaction regardless of the number of units purchased per product. Following equation 1, overprice is defined per unit of product, as the difference between the per-unit price paid for the product and the lowest per-unit price in that day's price distribution among products in the same reference group, scaled by the latter. A reference group is a set of products that are sufficiently similar in functionality and quality to be plausible close substitutes for the product, following the mixed strategy detailed in Section **??**. Overprice is calculated for purchases from the Food, Office Supplies, and Computers Framework Agreement catalogs, for which a reference group can be identified. All models control for the average outcome in the procurement unit during the pre-intervention period, strata FE, product-type FE, and week FE. Only purchases made by buyers that issued purchases during the pre-intervention period are considered. Standard errors clustered at the procurement unit level in parentheses. The bottom rows report sum of coefficients for the corresponding subgroup. $^{***}p < 0.01$, $^{**}p < 0.05$, $^{*}p < 0.10$

A. Appendix: Building Reference Groups

Table 14: Examples of Products Description and Identified Attribute Values

FA: Computers (Desktop)

Id	Product description	Type	Procesor	Brand	HD Cap	RAM	OS	HD Type	Screen
1537456	COMPUTADOR SP LABS INTEL CORE I5-7500, MSI B250M PRO-VH LGA 1151, DDR4-2400 M.2, HDMI, VGA, USB 3.1 MICRO-ATX, MEMORIA RAM 8GB DDR4 , HDD 1 TB , FREE DOS, GRABADOR DE DVD DESKTOP	Desktop	INTEL CORE I5	SP LABS	1TB	8GB	FREEDOS	HDD	None
1615084	LENOVO DESKTOP M720Q , PORCESADOR INTEL CORE I7-8700T,8GB DDR4 2666MHZ SODIMM, 256GB SOLID STATE DRIVE, WIFI INTEL 3165+BT 1X1AC, WINDOWS 10 PRO, MONITOR LENOVO S22E 21.5"" 1920X1080	Desktop	INTEL CORE I7	LENOVO	256GB	8GB	WINDOWS 10 PRO	SSD	21.5

FA: Office Supplies (Print Paper)

Id	Product description	Type	Size	Brand	Pages/Unit	Weight	Color	Units/Pkg
1533457	PAPEL MULTIPROPÓSITO CHAMEX ECO CARTA 75GR UNIDAD ECO CARTA 75GR UNIDAD MEDIDA 216X279MM, RESMA 500 HOJAS	Print Paper	Letter	CHAMEX	500	75	BLANCO	1
1006598	PAPEL MULTIPROPÓSITO XEROX OFICIO 75GR ALBURA 90-95% RESMA 500 HJS.	Print Paper	OFICIO	XEROX	500	75	BLANCO	1
1006553	PAPEL MULTIPROPÓSITO DIAZOL A4 SPECTRA 500HJ AMARILLO	Print Paper	A4	DIAZOL	500	NA	AMARILLO	1

FA: Food (Chicken)

Id	Product	Type	Brand	Maritated	Weight
552243	frozen chicken boneless breast not marinated caja 15 k	frozen chicken		not marinated	15000.0 gr.
552447	frozen chicken boneless breast las camelias not marinated caja 15 k	frozen chicken	las camelias	not marinated	15000.0 gr.
1471381	frozen chicken boneless breast not marinated bolsa 1k	frozen chicken		not marinated	1000.0 gr.
1505907	frozen chicken boneless breast seara not marinated bolsa 2k	frozen chicken	seara	not marinated	2000.0 gr.

FA: Food (Tomato)

Id	Product	Type	Model	Weight
1473761	fresh tomato cherry 1 kg. aprox.	fresh vegetables	cherry	1000.0 gr.
1473779	fresh tomato pomarola small size 1 kg. aprox.	fresh vegetables	pomarola small size	1000.0 gr.
1473794	fresh tomato pomarola medium size 1 kg. aprox.	fresh vegetables	pomarola medium size	1000.0 gr.
1473808	fresh tomato pomarola large size 1 kg. aprox.	fresh vegetables	pomarola large size	1000.0 gr.
1473819	fresh tomato pomarola extra large size 1 kg. aprox.	fresh vegetables	pomarola extra large size	1000.0 gr.
1473833	fresh tomato larga vida small size 1 kg. aprox.	fresh vegetables	pomarola small size	1000.0 gr.

B. Appendix Tables and Figures

Figure B.1: Buyer-level Report

DIPRES

CSP · INGENIERÍA INDUSTRIAL UNIVERSIDAD DE CHILE · CENTRO DE SISTEMAS PÚBLICOS

Monthly Overspending Report

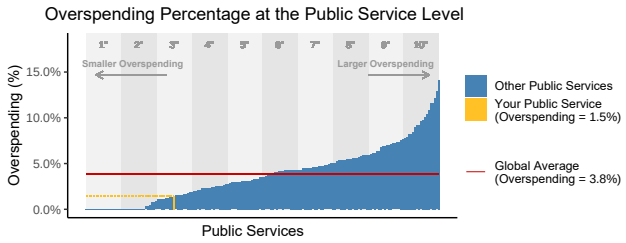
Public Service's name

October 2020

Dear Recipient,

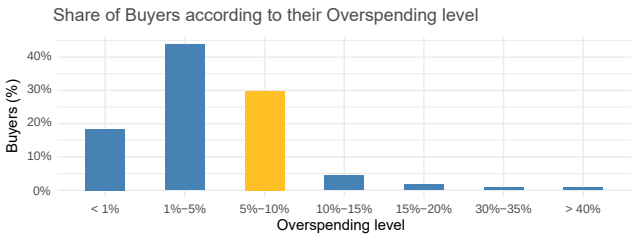
This report presents relevant information regarding the state of public spending executed by the Public Service in which you are employed, covering the period September 2019 - September 2020. This report comprises an analysis of aggregated purchases made by your Public Service, as well as a detailed breakdown of your purchases within the Public Market platform. We would like to inform you that **your average Overspending as a buyer is 7%. You are among the top 25% of buyers in your Service with the highest Overspending**. You are welcome to direct any inquiries regarding this report to gastoeficiente@dipres.gob.cl.

Remember that planning the expenses you make for your Service allows you to execute more efficient purchases, potentially leading to significant savings. For this purpose, ChileCompra's users have access to various procurement mechanisms such as Framework Agreement, Compra Ágil, or Procurement Auctions, which enable more strategic purchasing compared to other alternatives like Direct Purchases.



This chart displays the Overspending associated to each Public Service of the State, arranged in ascending order from the lowest to the highest Overspending (see definition in Appendix B). The relative position of your Public Service is indicated in yellow[1].

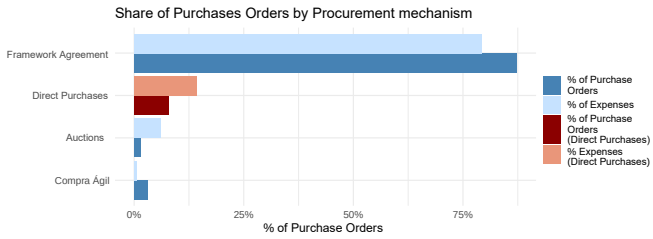
The average Overspending in your Public Service is equal to 1.5%, which positions it among the 50 Services with the lowest Overspending out of a total of 184 Services (excluding Autonomous Services and Subsecretariats). As an illustrative example, considering purchases made exclusively through Framework Agreement the Overspending in your Public Service amounts to approximately $200 million pesos during this period.



This chart displays the share of buyers in your Public Service according to their level of Overspending (%). For instance, 18% of the buyers within the Public Service have a level of Overspending below 1%.

The yellow bar indicates the range where you are positioned relative to the rest of buyers in your Service. Your average Overspending as a buyer is 7%.

On average, the buyers of your Service have incurred an Overspending level of 5.1%.



This chart illustrates the distribution of your purchases by procurement mechanism.

You executed the 87.3% of your purchases through Framework Agreement, which corresponds to 79.2% of the total expenditure you have incurred.

Remember to prioritize procurement procedures such as Framework Agreement, Auctions, and Compra Ágil, as they facilitate better expenditure planning and obtaining better prices.

Please remember that by conducting thorough searches and planning purchases, you can reduce your Overspending.

[1]In order to provide further information on best practices in the execution of public procurements, DIPRES offers an online course, which can be accessed through the following link. The course access is individualized. If you wish to grant access to the course for other individuals within your Service, please write to gastoeficiente@dipres.gob.cl.

1

Note: This is the Buyer-level report. It shows the information displayed in the monthly reports sent to buyers within Public and Private Treatment units.

Figure B.2: Backpage of Buyer-level Reports

Appendix B

Computation of Overspending

The Overspending is an efficiency measure for purchases of standardized products, i.e. products with comparable attributes to other products in the market. The Overspending associated with the purchase of a product measures how much could have been saved if the purchase had been made at a reference price. The reference price corresponds to the lowest price among the set of comparable products that were available in the purchasing region at the time of the purchase. For each purchase executed by a buyer, there exists an associated Overspending Amount and an associated Overspending Percentage.

