Governance of supply relationships: Evidence from Indian manufacturing

Aroon Narayanan
MIT

November 25, 2025

Please click here for the latest version

Abstract

How should firms dynamically manage a supplier portfolio to improve contracting outcomes? This paper analyzes a large Indian buyer's policy of promoting high-performing suppliers to higher tiers that receive increased order volume. I estimate a structural dynamic principal-agent model to quantify the policy's value and disentangle its two mechanisms: retaining better suppliers by sorting them based on a persistent type, and mitigating moral hazard by using relational value and induced competition to incentivize effort. A novel two-stage estimation strategy first recovers the buyer's policy using an iterative algorithm to handle unobserved supplier tiers, then estimates a dynamic effort choice model of supplier behavior conditional on this policy. The policy improves performance by 12% over random allocation, with selection effects dominating incentive effects (66% vs. 34% of the total value). The economic value of this relational approach is significant—achieving the same performance improvement in a spot-market benchmark would require 18% higher payments to suppliers. Counterfactual analysis shows that a policy optimized for the estimated environment could improve performance by 28%, but the firm's current policy is more robust to facing a less capable supplier pool.

I am grateful to Nikhil Agarwal, Robert Gibbons, Tavneet Suri, and Michael Whinston for their support and guidance throughout this project. I also thank Isaiah Andrews, Charles Angelucci, Vivek Bhattacharya, Alessandro Bonatti, Roberto Corrao, Glenn Ellison, Andreas Haupt, Bengt Holmström, Deivy Houeix, Namrata Kala, Ameet Morjaria, Stephen Morris, Thi Mai Anh Nguyen, Vishan Nigam, Tobias Salz, Scott Stern, Benjamin Vatter and participants of the MIT Industrial Organization seminar, the MIT Organization Economics seminar, and the Annual Conference on Economic Growth and Development for helpful comments. I acknowledge support from the George and Obie Shultz Fund, the Jerry Hausman Graduate Dissertation Fellowship, and the Sloan Africa Fellowship. Email: aroon@mit.edu.

1 Introduction

Modern supply chains are actively governed.[1] A common governance strategy, adopted by companies like Toyota, GE, and Walmart, is to reward high-performing suppliers with more business.[2] This approach leverages two key mechanisms: incentivization and selection. By awarding long-term partnerships only to relatively high performers, firms induce a competitive environment that incentivizes suppliers to engage in surplus-enhancing activities. At the same time, firms encounter large heterogeneity in the fundamental capabilities of their possible suppliers; performance-based selection allows firms to find and keep the most capable suppliers.[3]

Despite the prevalence of this governance approach, there is limited empirical evidence on how much value it creates or the relative importance of its underlying mechanisms. This motivates my paper's central research questions: How much value is created by dynamically allocating more business to better-performing suppliers? What drives this value—do firms primarily benefit from identifying and retaining more capable partners, or from incentivizing higher performance? And how does it compare to the value from alternative contracting arrangements?

I study these questions in the context of an innovative supplier management system developed by a large new venture in the market for fabricated steel products in India. Operating in a setting with weak contract enforcement, the firm has relied on a relational system to develop and sustain a high-quality supplier base. The system implements a dynamic allocation policy through a simple three-tier ranking structure—Gold, Silver, and Bronze—where suppliers that perform better on quality checks are promoted to better tiers, which come with the explicit promise of more future orders. The analysis utilizes a uniquely detailed administrative dataset that covers all orders between the buyer and her supplier network from the beginning of the buyer's operations in 2019 to late 2021. The data's strength lies in its comprehensive view into the full history of relationships, and the outcomes that realize in them: it contains part-level quality inspection outcomes for over 73,000 individual components, alongside three snapshots in time of the tier assigned to each supplier. This allows me to precisely measure performance and link it to the supplier management policy implemented by the buyer.

To capture and quantify the policy's interacting mechanisms of selection and incentiviza-

[1] Survey evidence shows that supplier performance reviews are common, with a majority of firms conducting such evaluations at least once per year (see, e.g., Gartner, 2025; Pech and Vaněček, 2020).