Consider the purchase of a product on a specific date. At the time of the purchase, there is a number of comparable products that meet the buyer's needs, whose availability will depend on the product's region of origin. As mentioned above, the reference price corresponds to the lowest price among those comparable products available in the purchasing region. The Overspending Amount for this purchase will then be calculated as follows:

$$Overspending\ Amount = (Purchase\ Price - Reference\ Price) \times Purchase\ Quantity$$

A buyer may execute multiple purchases within a specific period, and each purchase will have a different Overspending Amount. To compute the Overspending Percentage of all the purchases made by a buyer, we sum the Overspending Amount associated with each purchase and divide the total amount by the total expenditure incurred in those purchases.

Example: Let's assume that a user makes two purchases in a specific period. The first purchase is a ream of printer paper priced at 2100 CLP, with a reference price of 1800 CLP. The associated Overspending Amount for that purchase will be $300 \times 1 = 300$ CLP. The second purchase consists of 50 disposable face masks, priced at 190 CLP each, with a reference price of 130 CLP. The associated Overspending Amount for that purchase will be $60 \times 50 = 3000$ CLP. Thus, the Overspending Percentage associated to the two purchases executed by the buyer is computed as follows:

$$Overspending\ \% = \left(100 \times \frac{300 + 3000}{1800 + 130 \times 50}\right)\% = 39.7\%$$

Thus, if we wish to compute the Overspending Percentage at the Public Service level, we sum the Overspending Amount associated with all purchases made by each buyer of the Public Service, and divide it by the total expenditure associated with those purchases. Please notice that not every purchase executed by the Public Service is associated to a standardized product. Therefore, the results presented in this report may not comprise all the purchases executed by the Public Service during the analyzed period.

Recommendations

- Remember that you can use Compra Ágil as an alternative to Direct Purchases. In this manner, you will be able to access a wider range of products and sellers, leading to lower prices.
- When executing a purchase through Compra Ágil or Direct Purchases, clearly specify the product you want to procure. We recommend searching for the product in other catalogs and identifying its relevant characteristics, so you can specify them in your request. This way, you will ensure receiving a broad set of products meeting your purchasing needs.
- If you want to reduce your Overspending, consider planning your purchases and grouping them into a single purchase order using the electronic catalog of products for Framework Agreements, which might provide lower unit prices.

Note: This is the backpage included in the buyer-level reports. It informs how *ChileCompra* calculates Overspending as well as their recommendations to make efficient purchases.

Figure B.3: Manager-level Report

DIPRES

CSP INGENIERÍA INDUSTRIAL UNIVERSIDAD DE CHILE CENTRO DE SISTEMAS PÚBLICOS

Monthly Overspending Report

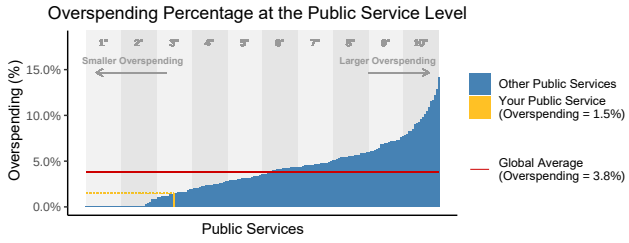
Public Service's name

October, 2020

Dear Recipient

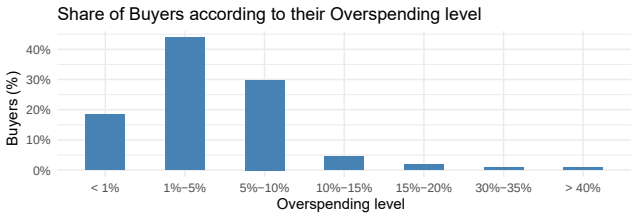
This report presents relevant information regarding the state of public spending executed by the Public Service in which you are employed, covering the period September 2019 - September 2020. The report comprises an analysis of aggregated purchases made by your Public Service. The purpose is to facilitate the management of Overspending within your Public Service. **Your Public Service ranks among the 50 Services with the lowest Overspending** out of a total of 184 Services (excluding Autonomous Public Services and Subsecretariats). You are welcome to address any inquiries or concerns related to this report by contacting us at gastoeficiente@dipres.gob.cl.

Please, keep in consideration that effective expenditure planning within your Service allows for more efficient purchases, leading to significant cost savings. For this purpose, ChileCompra's users have access to various procurement mechanisms such as Framework Agreements, Compra Agil, or Procurement Auctions, which enable more strategic purchasing compared to other alternatives like Direct Purchases.



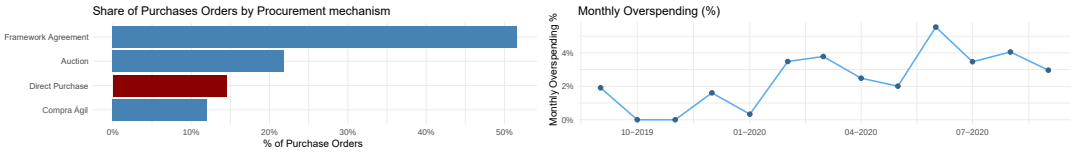
This chart displays the Overspending associated to each Public Service of the State, arranged in ascending order from the lowest to the highest Overspending (see definition in Appendix B). The relative position of your Public Service is indicated in **yellow**[1].

The average Overspending in your Public Service is equal to 1.5%, which positions it among the 50 Services with the lowest Overspending out of a total of 184 Services (excluding Autonomous Services and Subsecretariats). As an illustrative example, considering purchases made exclusively through Framework Agreement the Overspending in your Public Service amounts to approximately $200 million pesos during this period.



This chart displays the share of buyers in your Public Service according to their level of Overspending (%). For instance, 18% of the buyers within the Public Service have a level of Overspending below 1%.

On average, the buyers in your Service have incurred in an Overspending level of 5.1%. **In order to enhance efficiency in public expenditure, it may be crucial for you to remind the buyers about the importance of minimizing Overspending as much as possible in their purchases.**



The chart on the left illustrates the distribution of total purchases by procurement mechanism. 51.6% of the purchases made by your Service are through Framework Agreement. **Remember to prioritize procurement mechanisms such as Framework Agreement, Procurement Auctions, and Compra Ágil, which allow for better expenditure planning and better prices.** The chart on the right shows the monthly evolution of Overspending in your Service during the reported period.

Please remember that by conducting thorough searches and planning purchases, you can reduce your Overspending.

[1] In order to provide further information on best practices in the execution of public procurement, DIPRES offers an online course, which can be accessed through the following link. The course access is individualized. If you wish to grant access to the course for other individuals within your Service, please write to gastoeficiente@dipres.gob.cl.

1

Note: This is the Manager-level report. It shows the information displayed in the monthly reports sent to managers of Public and Private Treatment units.

Figure B.4: Ranking of Buyers by Overspending, received only by Managers in Public Treatment

Appendix A

Please find below the list of active buyers of your service. The list indicates each buyer's Overspending level and the buyer's share of expenditure within the Public Service associated to the analyzed purchases. The buyers are organized in descending order according to their Overspending level. This arrangement aims to simplify the identification of individuals in need of more immediate adjustments regarding their purchasing behavior.

Name	Email	Overspending (%)	Exp. (%)
		> 40%	< 0.1%
		30%-35%	< 0.1%
		15%-20%	0.1%
		15%-20%	< 0.1%
		10%-15%	0.1%
		10%-15%	0.5%
		10%-15%	< 0.1%
		10%-15%	0.3%
		10%-15%	0.6%
		5%-10%	1.1%
		5%-10%	0.4%
		5%-10%	0.1%
		5%-10%	0.2%
		5%-10%	0.3%
		5%-10%	0.7%
		5%-10%	0.4%
		5%-10%	0.2%
		5%-10%	0.9%
		5%-10%	0.2%
		5%-10%	0.2%
		5%-10%	0.5%
		5%-10%	2.9%
		5%-10%	0.4%
		5%-10%	0.4%
		5%-10%	< 0.1%
		5%-10%	0.3%
		5%-10%	0.3%
		5%-10%	< 0.1%
		5%-10%	0.5%
		5%-10%	0.5%
		5%-10%	0.1%
		5%-10%	0.1%
		5%-10%	< 0.1%
		5%-10%	< 0.1%
		5%-10%	0.1%
		5%-10%	< 0.1%
		5%-10%	0.2%
		5%-10%	0.1%
		5%-10%	0.4%
		5%-10%	3.9%
		5%-10%	0.2%
		5%-10%	0.2%
		5%-10%	0.5%
		1%-5%	3.6%
		1%-5%	0.4%
		1%-5%	0.6%
		1%-5%	1.6%
		1%-5%	7.1%
		1%-5%	3.8%
		1%-5%	3.5%
		1%-5%	0.1%
		1%-5%	0.5%

Note: The list contains the names and emails of all buyers in the procurement units and the range of their overspending, ranked by largest to lowest, as well as the share of expenditures incurred by each buyer. This information is displayed in an additional backpage of monthly reports, but it is included only for managers in units assigned to the Public treatment group.

Table B.1: Experimental Sample

Ministry	No. of Service Units
Ministry of Agriculture	5
Ministry of Science and Technology	1
Ministry of Defense	14
Ministry of Social Development and Family	5
Ministry of Economics, Development, and Tourism	12
Ministry of Education	19
Ministry of Energy	3
Ministry of Finance	8
Ministry of Justice and Human Rights	5
Ministry of Women and Gender Equality	1
Ministry of Culture and Arts	1
Ministry of Environmental Protection	2
Ministry of Mining	2
Ministry of Public Infrastructure	14
Ministry of International Relationships	4
Ministry of Health	36
Ministry of Transport and Telecommunications	1
Ministry of Housing and Urbanism	17
Ministry of Sport	1
Ministry of Interior and Public Safety	23
Ministry of Labor and Social Security	9
Ministry of General Secretariat of Government	1
Total No. of Experimental Service Units	184

Notes: List of Ministries and number of selected Service Units per Ministry to participate in the experiment.

Table B.2: Baseline Balance – Purchasing Behavior

	Mean Control (1)	Mean Public (2)	Mean Private (3)	Public vs. Control (4)	Private vs. Control (5)	Public vs. Private (6)
Panel A. Buyer-level						
Extensive margin						
If Buyer purchased in FA channel in pre-treat. period – FOSC	0.590	0.637	0.636	0.058	0.065	-0.008
				(0.039)	(0.040)	(0.034)
If Buyer purchased in any channel in the pre-treatment period	0.924	0.948	0.941	0.001	0.007	-0.006
				(0.012)	(0.015)	(0.015)
Intensive margin: If Buyer purchased in FA channel in pre-treat. period – FOSC						
No. FA Purchase Orders – FOSC	114.797	103.438	108.376	-21.092	11.767	-32.859*
				(16.890)	(19.226)	(18.649)
Average Overprice per unit purchased in FA – FOSC	0.087	0.091	0.098	-0.001	0.000	-0.001
				(0.010)	(0.007)	(0.008)
Average Log Number of units purchased per product in FA – FOSC	6.823	6.811	6.753	-0.151	0.117	-0.268
				(0.321)	(0.335)	(0.321)
Average Price Ranking Std. Dev. in the Ref. group – FOSC	0.847	0.955	0.984	-0.025	0.081	-0.106
				(0.090)	(0.076)	(0.081)
Average Between Price CV in the Ref. group – FOSC	10.711	10.370	10.641	0.216	0.120	0.096
				(0.537)	(0.678)	(0.605)
Within Unit Baseline Overprice Ranking (%) – FOSC	0.537	0.545	0.564	0.021	0.016	0.005
				(0.025)	(0.018)	(0.022)
Within Unit Lowest 25% Baseline Overprice – FOSC	0.213	0.178	0.183	-0.040	-0.008	-0.032
				(0.026)	(0.019)	(0.025)
Within Unit Lowest 50% Baseline Overprice – FOSC	0.457	0.479	0.429	0.000	-0.020	0.020
				(0.036)	(0.030)	(0.035)
Joint orthogonality F-test (Asymptotic SEs)				0.522	0.866	0.488
Joint orthogonality F-test (Clustered Wild Bootstrapped SEs)				0.543	0.904	0.681
Panel B. Purchase-level						
Overprice per unit purchased in FA – FOSC	0.108	0.110	0.117	-0.012	-0.000	-0.012
				(0.011)	(0.008)	(0.007)
Panel C. Unit-Level						
No. Buyers who purchased in FA channel in pre-treat. period – FOSC	13.685	14.558	16.120	0.364	-3.447	3.811
				(6.535)	(6.815)	(8.296)
No. Buyers who purchased in FA channel in pre-treat. period – any product	29.037	30.962	33.220	1.009	-5.296	6.305
				(12.191)	(11.392)	(14.423)
No. Buyers who purchased in any channel in pre-treat. period	37.778	42.538	46.480	3.431	-3.938	7.369
				(15.639)	(13.972)	(18.321)
No. FA Purchase Orders – FOSC	2272.556	2002.481	2592.160	-505.135	-648.305	143.170
				(691.879)	(767.216)	(740.127)
Average Overprice per unit purchased in FA – FOSC	0.101	0.100	0.096	-0.011	-0.019	0.008
				(0.016)	(0.014)	(0.016)
Average Log Number of units purchased per product in FA – FOSC	8.479	8.794	9.071	0.273	0.230	0.044
				(0.485)	(0.463)	(0.489)
Joint orthogonality F-test (Asymptotic SEs)				0.022	0.003	0.581
Joint orthogonality F-test (Clustered Wild Bootstrapped SEs)				0.380	0.184	0.843

Notes: Balance test using *ChileCompra* administrative data covering the baseline period from February to May 2020. In all, we consider only buyers in experimental procurement units who issued FA purchase orders in Food, Office Supplies, or Computers (abbreviated as FOSC) throughout the post-treatment period (July 2020 - January 2021), for a total of 2,076 buyers. Columns (1)–(3) report the baseline mean for each treatment arm. Columns (4)–(6) report OLS estimates of differences in means across groups. Outcomes in Panel A are at the buyer level: mean differences are obtained by regressing the baseline outcome against the treatment dummies and controlling for strata fixed effects (standard errors clustered at the procurement unit level are shown in parentheses). Outcomes in Panel B are at the purchase level: mean differences are obtained by regressing the baseline outcome against the treatment dummies and controlling for strata fixed effects, product-type fixed effects, and week fixed effects (standard errors clustered at the procurement unit level shown in parentheses). Outcomes in Panel C are aggregated at the procurement unit level: mean differences are obtained by regressing the baseline outcome against the treatment dummies and controlling for strata fixed effects (robust standard errors are shown in parentheses). Joint orthogonality F-test are shown at the bottom of each family of outcomes, using both asymptotic SEs as well as clustered wild bootstrapped SEs with 10,000 replications (Cameron, C., Gelbach, J., and Miller, D., 2008).$^{***}p < 0.01$, $^{**}p < 0.05$, $^{*}p < 0.10$.

Table B.3: Baseline Balance – Buyer-level Surveys

	Mean Control (1)	Mean Public (2)	Mean Private (3)	Control vs Public (4)	Control vs Private (5)	Public vs Private (6)
Panel A. Survey Response and Attrition (all ever respondents)						
Responded baseline survey (=1)	0.838	0.831	0.845	-0.004	-0.004	0.000
				(0.024)	(0.020)	(0.022)
Responded follow-up survey (=1)	0.466	0.444	0.434	0.045	0.048	-0.003
				(0.028)	(0.033)	(0.029)
Responded baseline and follow-up survey (=1)	0.303	0.276	0.279	0.041*	0.044	-0.003
				(0.024)	(0.029)	(0.024)
Joint orthogonality F-test (Asymptotic SEs)				0.391	0.241	0.527
Joint orthogonality F-test (Clustered Wild Bootstrapped SEs)				0.575	0.434	0.636
Panel B. Balance on Buyers' Stated Persuadability toward Efficiency – Only if responded baseline survey						
Prioritize Price over Quality if Audit 1%	0.500	0.524	0.468	0.029	-0.038	0.068*
				(0.034)	(0.036)	(0.040)
Prioritize Price over Quality if Audit 5%	0.553	0.584	0.526	0.028	-0.018	0.046
				(0.035)	(0.035)	(0.039)
Prioritize Price over Quality if Audit 10%	0.594	0.616	0.550	0.020	-0.039	0.059
				(0.036)	(0.036)	(0.038)
Joint orthogonality F-test (Asymptotic SEs)				0.657	0.256	0.222
Joint orthogonality F-test (Clustered Wild Bootstrapped SEs)				0.824	0.523	0.401
Panel C. Balance on Buyers' Stated Persuadability toward Efficiency – Only if responded baseline and follow-up survey)						
Prioritize Price over Quality if Audit 1%	0.515	0.567	0.452	0.039	-0.082	0.121*
				(0.058)	(0.066)	(0.069)
Prioritize Price over Quality if Audit 5%	0.570	0.600	0.513	0.000	-0.063	0.063
				(0.055)	(0.061)	(0.064)
Prioritize Price over Quality if Audit 10%	0.603	0.639	0.539	0.004	-0.074	0.077
				(0.054)	(0.056)	(0.063)
Joint orthogonality F-test (Asymptotic SEs)				0.123	0.884	0.449
Joint orthogonality F-test (Clustered Wild Bootstrapped SEs)				0.210	0.961	0.687

Notes: Balance test using the baseline survey data at the buyer-level. Sample of analysis comprises buyers in experimental procurement units who responded any survey (N=2,649). Some responded only the baseline survey, some only the follow up survey, and some responded both. Panel A reports response rates by type of respondents and arm. Panel B tests balance among those who responded the baseline survey (N=1960), regardless of whether responded the follow up survey; Panel C tests balance among those who responded baseline and follow-up surveys (N=715). Prioritize-price outcomes are built from the raw unfolding-bracket audit questions. Buyers were asked: *"If DIPRES audits X% of your purchases, would you prioritize price over quality?"*, with X set to 1%, 5%, and 10%. Responses are coded as a binary dummy equal to 1 if the buyer answered yes, and zero otherwise. The lower the audit margin, the more informative the response is about the buyer's underlying commitment to efficiency: prioritizing price over quality despite a low audit probability is less likely to reflect compliance with audit risk and more likely to reflect a genuine preference for cost savings. Buyers who answer yes are classified as more persuadable by managers, since their stated preferences align more with the price-efficiency objective. Columns (1)–(3) report means by arm; (4)–(6) OLS differences controlling for strata fixed effects. Standard errors clustered at the procurement unit level are shown in parentheses. Joint orthogonality F-tests are shown using both asymptotic and wild cluster bootstrap (10,000 reps., Cameron, C., Gelbach, J., and Miller, D., 2008). $^{***}p < 0.01$, $^{**}p < 0.05$, $^{*}p < 0.10$.