[2] See Liang and Atkins (2021) for a discussion on the prevalence of dynamic business allocation.

[3] See Syverson (2011) for a survey of the literature that documents major, persistent differences in productivities across firms within the same industry, with a wider dispersion in developing countries.

tion, I estimate a structural dynamic principal-agent model. In the model, the principal, i.e. the buyer, establishes a policy that maps a supplier's performance in the current period to his tier in the following period, and then uses the realized tiers to specify the probability of receiving orders. The agents, i.e. the suppliers, who possess persistent performance types, choose a level of costly private effort as the solution to a dynamic programming problem, where the value of a higher tier tomorrow is weighed against the cost of effort today. Performance is realized as a function of both the persistent type and the chosen effort.

The estimation of this model proceeds in a novel two-stage approach that separates the recovery of the principal's policy from the estimation of agent behavior. This structure circumvents the need to take a stand on the buyer's feasible policy space and constraints (such as commitment), and avoids the computational burden of repeatedly solving for equilibrium during estimation. In the first stage, I estimate the parameters of the buyer's order-allocation rule and tier-transition rule by maximizing the likelihood of these parameters conditional on the supplier selections, realized performance, and the three observed tier snapshots. The intuition behind the identification of the policy is that the distribution of supplier selections as a function of performance is an informative signal of the buyer's policy. The policy's Markovian structure then allows the the allocation and transition rules to be jointly identified from the distribution of supplier selections. The main implementation challenge is that supplier tiers are observed only at the three snapshot times, making them a latent state between these observations. I overcome this by using a Monte Carlo Expectation-Maximization (MCEM) algorithm to integrate over the distribution of all possible latent tier paths.

In the second stage, I estimate the parameters of the supplier-side model—including the effort cost function and the distribution of supplier types in the outside pool—by Simulated Maximum Likelihood. This involves solving each supplier's dynamic problem to recover optimal efforts and the resulting pass/fail probabilities for each delivered part, integrating over unobserved supplier types and the latent tier paths implied by the estimated buyer policy, and then selecting the parameter values that maximize the likelihood of the observed performance histories. The panel nature of the data helps separate the effect of endogenous effort from supplier type. Since a supplier's type is persistent, within-supplier responses in performance to exogenous cost shocks and the policy's dynamic incentives identify the parameters of effort choice. Then, the distribution of across-supplier performance on first orders identifies the distribution of supplier types in the outside pool. The estimation of the supplier-side model leverages the fact that the buyer's dynamic incentive structure can be treated as known, which is a major advantage of the two-stage approach.

The estimates imply that the buyer's tiered relational policy places a significant premium

on performance—for example, a new supplier goes from practically zero probability of becoming a highest-tier Gold supplier at the 5-th percentile of performance, to being almost certain to become one at the 95-th percentile. This leads to a clear sorting of types into tiers, with higher tiers being populated with higher types. Gold suppliers are also more likely to be chosen for an order than the next tier Silver and the lowest-tier Bronze suppliers. For example, a Gold supplier is likely to be chosen with 50% higher probability on average than a Bronze supplier.

The tier system also functions as an incentive scheme: lower tier suppliers, facing a high risk of being removed from the supplier pool, exert the most effort. In contrast, higher tier suppliers are safer and thus exert less effort, but their higher underlying type ensures higher performance nonetheless. This further interacts with the overall level of internal competition among the suppliers induced by the buyer's policy. The estimates imply that supplier effort is generally an inverted U-shaped function of the number of active competitors, which is a result of opposing forces—the threat of replacement by rivals boosts effort but an oversaturation of suppliers dilutes the future value of the relationship. This result highlights that a unique feature of managing a system of relationships is that the buyer must carefully choose the level of internal competition. Moreover, it demonstrates that the competition-cooperation dynamic operates differently at the firm-level than in typical market settings, where increased competition often comes at the cost of reduced relational cooperation;[4] here, competition is a tool that the buyer deploys to increase cooperation.