Table B.4: Baseline Balance – Manager-level Surveys

	Mean Control (1)	Mean Public (2)	Mean Private (3)	Control vs. Public (4)	Control vs. Private (5)	Public vs. Private (6)
Panel A. Survey Response and Attrition (184 procurement units)						
Responded baseline survey (=1)	0.918	0.852	0.887	-0.020	0.033	-0.053
				(0.053)	(0.061)	(0.062)
Responded follow-up survey (=1)	0.902	0.820	0.903	-0.028	0.042	-0.070
				(0.060)	(0.060)	(0.060)
Responded baseline and follow-up survey (=1)	0.820	0.738	0.790	0.002	0.039	-0.037
				(0.075)	(0.080)	(0.077)
Joint orthogonality F-test (Asymptotic SEs)				0.907	0.822	0.513
Joint orthogonality F-test (Clustered Wild Bootstrapped SEs)				0.907	0.822	0.502
Panel B. Balance on Managers' Beliefs toward Efficiency – Only if responded baseline survey						
DIPRES values savings (1–7)	4.018	3.558	3.836	-0.393	-0.206	-0.187
				(0.364)	(0.372)	(0.408)
DIPRES monitors purchases (1–7)	3.982	3.827	4.055	-0.178	0.010	-0.188
				(0.348)	(0.380)	(0.370)
Budget execution is a pressure (1–5)	4.196	4.212	4.382	0.074	0.398**	-0.324
				(0.159)	(0.193)	(0.196)
Budget depends on execution (1–5)	3.321	3.327	3.200	0.206	-0.094	0.300
				(0.270)	(0.303)	(0.289)
Savings lead to budget cut (1–5)	3.018	2.692	2.509	-0.228	-0.533*	0.304
				(0.273)	(0.274)	(0.255)
Budget stable if under-executed (1–3)	1.982	2.019	1.855	0.120	-0.068	0.188
				(0.162)	(0.179)	(0.175)
Budget stable if savings shown (1–3)	1.875	2.077	2.018	0.221	0.118	0.102
				(0.149)	(0.134)	(0.144)
Joint orthogonality F-test (Asymptotic SEs)				0.745	0.599	0.146
Joint orthogonality F-test (Clustered Wild Bootstrapped SEs)				0.727	0.539	0.184
Panel C. Balance on Managers' Beliefs toward Efficiency – Only if responded baseline and follow-up surveys						
DIPRES values savings (1–7)	3.900	3.533	3.857	-0.286	0.122	-0.409
				(0.414)	(0.410)	(0.455)
DIPRES monitors purchases (1–7)	3.860	3.800	4.082	-0.159	0.216	-0.374
				(0.379)	(0.403)	(0.399)
Budget execution is a pressure (1–5)	4.200	4.244	4.388	0.011	0.442*	-0.431*
				(0.173)	(0.224)	(0.224)
Budget depends on execution (1–5)	3.320	3.267	3.143	0.318	-0.099	0.417
				(0.296)	(0.327)	(0.311)
Savings lead to budget cut (1–5)	3.020	2.778	2.490	-0.080	-0.437	0.357
				(0.319)	(0.320)	(0.279)
Budget stable if under-executed (1–3)	1.980	1.978	1.837	0.049	-0.057	0.106
				(0.183)	(0.210)	(0.204)
Budget stable if savings shown (1–3)	1.860	2.000	2.000	0.088	0.140	-0.052
				(0.166)	(0.151)	(0.163)
Joint orthogonality F-test (Asymptotic SEs)				0.929	0.758	0.080
Joint orthogonality F-test (Clustered Wild Bootstrapped SEs)				0.914	0.664	0.122

Notes: Balance test using the baseline survey data at the manager-level. Sample of analysis comprises the 184 experimental procurement units, with one manager per service. Panel A reports the share of procurement units whose manager responded the baseline and follow-up surveys, by arm (attrition check). Panel B tests balance among those who responded the baseline survey (N=163); Panel C tests balance among those who responded baseline and follow-up surveys (N=144). Columns (1)–(3) report means by arm; (4)–(6) OLS differences controlling for strata fixed effects. Robust standard errors in parentheses. For each subsample, joint orthogonality F-tests are shown using both asymptotic and wild cluster bootstrap (10,000 reps., Cameron, C., Gelbach, J., and Miller, D., 2008). $^{***}p < 0.01$, $^{**}p < 0.05$, $^{*}p < 0.10$.

Figure B.5: Baseline Number of Purchase Orders (P.O.) between February 2020 and May 2020

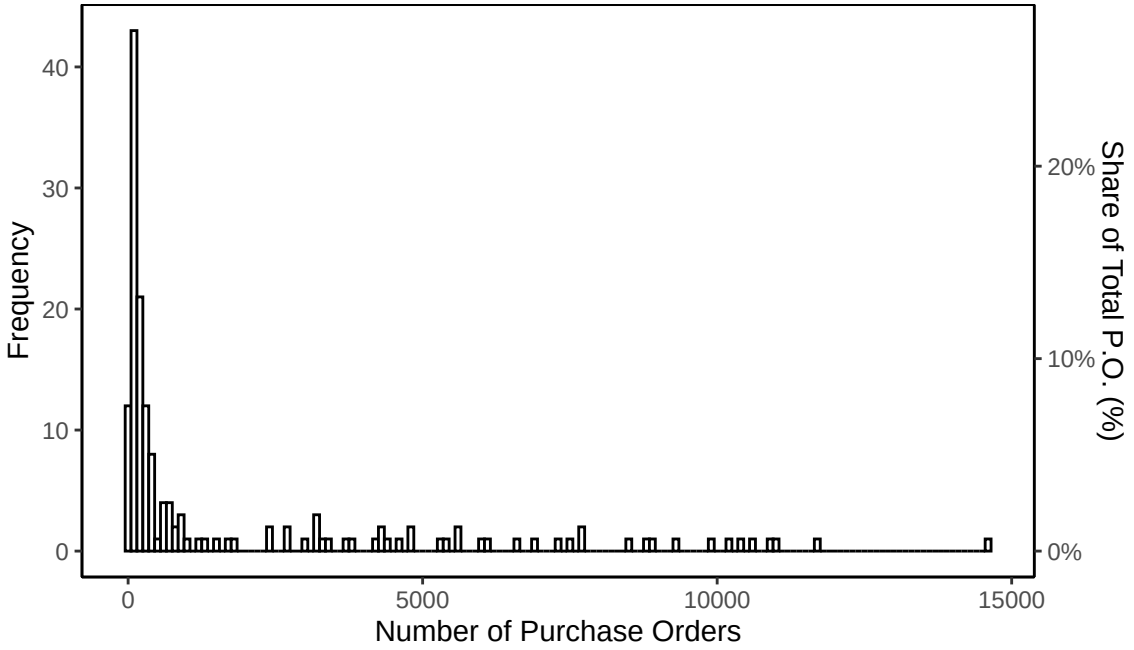


Table B.5: Extensive and Intensive Margin Effects of being exposed to Training Video.

	Intensive Margin (FA only)			Extensive Margin
	Overprice	Log(Q)	#P.O.	% P.O. in FA
	OLS	OLS	Poisson	OLS
	(1)	(2)	(3)	(4)
Training Video	-0.001	-0.053	-0.017	-0.017
	(0.004)	(0.096)	(0.500)	(0.014)
No. Observations	24,340	24,340	1,128	115,583
No. Buyers	1,128	1,128	1,128	4,699
Control Mean	0.103	3.448	4.840	0.360

Notes: This table shows the intention-to-treat effects of being exposed to the Training Video. Both Public and Private assigned-to-treatment groups were exposed to the very same Training Video, thus we report the training effects through a single dummy that equals one if the purchase was made in an assigned-to-treatment unit (either Public or Private) and zero otherwise. The Training Video effects are estimated for the two-month post-training-pre-reports period (July to August 2020). Intensive margin outcomes all refer to purchases made through the FA mechanism, which naturally limit the number of buyers and procurement units under consideration. Regressions in models (1) and (2) are at the unit-transaction level, and consider only purchases made through Food FA, Office Supplies FA, and Computers FA, and for which overprice can be calculated (i.e., we can find a reference group of products). Following equation 1, overprice of a unit transaction is defined as the relative difference between the per-unit price paid and the lowest price in the daily price distribution for products from the reference group. Log(Q) stands for the amount of items purchased per unit transaction (in logs). In buyer-level model (3), estimates are derived through a Poisson regression model, and #P.O. counts the number of Purchase Orders made by a buyer, which may contain more than one product (wholesales). Model regressions (1)-(3) control for strata fixed effects and the average outcome in procurement unit during the pre-intervention period. Since models (1) and (2) are at the unit-transaction level, we also control for product type fixed effects and week fixed effects. The extensive margin outcome (column 4) refers to the share of Purchase Orders made through FA (Framework Agreement) instead of alternative procurement mechanisms like *Procurement Auctions* or *Direct Purchases*. Outcome includes the universe of P.O. made by buyers in either of the three purchasing mechanisms, and the regression controls for strata fixed effects and the average outcome in the procurement unit during the pre-intervention period (Feb. - May 2020). In all, standard errors clustered at the procurement unit level are shown in parentheses. $^{***}p < 0.01$, $^{**}p < 0.05$, $^{*}p < 0.10$.

Table B.6: Extensive and Intensive Margin Effects on Placebo Buyers

	Intensive Margin (FA only)			Extensive Margin
	Overprice	Log(Q)	#P.O.	% FA
	OLS	OLS	Poisson	OLS
	(1)	(2)	(3)	(4)
Public Reports	0.018	0.015	2.383	0.023
	(0.018)	(0.126)	(3.261)	(0.022)
Private Reports	0.003	0.156	2.694	0.014
	(0.007)	(0.123)	(2.049)	(0.023)
No. Observations	61,928	61,928	1,285	3,124
No. Buyers	1,285	1,285	1,285	3,124
Control Mean	0.103	3.408	9.220	0.419
p-val. Public=Private	0.351	0.319	0.899	0.709

Notes: This table shows the intention-to-treat effects for placebo buyers, i.e., it compares purchases made by placebo buyers in Public and Private treatment groups with purchases made by control group buyers, considering the full post-treatment period of analysis (July 2020 to January 2021). Every month, placebo buyers received a message indicating that her overspending is being monitored, but do not provide any type of information about individual performance. The extensive margin outcome refers to the share of purchases per buyer made through FA (Framework Agreement) instead of alternative procurement mechanisms like *Auctions* or *Direct Purchases*. Intensive margin outcomes all refer to purchases made through the FA mechanism, which naturally limit the number of buyers and procurement units under consideration. Regressions in models (1) and (2) are at the unit-transaction level, and consider only purchases made through Food FA, Office Supplies FA, and Computers FA, and for which overprice can be calculated (i.e., we can find a reference group of products). Following equation 1, overprice of a unit transaction is defined as the relative difference between the per-unit price paid and the lowest price in the daily price distribution for products from the reference group. Log(Q) stands for the amount of items purchased per unit transaction (in logs). In buyer-level model (3), estimates are derived through a Poisson regression model, and #P.O. counts the number of Purchase Orders made by a buyer, which may contain more than one product (wholesales). Model regressions (1)-(3) control for strata fixed effects and the average outcome in procurement unit during the pre-intervention period. Since models (1) and (2) are at the unit-transaction level, we also control for product type fixed effects and week fixed effects. The extensive margin outcome (column 4) refers to the share of Purchase Orders made through FA (Framework Agreement) instead of alternative procurement mechanisms like *Procurement Auctions* or *Direct Purchases*. Outcome includes the universe of P.O. made by buyers in either of the three purchasing mechanisms, and the regression controls for strata fixed effects and the average outcome in the procurement unit during the pre-intervention period (Feb. - May 2020). In all, standard errors clustered at the procurement unit level are shown in parentheses. $^{***}p < 0.01$, $^{**}p < 0.05$, $^{*}p < 0.10$. Bottom row shows the p-value corresponding to the null hypotheses of no differences between Public and Private treatment effects.

Table B.7: Baseline Survey Questions

Question
1. Suppose that the unit in which you work fails to execute 100% (or a percentage close to 100%) of the budget allocated for this year. Regardless of the economic projections that DIPRES manages for the following year, what would you expect to happen with the budget approved for your unit the following year? *Alternatives: 1) I would expect the unit's budget to decrease from this year's budget, 2) I would expect the unit's budget to remain the same as this year's budget, 3) I would expect the unit's budget to increase from this year's budget.*
2. This question is about what you think the head of purchases of the unit in which you work believes. What do you think he or she would answer about what would happen to next year unit's budget if 100% of the budget (or close to it) allocated for this year is not executed? *Alternatives: 1) I think he/she believes the unit's budget would decrease next year, 2) I think he/she believes the unit's budget would stay the same next year, 3) I think he/she believes the unit's budget would increase next year.*
3. Suppose that the unit in which you work does not manage to execute 100% (or a percentage close to 100%) of the budget allocated for this year, but manages to demonstrate that it generated considerable savings in the use of resources. Regardless of the economic projections that DIPRES manages for the following year, what would you expect to happen with the unit's budget the following year? *Alternatives: 1) If savings are generated, I would expect the unit's budget to decrease from this year's budget, 2) If savings are generated, I would expect the unit's budget to remain the same as this year's budget, 3) If savings are generated, I would expect the unit's budget to increase from this year's budget.*
4. This question is about what you think the head of purchases of the unit in which you work believes. What do you think he/she would answer about what would happen to next year unit's budget if the unit does not manage to execute 100% of the budget (or a close percentage) allocated for this year, but manages to demonstrate that it generated considerable savings in the use of resources? *Alternatives: 1) If savings are generated, I think he or she believes the unit's budget would decrease next year, 2) If savings are generated, I think he or she believes the unit's budget would remain the same next year, 3) If savings are generated, I think he or she believes the unit's budget would increase next year.*
5. Next year's budget is largely determined by how much of this year unit's budget is executed. *Alternatives: 1) Strongly disagree, 2) Disagree, 3) Neither agree nor disagree, 4: Agree, 5: Strongly agree*
6. If savings are generated, the unit's budget will be reduced next year. *Alternatives: 1) Strongly disagree, 2) Disagree, 3) Neither agree nor disagree, 4: Agree, 5: Strongly agree*
7. Executing the budget allocated to your unit involves making spending decisions and savings strategies. Some believe that the level of public savings is a relevant variable for DIPRES in the process of negotiating the budget for the following year. Others, on the other hand, believe that DIPRES has little interest in the level of public savings generated by each unit and does not consider this variable in the budget negotiation process. From your perspective, what is the level of relevance that DIPRES gives to the public savings generated by your unit when making budget allocation decisions? Being 1 "Not relevant" and 7 "Extremely relevant"
8. One of the main mechanisms through which the budget is executed is the Public Procurement System ("Auctions", "Framework Agreements" or "Direct Purchases"). Some believe that DIPRES monitors and supervises on a recurrent basis the purchases made by each unit. Others, on the other hand, believe that DIPRES does NOT monitor or supervise the purchases made by each unit. From 1 to 7, what do you think is the level of monitoring exercised by DIPRES on public purchases made by your unit? Being 1 "No monitoring" and 7 "High level of monitoring".
9. Utilizing the maximum unit's budget is a pressure. *Alternatives: 1) Strongly disagree, 2) Disagree, 3) Neither agree nor disagree, 4: Agree, 5: Strongly agree*
10. Sometimes purchases are made at high prices to comply with budget execution. *Alternatives: 1) Strongly disagree, 2) Disagree, 3) Neither agree nor disagree, 4: Agree, 5: Strongly agree*

C. Appendix: Model's Details

C.1. Set Up

Consider a procurement unit composed of a Manager and a Buyer. The Buyer is tasked with purchasing a standardized product at the lowest possible price. In the absence of corruption or active waste, passive waste (inefficiency) may still arise if the Buyer's search effort to seek out the lowest price is insufficient. The Manager can try to persuade the Buyer to put in more search effort through a series of managerial actions such as setting efficiency goals, holding regular meetings, supervising purchases, creating purchasing manuals, training programs, and performance-based rewards, to name a few. However, the manager only observes the price paid and not the Buyer's search effort.

Suppose that the Buyer is extrinsically motivated and derives no utility from low prices *per se*. In this case, he only supplies search effort in response to rewards for finding lower prices and managerial pressure. By default, he exerts minimal search effort and obtains a purchase price μ for the product[34]. The Manager can try to persuade the buyer to provide a higher search effort through a set of actions. Doing so can potentially lower the price down to a minimum of $p^m \leq \mu$.[35] Let $a \in [0, A]$ represent the intensity of managerial actions, with $c(a)$ being the cost of a, which we assume strictly increasing and convex.[36]

The effectiveness of the Manager's actions depends on the Buyer's responsiveness, captured by the Buyer's type $\theta_B \in \Theta \subseteq [0, 1]$. A Buyer with a higher θ_B is *more receptive to managerial persuasion*, making it more likely that the Manager's actions successfully induce the price p^m. The Manager persuades the Buyer to buy at price p^m with probability $P(\theta_B, a) = \theta_B \phi(a)$, with $\phi(a)$ strictly increasing and concave function of a.