Using the present discounted measure of performance on quality checks as the value created by the policy, I find that the buyer's chosen policy increases value by 12% relative to a baseline of random allocation. Decomposing this gain, I find that selection is the dominant mechanism, accounting for approximately 66% of the value added, with incentivization accounting for the remaining 34%. The policy's value relative to a spot benchmark with no relationships is economically significant: to achieve the same performance improvement in a series of one-shot interactions, the buyer would need to offer a monetary performance bonus equal to 1.6 times the average margin that a supplier gets on an order. This represents an 18% increase in total payments to suppliers, which provides a simple justification as to why the buyer has chosen to implement a relationship-based policy instead.

Since the two-stage estimation approach recovers the buyer's policy without imposing optimality, I am able to speak to optimal policy design in the estimated supplier environment, where the optimal policy maximizes the present discounted measure of performance taking into account supplier optimal responses to the policy. I find that the optimal policy is approximated by a dual-tier policy, where suppliers either attain the top tier or are removed from the pool,

[4]See, e.g., Macchiavello and Morjaria (2021) and Harris and Nguyen (2025).

with very steep rewards for performance. It improves performance by 28% over baseline, primarily by filling the top tier with very high types and then almost exclusively selecting the top tier for orders.

Interestingly, the optimal policy is not a cutoff policy, i.e., it doesn't simply use a performance cutoff to either keep or drop suppliers. The reason is that a cutoff policy provides very little incentive for effort to types that are further away to both sides of the chosen cutoff, which reduces the incentivization value of the policy. This is a critical distinction between a pure learning setting—a standard "bandit" environment—and this paper's model, where the bandit arms have endogenous responses to the policy. Cutoff policies are optimal in bandit environments, while in this setting, the optimal policy is smooth since it must strike the right balance between incentivization and selection.

Finally, I perform counterfactuals to understand why the buyer might be justified in deploying the estimated policy instead of the optimal one. The main takeaway from these is that in counterfactual environments where either the type distribution is significantly worse than the estimated one, or where the buyer has to draw initially from a worse distribution before getting access to more capable suppliers later on, the estimated policy can either close the gap or even perform better than the policy that was optimal in the estimated environment. Thus, one rationale for the chosen policy is robustness to the economic environment. Such a strategy could have been especially appealing if the buyer initially had high levels of uncertainty about the specific supplier environment that she was going to face.[5]

Related literature The model in this paper is related to theoretical work in Economics and Operations Management that explores the use of dynamic allocation policies in contracting. This work has shown that the promise of future business creates endogenous switching costs that prevent opportunism (Board, 2011), that dynamic allocation can make incentives for effort credible under private monitoring by tying an agent's maximum reward to their future value to the principal (Andrews and Barron, 2016), and that performance-based supplier rating can be combined with allocation incentives to improve supply chain functioning (Li, Zhang, and Fine, 2013). In contrast to the focus on moral hazard in these papers, the analysis in this paper focuses on a model that incorporates both selection and moral hazard.

The model also connects to two foundational literatures on incentive mechanisms: efficiency wages and tournaments. In the literature on efficiency wages, Shapiro and Stiglitz

[5]While the buyer could theoretically overcome this with a policy that continually updates as uncertainty resolves, this is logistically challenging because it takes significant time to implement and evaluate the results of any business strategy. Frequent changes can also potentially backfire by damaging trust with suppliers.

(1984) shows how rents attached to an employment relationship can induce effort via the threat of job loss. In the literature on tournaments, Lazear and Rosen (1981) demonstrates how a firm can use rank-order prize spreads to induce effort and sort workers by ability when output is noisy. In this paper's model, the tiers in the buyer's policy create continuation-value rents from future orders (the efficiency-wage channel) and function as rank-based prizes (the tournament channel).