We assume that θ_B is unobservable to the Manager, i.e., she does not know how responsive the Buyer is to her persuasive actions. However, the Manager can learn about θ_B through information $I \in [0, \infty)$ accumulated from the Buyer's purchase history and his efficiency performance. The more information the Manager accumulates, the more accurate is her estimate of θ_B, allowing her to better adjust actions. In the absence of any information, the Manager's best guess is the population average $\bar{\theta}$. Using information I, the Manager's belief about θ_B is given by:

$$\hat{\theta}_B(I) = \alpha(I)\bar{\theta} + (1 - \alpha(I))\theta_B \tag{B.1}$$

with $\alpha(I)$ strictly decreasing in I, $\alpha(0) = 1$ and $\alpha(I) \xrightarrow{I \to \infty} 0$. Thus, if the Manager has no information, the Manager's belief is $\bar{\theta}$. As information goes to infinity, $\hat{\theta}(I)$ converges to θ_B.[37]

[34] One can think of minimal effort as a single draw in a sequential search process: the Buyer samples one price from the distribution of available prices and accepts it, so the expected purchase price is the mean μ.

[35] In principle, the buyer can face a distribution of multiple prices for the sought product, yet for simplicity we assume the buyer decides between two potential prices, either μ or the minium price.

[36] The intensity of managerial actions could refer to the number of actions taken or the quality of those actions, yet for tractability we do not make a specific distinction of attributes of the actions.

[37] How information is used by the Manager to update her belief can be rationalized as follows. Suppose the Manager

C.2. The Manager's Problem

The Manager's objective is to minimize overspending net of the cost of persuasion effort. Let $v(p)$ be the Manager's utility from purchasing the item at price $p \in (p^m, \mu)$. We assume that $v(\cdot)$ is strictly decreasing and concave. Then, the Manager's payoff is given by:

$$u_m(a, \hat{\theta}_B(I)) = P(\hat{\theta}_B(I), a)\bar{v} + (1 - P(\hat{\theta}_B(I), a))\underline{v} - c(a), \tag{B.2}$$

with $\bar{v} = v(p^m)$, and $\underline{v} = v(\mu)$.

Buyers are required to follow a purchasing protocol, and at the minimum, Managers are required to oversee that the Buyer correctly follows this protocol. Thus, we will assume that the Manager must exert at least a minimum intensity of managerial actions $a_0 = \arg\max\{u_m(a, \bar{\theta}) \mid a \geq 0\}$. Thus, the Manager's problem is defined as:

$$a^*(I) \in \arg\max\{u_m(a, \hat{\theta}_B(I)) \mid a \geq a_0\}, \tag{B.3}$$

The Manager's optimal action intensity depends on her belief about the Buyer's type, $\hat{\theta}_B(I)$. If the Buyer's true type is lower than $\bar{\theta}$, then she will realize that the Buyer is less responsive to her actions. Therefore, her optimal action profile will not change, $a^*(I) = a_0$. Conversely, if $\theta_B > \bar{\theta}$, new information reveals that the Buyer is more responsive than the Manager initially believed, and hence will optimally increase the intensity of her actions above the minimum a_0.[38]

The Manager's utility function can be written as:

$$u_m(a, \hat{\theta}_B(I)) = (\bar{v} - \underline{v})\hat{\theta}(I)\phi(a) + \underline{v} - c(a).$$

Let $I > 0$ be the Manager's updated information, and let $a^* = a^*(I, \theta_B)$ the solution for the Manager's problem. First order conditions imply:

$$\frac{\partial u_m(a^*(I), \hat{\theta}_B(I))}{\partial a} = (\bar{v} - \underline{v})\hat{\theta}(I)\phi'(a^*) - c'(a^*) = -\lambda, \tag{B.4}$$

where $\lambda \geq 0$ is the Lagrange multiplier associated with the constraint $a - a_0 \geq 0$. The assumptions over ϕ and c imply that $a_0 > 0$ and thus

$$\frac{\partial u_m(a_0, \bar{\theta})}{\partial a} = (\bar{v} - \underline{v})\bar{\theta}\phi'(a_0) - c'(a_0) = 0.$$

holds a Normal prior $\theta_B \sim \mathcal{N}(\bar{\theta}, \sigma_0^2)$ and each period observes a noisy signal $s_t = \theta_B + \varepsilon_t$, with $\varepsilon_t \sim \mathcal{N}(0, \sigma_\varepsilon^2)$ i.i.d. After accumulating I signals with sample mean $\bar{s}_I$, the posterior yields $\hat{\theta}_B(I) = \alpha(I)\bar{\theta} + (1 - \alpha(I))\bar{s}_I$, where $\alpha(I) = \sigma_\varepsilon^2/(\sigma_\varepsilon^2 + I\sigma_0^2)$. Notice that the posterior mean depends on the realized signals $\bar{s}_I$, which are stochastic. As $I \to \infty$, $\bar{s}_I \to \theta_B$ almost surely. Equation B.1 is a simplification that keeps the model tractable and is sufficient to capture the key dynamic: as the Manager accumulates more information, her belief transitions from the prior $\bar{\theta}$ toward the Buyer's true type θ_B, allowing her to better tailor her actions over time.

[38] To avoid uninteresting cases, we assume that $\lim_{a\to 0} \phi'(a) = +\infty$, $\lim_{a\to\infty} \phi'(a) = 0$, $\lim_{a\to 0} c'(a) = 0$, and $\lim_{a\to+\infty} c'(a) = +\infty$.

First, notice that if $\theta_B = \bar{\theta}$, then $\hat{\theta}(I) = \hat{\theta}(0) = \bar{\theta}$, which implies that $a^*(I) = a_0$. Suppose now that $\theta_B < \bar{\theta}$. It follows that

$$\frac{\partial u_m(a^*(I), \hat{\theta}_B(I))}{\partial a} < \frac{\partial u_m(a^*(I), \bar{\theta})}{\partial a} < \frac{\partial u_m(a_0, \bar{\theta})}{\partial a} = 0$$

Therefore, $\lambda > 0$. By complementary slackness, we conclude that $a^* = a_0$. Hence, if $\theta_B \leq \bar{\theta}$, the Manager will maintain her baseline level of effort.

Now consider the case $\theta_B > \bar{\theta}$. Then, $\frac{\partial u_m(a_0, \hat{\theta}_B)}{\partial a} > 0$. First order conditions, and strict concavity of u_m in a imply that $a^*(I, \theta_B) > a_0$. In addition, the partial cross derivative of the Lagrange operator is given by:

$$\frac{\partial^2}{\partial a \partial I} u_m(a, \hat{\theta}_B(I)) \quad = \quad (\bar{v} - \underline{v})\phi'(a)\frac{\partial}{\partial I}\hat{\theta}_B(I). \tag{B.5}$$

Since $\theta_B > \bar{\theta}$, $\frac{\partial}{\partial I}\hat{\theta}(I, \theta_B) > 0$. Therefore, by Topkis Theorem, Manager's persuasion actions $a^*(I, \theta_B)$ are increasing in I if $\theta_B > \bar{\theta}$. We summarize the previous conclusion in the following predictions:

C.3. Exogenous efficiency measure

Once the action profile $a^*(I)$ is chosen, the probability of achieving the lowest price depends on the true Buyer's type θ_B. Let $f : \mathbb{R}_+ \to \mathbb{R}$ be an exogenous efficiency measure such that $f(p)$ is the efficiency of a purchase of the item at price $p \in (p^m, \mu)$, with f strictly decreasing in p. (For instance, $f(p)$ could be defined as the overprice of the purchase: $f(p) = -(p - p^m)/p^m$.) Then, the purchase efficiency from Manager's actions profile is given by:

$$\pi(I) = P(\theta_B, a^*(I))f(p^m) + (1 - P(\theta_B, a^*(I)))f(\mu). \tag{B.6}$$

It follows that the effect of information in purchase efficiency is driven solely by Manager's actions. Purchase efficiency $\pi(I)$ is constant in I for low types ($\theta_B < \bar{\theta}$) and strictly increasing in I for high types ($\theta_B > \bar{\theta}$). Indeed, the expected purchase efficiency from Manager's efforts can be written as:

$$\pi(I) = [f(p^m) - f(\mu)]\theta_B\phi(a^*) + f(\mu)$$

If $\theta_B \leq \bar{\theta}$, for any $I > 0$ the Manager will exert the baseline level of persuasion actions a_0. Thus, in this case $\pi(I)$ is constant in I. If $\theta_B > \bar{\theta}$, then $a^*(I) > a_0$, and:

$$\frac{\partial \pi(I)}{\partial I} = [f(p^m) - f(\mu)]\theta_B\phi'(a^*(I))\frac{\partial a^*(I)}{\partial I} > 0$$

Thus if $\theta_B > \bar{\theta}$ and $I > 0$, then $\pi(I)$ is strictly increasing in I.

C.4. Price dispersion and potential reduction in paid price

We model the Buyer's search process as a sequential search (Weitzman, 1979). Each search is a random sample of a price at cost $\kappa > 0$ from a distribution $G(\sigma)$ with support $\Gamma = [p^m, \bar{p}]$. Let μ and σ^2 the mean and variance of $G(\sigma)$. By default, absent any managerial action, the Buyer exerts a minimal effort and samples only once, yielding the expected price μ. The larger is the number of searches the lower is the expected price obtained from searching. The Manager's actions persuade the buyer to increase the number of searches such that minimum price p^m is obtained.