There is a related empirical literature that studies relational sourcing strategies. There is work that shows suppliers forgo significant spot-market rents to preserve relationships (Macchiavello and Morjaria, 2015) and that increased competition erodes the rents needed to sustain relational cooperation (Macchiavello and Morjaria, 2021). Other papers provide reduced-form analysis of firm-level sourcing strategies: Cajal-Grossi, Macchiavello, and Noguera (2023) show that relational buyers systematically generate higher supplier markups, while Helper and Munasib (2022) show that organization-level strategies shape supplier governance more than product-level ones. Recent structural work has quantified the equilibrium trade-off between relationship-based sourcing and spot market transactions (Harris and Nguyen, 2025). This paper analyzes the design of a firm-level relational sourcing system, and highlights the role of internal competition in the supplier portfolio as a design tool.

This paper's methodology relates to the literature on two-stage methods for estimating dynamic structural models. These methods first estimate player policies from data and then recover structural parameters using techniques that circumvent the need to solve the full dynamic problem, such as policy function inversion or forward simulation (Hotz and Miller, 1993; Bajari, Benkard, and Levin, 2007; Pakes, Ostrovsky, and Berry, 2007; Aguirregabiria and Mira, 2007). Building on these, the two-stage method in this paper recovers the principal's policy from the data in the first stage, but then uses that to enable a full solution of the agents' optimization problems. It provides a template for dynamic games where one player moves first by choosing a policy (with possibly latent and endogenous policy states) and the counterparty responds to that policy: recover the first player's policy from her behavior, then estimate the other side's primitives conditional on that recovered policy. In particular, it can be applied more broadly in other dynamic principal-agent settings.

There is a related empirical literature on contracting with both adverse selection and moral hazard. Reduced-form work has provided evidence on the importance of selection in incentive contracts, showing how a shift to performance pay can sort workers by productivity (Lazear, 2000) and how contract choices can reveal private information in various settings (e.g., Ackerberg and Botticini, 2002; Einav et al., 2013; Ashraf et al., 2020). Recent structural work has utilized dynamic data, including papers separating inherently risky consumers from

careless ones in insurance markets (Bajari, Dalton, et al., 2014; Jeziorski, Krasnokutskaya, and Ceccarini, 2017). This paper's approach recovers the policy and its dynamic continuation value incentive from the data and uses it to disentangle effort from type.

A related literature in Operations Management studies supplier performance measurement and evaluation systems. Maestrini et al. (2018) use survey and dyadic data to examine how supplier performance measurement systems, monitoring, and incentives affect supplier performance and relational outcomes, while Jääskeläinen (2021) shows that the value created by performance information depends critically on relational governance and the supplier's strategic status. Zhou et al. (2022) models how evaluation-based development programs are allocated across suppliers to mitigate risk in a stochastic programming framework. This paper complements these studies by treating the evaluation-and-allocation rule itself as a design choice that shapes the long-run composition of the supplier base.

The remainder of the paper is organized as follows. Section 2 describes the institutional setting and the data. Section 3 presents some descriptive evidence that informs the structural model. Section 4 lays out the model. Section 5 details the structural estimation procedure. Section 6 presents the main results, and Section 7 analyzes the policy and presents counterfactuals. Section 8 concludes.

2 Setting and data

Institutional details I use comprehensive administrative datasets from one of the largest buyers of fabricated steel products in India. The buyer acts as an intermediary, contracting with mostly small and medium scale suppliers to deliver the products to downstream customers. The downstream customers contract with the buyer for the products that they need, and then the buyer contracts with the suppliers to procure them. The suppliers fabricate a wide range of products for the buyer, including core beams of residential and office buildings, bases and blades of wind turbines, mounts for solar panel arrays, parts of boilers for power plants, and railway tracks. Since the products need to conform to engineering specifications, the buyer inspects each of them to certify whether it meets or fails to meet the required specifications. The outcomes of these inspections are the fundamental measure of quality for the order; failed quality checks, whether due to a lack of readiness by the stipulated date or a failure to meet engineering specifications, can lead to significant and costly delays for the downstream customer. The equipment is delivered to the customer site once it passes inspections.

The buyer began operations in 2018, expanding quickly due to venture capital invest-

ment.[6] Over the next few years, she built up a base of suppliers with whom she maintains relationships. Given the smaller scope of operations of the suppliers, and given the lack of formal enforcement in India, the buyer influences outcomes primarily via active relationship management. In particular, the buyer awards active suppliers a status based on past performance - “Gold Supplier", “Silver Supplier" or “Bronze Supplier" - each of which comes with explicit promises about future consideration for orders. Suppliers can be upgraded, downgraded or completely dropped from the supplier pool depending on their performance on orders. New suppliers, who don’t have any performance history, are treated as a separate tier. Moreover, dropped suppliers are treated effectively as part of a separate, permanently inactive tier. When an order needs to be assigned, the buyer reaches out to active suppliers to check for availability and then assigns the order to an available supplier, with higher tiers given higher consideration. Suppliers charge a standard amount for specific kinds of orders. Importantly, the amount paid to a supplier for an order does not condition on the performance on that order.

Once an order is assigned, the suppliers take 1-2 months to deliver all parts of an order. In this time, they work on the different stages of production to convert raw steel sheets into the finished intermediate product, such as cutting, grinding, welding, painting, etc. Being mostly small scale enterprises, the techniques and machinery are basic and standard across most of the suppliers, with the differentiator being the skill of the workforce and the quality of the firm’s management of the workforce and equipment to ensure timely and competent delivery. The primary raw material for each order is steel, and a key characteristic of this production process is that the value of the steel that is required for an order is a large fraction of the total value of the order, being around 60% on average. Production is followed by inspection and delivery. Since transportation costs are extremely high due to the high weight of the fabricated equipment, orders are always assigned to suppliers in the same region as the downstream customer. The suppliers exhibit varying degrees of specialization and certification, ranging from none to being a certified manufacturer for the Indian Railways. These specializations and certifications further classify suppliers within a region into clusters. Suppliers within a cluster form the consideration set for any given order that I observe in the data.

Data The data fall into two categories. The first contains information on all the orders between the buyer and suppliers on fabrication of metal parts spanning from January 2019 to October 2021, while the second contains information on the assigned tiers of suppliers at three different points in time. I label each time period of observed tiers as a *snapshot*. The three

[6] This enabled them to build a reputation for timely payment to suppliers, which in turn enabled them to sign on suppliers quickly. In an industry dominated by one-on-one contracting between customers and suppliers, a large pool of suppliers enabled the buyer to convince customers to use their intermediated service instead.

snapshots correspond to August 2019, February 2020, and September 2020 respectively.

The records on orders detail the interactions between the buyer and approximately 270 suppliers in 9 clusters on close to 900 orders. Each order is composed of multiple parts, amounting to a total of approximately 74,000 parts across all orders in the dataset. In most cases, parts are individual manufactured items, such as a steel beam. Multiple parts are worked on, inspected, and delivered in lots. There can be multiple lots in an order. The data on orders includes the customer number, the supplier number, order number, part-level binary (0/1) inspection outcomes, lot-level steel purchase dates and delivery dates, payments made to the supplier, and order size (in metric tonnes, MT). To this dataset, I merge in steel price data at the month level obtained from the Indian Ministry of Steel. In addition, I observe the contract terms between the buyer and suppliers, which are decided at the beginning of the relationship and kept fixed thereafter. These mainly relate to when the supplier will be paid for orders and what auxiliary services will be provided by the buyer, such as credit or cash upfront.

Statistics on the main variables in the orders data are reported in Table 1. On average an order sees 79% of the parts in an order pass the first inspection - I define this metric as the *pass rate*, and this is the measure that I will use to denote performance. On average a supplier has a pass rate of 75%. There is also a fairly large variation in pass rates across both suppliers and orders. Statistics on the tiers dataset are reported in Table 2. The statistics are reported at the cluster level, since that is the level at which suppliers compete with each other for orders. The first snapshot sees an average of 9 active suppliers in a cluster, which drops to 8 and 4 in the later snapshots. Suppliers in each active tier decrease, but more so in the lower tiers. As expected, the ranks of the Dropped tier swell substantially across the snapshots.