The return to searching depends on the price dispersion, captured by the variance of the price distribution σ^2. When the distribution G widens, the best available price p^m falls. Thus, the Manager has more incentives to persuade the Buyer to purchase at p^m, implying the intensity of her managerial actions is increasing in price dispersion σ^2. As both p^m and a^* can change with σ^2, then efficiency $\pi(I)$ indirectly depends on σ^2. The cross partial derivative of the Manager's utility is given by:

$$\frac{\partial^2 u_m}{\partial a \partial \sigma} = \hat{\theta}_B \phi'(a) \left(v'(p^m) \frac{dp^m}{d\sigma} \right).$$

Since $v'(p^m) < 0$, and p^m is strictly decreasing in σ, the Manger's persuasion action $a^*(I)$ is strictly increasing in σ, for all $I \geq 0$. This allows us to analyze the effect of price dispersion on purchase efficiency. Consider the Overprice measure $f(p) = -(p - p^m)/p^m$ as a definition of efficiency $\pi(I)$. Then,

$$\pi(I) = (1 - \theta_B \phi(a^*(I)) \left(1 - \frac{\mu}{p^m} \right),$$

from which we obtain the partial derivative with respect to σ,

$$\frac{\partial \pi(I)}{\partial \sigma} = (1 - \theta_B \phi(a^*(I)) \frac{\mu}{(p^m)^2} \frac{dp^m}{d\sigma} - f(\mu) \theta_B \phi'(a^*(I)) \frac{\partial a^*(I)}{\partial \sigma}.$$

Hence, the magnitude of the effect of information on purchase efficiency π is also shaped by price dispersion, through similar mechanisms. On the one hand, information raises the baseline probability that the Manager successfully persuades the Buyer, it reduces the losses generated by increasing the gap between μ and p^m. On the other hand, the effect of information on π is also shaped by σ^2 through the Manager's actions through two opposite effects. Indeed,

$$\begin{aligned} \frac{\partial \pi(I)}{\partial \sigma} - \frac{\partial \pi_0}{\partial \sigma} &= \frac{\theta_B \mu}{(p^m)^2} \frac{dp^m}{d\sigma} (\phi(a_0) - \phi(a^*(I))) \\ &+ f(\mu) \theta_B \left(\phi'(a_0) \frac{\partial a_0}{\partial \sigma} - \phi'(a^*(I)) \frac{\partial a^*(I)}{\partial \sigma} \right) \end{aligned}$$

The first term is unambiguously positive, since price dispersion increases the Manager's incentive to exert further actions in response to new information, since dispersion widens the gap $\bar{v} - \underline{v}$. However, due to diminishing marginal returns, information may not raise the intensity of her actions by as much in a higher-dispersion setting as it would in a lower-dispersion setting. As a result, whether the effect of information on purchase efficiency is strengthened or weakened by price dispersion depends on which of these two forces dominates. If the marginal return of increasing further a^* is large enough, the gains in efficiency from increasing information are larger when price dispersion is higher.

The effect of information on efficiency is affected by the potential price reduction from persuasion: $\Delta p = \mu - p^m$ through an equivalent mechanism. The gap Δp is reduced whenever Buyers purchase closer to p^m in absence managerial actions. That is, when μ moves closer to p^m. If μ is closer to p^m there is less room left for the Manager to improve upon the untreated outcome. Conversely, the gap widens as μ moves further from p^m, leaving more room for the Manager's actions to matter. The direct effect of increasing μ on π, and its effect on the efficiency gains from information are given by:

$$\frac{\partial \pi(I)}{\partial \mu} = (1 - \theta_B \phi(a^*(I)) f'(\mu) - \theta_B \phi'(a^*(I)) \frac{\partial a^*(I)}{\partial \mu} f(\mu).$$

$$\begin{aligned} \frac{\partial \pi(I)}{\partial \mu} - \frac{\partial \pi_0}{\partial \mu} &= \theta_B f'(\mu)(\phi(a_0) - \phi(a^*(I))) \\ &+ f(\mu)\theta_B \left(\phi'(a_0) \frac{\partial a_0}{\partial \mu} - \phi'(a^*(I)) \frac{\partial a^*(I)}{\partial \mu} \right) \end{aligned}$$

This is an equivalent channel through which price dispersion operates: dispersion widens the potential gain by lowering p^m. As with dispersion, a wider potential gain raises the Manager's incentive to implement persuasion actions, but diminishing returns to managerial actions mean this incentive may not translate into much additional persuasion when the Manager is already acting in response to information. Hence, the effect of both, price dispersion and Δp, are generally ambiguous.

C.5. Intrinsically motivated Buyers

Thus far, the Buyer is extrinsically motivated, as he only searches for low prices if the Manager undertakes managerial actions. Suppose now that the Buyer has no extrinsic motivation but intrinsically values low prices by $u_b(p) = -p$. To search for low prices, the buyer sequentially samples prices from a distribution G with support in $\Gamma = [p^m, \bar{p}]$ [39]. The expected price obtained from sampling x times prices is $p(x) = \mathbb{E}[\min\{p_1, \ldots, p_x\}]$, where $p_1, \ldots, p_x$ are i.i.d. draws from G.

[39] We build our search model on Weitzman (1979)

Importantly, when the Buyer is only intrinsically motivated, Manager's persuasion actions have no effect on the Buyer's search decision, implying information I acquired by the manager has zero effect on the purchased price through Manager's actions. Therefore, Predictions **??** and **??** do not hold in this case. The Buyer follows an optimal stopping rule characterized by a reservation price r, accepting the first sampled price below r. The optimal reservation price satisfies:

$$\kappa = \int_{p^m}^{r} (r - p)\, dG(p),$$

where G is the distribution of prices. However, if a large number of sellers participate in this market, available options can exhibit a wide range of prices. Further, sellers adjust their prices over time, and the Buyer must actively search to identify the cheapest option. Without guidance, the Buyer can only sample from a limited range of available prices, and potentially missing lower prices. We assume that Buyer has incomplete information about the true price distribution G. However, information helps the Buyer to find lower prices and adjust his reservation price accordingly. The Manager can share information $I \geq 0$ that informs the Buyer about the true price dispersion, expanding the range of prices the Buyer perceives as accessible. Knowing the availability of lower prices allows the Buyer to optimally decrease his reservation price.

To model the role of information in searching, consider a location-scale family of random variables $\mathcal{G} = \{p(\gamma) = \mu + \gamma Z : \sigma_0 \leq \gamma \leq \sigma\}$, with $Z \sim \varphi$ is a standardized random variable with bounded support. The family $\mathcal{G}$ comprises price distributions with mean μ, and standard deviation γ. As γ increases, the distribution preserves its shape but spreads further around μ. Assume that the true distribution of prices is $p(1) = \mu + \sigma Z \sim G$, with support $\Gamma = [p^m, \bar{p}]$. Let $\Gamma(\gamma)$ denote the support of $p(\gamma)$. It follows that $\Gamma(\gamma) = [\underline{p}(\gamma), \bar{p}(\gamma)]$ with $\underline{p}(\gamma) = \mu + \gamma(p^m - \mu)$ and $\bar{p}(s) = \mu + \gamma(\bar{p} - \mu)$. Therefore $\{\Gamma(\gamma)\}_{\gamma \in [\sigma_0, \sigma]}$ is a nested family of supports, $\Gamma(\gamma) \subsetneq \Gamma(\gamma')$ for all $\gamma < \gamma'$. Define:

Define $\gamma(I) = \sigma_0 \left(\frac{1}{1+I}\right) + \sigma \left(\frac{I}{1+I}\right)$, with $0 < \sigma_0 < \sigma$. The value $\gamma(I)$ captures the fraction of price dispersion accessible to the Buyer, while σ captures the standard deviation of the true distribution of prices. When $\gamma(I)$ is small, prices are concentrated near the mean μ and the Buyer accesses to limited price options. As $\gamma(I)$ increases from σ_0 toward σ, the Buyer gains access to a wider range of prices, including lower prices further below the mean. The nested structure of supports reinforces this intuition. If $\gamma(I)$ is larger, the Buyer faces a strictly richer price distribution. Hence, $\gamma(I)$ reflects the quality of the Buyer's information about the market. Better information allows the Buyer to identify options with greater price dispersion, increasing the potential gains from search. The true distribution G is recovered at $\lim_{I \to \infty} \gamma(I) = \sigma$, representing the case of full information.

Denote $G(p|\gamma)$ the distribution of $p(\gamma)$. Let $r(\gamma)$ be the reservation price with respect to the distribution $G(p|s)$. We assume that κ is such that $r(\gamma(0)) < \mu$. This means that the cost of searching κ is small enough to allow the Buyer to be selective, only accepting prices below the mean. This assumption is reasonable in markets where the number of available prices is large,

and the cost of each sample is low, such as online markets. While the information affects the reservation price of the Buyer, the prices are actually sampled from the true underlying distribution G. Thus, the expected price obtained by the Buyer is given by:

$$p^*(\gamma) = \mathbb{E}[p|p \leq r(\gamma)] = \frac{\int_{p^m}^{r(\gamma)} pdG(p)}{G(r(\gamma))}.$$

We will show that the reservation price $r(\gamma)$ and the expected purchase price $p^*(\gamma)$ are decreasing in γ, and therefore in information I. Notice that:

$$\kappa = [pG(p|\gamma)]_{\underline{p}(\gamma)}^{r(\gamma)} - \int_{\underline{p}(\gamma)}^{r(\gamma)} pdG(p|\gamma) = \int_{\underline{p}(\gamma)}^{r(\gamma)} G(p|\gamma)dp.$$

Since $p(\gamma) = \mu + \gamma Z$,

$$\kappa = \int_{\underline{p}(\gamma)}^{r(\gamma)} \varphi\left(\frac{p-\mu}{\gamma}\right) dp = \gamma \int_{\frac{\underline{p}(\gamma)-\mu}{\gamma}}^{\eta(\gamma)} \varphi(t)dt,$$

with $\eta(\gamma) = \frac{r(\gamma)-\mu}{\gamma}$. Then, differentiating with respect to γ:

$$0 = \int_{\frac{\underline{p}(\gamma)-\mu}{\gamma}}^{\eta(\gamma)} \varphi(t)dt + \gamma\left[\varphi(\eta(\gamma))\eta'(\gamma) + \varphi\left(\frac{\underline{p}(\gamma)-\mu}{\gamma}\right)\frac{d}{d\gamma}\left\{\frac{\underline{p}(\gamma)-\mu}{\gamma}\right\}\right].$$

Since $\frac{\underline{p}(\gamma)-\mu}{\gamma} = p^m - \mu$, $\varphi\left(\frac{\underline{p}(\gamma)-\mu}{\gamma}\right) = 0$. It follows that:

$$0 = \int_{\frac{\underline{p}(\gamma)-\mu}{\gamma}}^{\eta(\gamma)} \varphi(t)dt + \gamma\varphi(\eta(\gamma))\eta'(\gamma).$$

From which we get and expression for $\eta'(\gamma)$:

$$\eta'(\gamma) = -\left(\frac{1}{\gamma\varphi(\eta(\gamma))}\right)\int_{\frac{\underline{p}(\gamma)-\mu}{\gamma}}^{\eta(\gamma)} \varphi(t)dt.$$

Therefore, $\eta'(\gamma) < 0$. Furthermore, by definition of $\eta(\gamma)$,

$$r'(\gamma) = \eta(\gamma) + \gamma\eta'(\gamma) < 0.$$

Hence, $r(\gamma)$ is decreasing in γ. Moreover, higher dispersion also translates in smaller expected price $p^*(\gamma)$. This follows directly from the reservation price condition. Indeed,

$$p^*(\gamma) = r(\gamma) - \frac{\kappa}{G(r(\gamma))},$$

which implies that $p^*(\gamma)$ is decreasing in γ. Thus, the higher the information, the lower is the reservation price. This directly implies that the expected price $p^*(\gamma(I))$ is strictly decreasing in I.

Furthermore, in the absence of extrinsic motivation, the Manager's actions have no effect on the Buyer's effort, so $P(\theta_B, a) = 0$. The Manager's payoff is:

$$u_m(a, \hat{\theta}_B(I)) = v(p^*(\gamma(I))) - c(a) \Rightarrow a^*(I) = a_0.$$

Hence, purchase efficiency is driven solely by the Buyer's search effort, $\pi(I) = f(p^*(\gamma(I)))$, and is increasing in information I. The expected purchase price $p^*(\gamma(I))$ is therefore a strictly decreasing function of information I, which in turn determines purchase efficiency $\pi(I) = f(p^*(\gamma(I)))$.

The same result extends to price dispersion of the underlying true distribution G, σ: greater price dispersion leads to a lower expected purchase price, as the Buyer faces a wider range of prices and becomes more selective in his search. Analogously to the potential price reduction from persuasion in the extrinsic model, define $\Delta p = p^*(\sigma_0) - p^*(\sigma)$ as the potential reduction in price available to the Buyer, the gap between the price attainable under full information and the price attainable at baseline. Since $\frac{\partial \gamma}{\partial \sigma}, \frac{\partial \gamma}{\partial \sigma_0} > 0$, the expected purchase price $p^*(\gamma)$ is decreasing in both σ and σ_0, so Δp is positive and shrinks as the buyer's baseline dispersion σ_0 approaches the true dispersion σ.

Price dispersion σ and the gap Δp also shapes how effective information is. A wider gap between σ and σ_0 tends to make information more valuable in lowering the reservation price. However, these effects are not unambiguous:

$$\frac{\partial^2 r(\gamma)}{\partial I \partial \sigma_0} = r''(\gamma)\frac{\partial \gamma}{\partial I}\frac{\partial \gamma}{\partial \sigma} + r'(\gamma)\frac{\partial^2 \gamma}{\partial I \partial \sigma}.$$

The term $r''(\gamma)$ represents how sharply the reservation price responds to further increases in γ, an object we cannot sign without further assumptions on the distribution φ. The same ambiguity carries over to the expected purchase price p^*.

C.6. Extrinsically and Intrinsically motivated Buyers

When the Buyer is both intrinsically and extrinsically motivated, information plays a dual role that creates a tension in the Manager's optimal response. The Manager has incentives to take actions as long as $p^*(I) > p^m$ for any given $I > 0$. Information helps the Manager learn that a high-type Buyer is more responsive to her actions, raising the marginal return to extrinsic effort. In addition, more information also reveals that the Buyer is already exerting search effort on his own, reducing the marginal gain from the Manager's intervention. How much the Manager increases her effort as information grows depends on which effect dominates. If intrinsic motivation is strong enough, the Buyer's search effort drives the purchase price sufficiently close to p^m, making it suboptimal for the Manager to undertake costly persuasion actions, even when θ_B is high.

The objective function of the Manager is given by:

$$u_m(a, \hat{\theta}_B(I)) = [v(p^m) - v(p^*(\gamma))]\hat{\theta}(I, \theta_B)\phi(a) + v(p^*(\gamma)) - c(a).$$

Let $a^* = a(I)$ the optimal Manager persuasive effort. Notice that if $p^m > p^*(I)$, then $a^*(I') = a_0$ for any $I' > I$. Then

$$\frac{\partial^2}{\partial a \partial I} u_m(a, \hat{\theta}_B(I)) = [v(p^m) - v(p^*(\gamma))]\phi'(a)\frac{\partial \hat{\theta}(I, \theta_B)}{\partial I} - \hat{\theta}(I, \theta_B)\phi'(a^*)v'(p^*(\gamma)(p^*(\gamma))'\frac{\partial \gamma}{\partial I}.$$

The first term captures the effect of information from Manager's actions (extrinsic motivation), while the second term captures the effect of information from the marginal gain due to search effort (intrinsic motivation). By Topkis' Theorem, whether the Manager's actions are increasing in information I will be determined by the magnitude of the mentioned effects. If the effect of information on the marginal utility of Buyer's search effort is large enough, the Manager will choose a lower level of actions.

Further, in this context, the purchase efficiency is:

$$\pi(I) = \theta_B \phi(a^*)[f(p^m) - f(p^*(\gamma))] + f(p^*(\gamma)).$$

Thus, the effect of information will also depend on the relative importance of the intrinsic and extrinsic motivation. Indeed, the partial derivative of $\pi(I)$ with respect to I is given by:

$$\frac{\partial \pi(I)}{\partial I} = \theta_B \phi'(a^*)[f(p^m) - f(p^*(\gamma))]\frac{\partial a^*}{\partial I} + (1 - \theta_B \phi(a))f'(p^*(\gamma))(p^*(\gamma))'\frac{\partial \gamma}{\partial I}$$

Thus, the effect of information on purchase efficiency is given by its effect on Manager's action and its effect on Buyer's search efforts. Notice that while the purchased price from Buyer's intrinsic motivation $p^*(\gamma)$ is decreasing in I, the effect of information can deter the Manager's actions for highly motivated Buyers.

The same tension arises with respect to price dispersion. A more dispersed market improves the Buyer's search outcomes directly, as a higher dispersion parameter σ lowers the expected purchase price. Thus, the effects of both information and price dispersion on purchase efficiency are in general ambiguous, as each improves efficiency through the Buyer's own search but may simultaneously discourage the Manager's intervention. Hence, Predictions **??**–**??** do not hold in general when the Buyer has both intrinsic and extrinsic motivation. Nevertheless, since Manager's actions are bounded below, any increase in information yields weakly higher efficiency compared to the baseline, so $\pi(I) > \pi(0)$ for any $I > 0$. In contrast, no such guarantee holds for price dispersion when the Manager has information $I > 0$: a market with higher dispersion σ may induce lower Manager persuasion actions $a^*(I)$ compared to a market with lower dispersion, offsetting the direct gains from the Buyer's search and implying that the effect on purchase efficiency is

ambiguous.