3 Descriptive evidence

This section presents descriptive evidence from the data that motivates the subsequent structural model and highlights the key variation that will be leveraged for its estimation. The evidence points to four important features of the environment: persistent type heterogeneity of suppliers, explicit rewards for high performance, a dynamic tier policy that implements the performance rewards, and moral hazard in supplier behaviour.

3.1 Suppliers performance is heterogeneous

In Figure 1, I plot a histogram of shrinkage-adjusted supplier-level average pass rate. The large variation in supplier-level performance is indicative of there being substantial underlying heterogeneity in the capabilities of the suppliers. The model incorporates this via persistent, private types for the suppliers. The distribution of their types, in interaction with the buyer's policy and the suppliers' chosen behaviours, leads to the observed distribution of performance across orders.

Turning to an exploration of supplier performance by tier, in Figure 2a, I plot the performance of suppliers by tier on orders belonging to the same time period as the three snapshots. The heterogeneity in performance across tiers is indicative of selection, but there is further heterogeneity that this plot hides. In particular, the same Bronze tier bar can contain both a bad supplier and a good supplier who hasn't proved himself yet. This conditional plot is displayed in Figure 2b, where we see that suppliers who were observed as Bronze in a snapshot but were observed to be upgraded later performed better as Bronze than those that were dropped later. This also provides a natural segue into the next observation.

3.2 The buyer is responsive to performance

I show in Table 3 that suppliers that have lifetime better performance also have higher lifetime orders. This indicates that the buyer is selecting suppliers for longer relationships based on performance. Thus, a straightforward interpretation is that the buyer is engaging in selection of suppliers based on performance, in order to build a long-term supplier pool. An illustration of its implementation is in Figure 3, which shows that the buyer is more likely to give another order to suppliers who do better than average on the first order, with the buyer becoming less discriminating on the second order.[7]

A more direct question is whether the buyer prefers higher performing suppliers when making the decision of who to select for an order. Table 3 also shows that in any given month, suppliers that have a higher average past performance leading up to that month have a higher likelihood of getting orders in that month. This reinforces that the history of performance is an integral part of the buyer's supplier management policy, and gets closer to the variation that will be utilized in the estimation to get the magnitude of how much performance matters in

[7]The latter effect can be interpreted as a result of tier dynamics. Suppliers that do well in the first order earn themselves a higher tier, leading to a higher likelihood of remaining in the pool regardless of immediate subsequent performance.

the buyer's policy.

3.3 A dynamic tier system

The tier system is the mechanism used by the buyer to implement these performance-based rewards. Higher performance leads to higher tiers, and higher tiers lead to more orders. The fact that this system is explicitly dynamic is most easily seen in Figure 4, which shows that the mix of supplier tiers changes over the snapshots, and in Figure 5, which demonstrates that the mix changes because suppliers move across tiers.

In Figure 6, we see that the movements across snapshots depend in intuitive ways on performance, with suppliers that do better being more likely to get into better tiers, and also more likely to remain in the active supplier pool. This is one source of variation that helps pin down the buyer's tier-transition rule. Of course, this doesn't capture all the dynamism, since many orders and many intermediate tier changes can happen between snapshots.

The final piece—that being in a higher tier is associated with a higher likelihood of getting an order in a snapshot month—is illustrated in Figure 7. The magnitude of this preference, e.g., the extent to which a Gold supplier is likelier to get an order in and around a snapshot, is an important source of identification for the buyer's order-allocation rule.

3.4 There is suggestive evidence of moral hazard

Finally, I present some evidence that points towards the existence of moral hazard on the supplier side. The first observation is that supplier performance drops when there is a sudden steel price increase during order processing. The value of steel is on average 60% of the total value of the order, so suppliers have a strong incentive to cut steel usage when faced with an unanticipated steel price shock. Since the parts still need to adhere to strict engineering specifications, this implies that more effort cost needs to be incurred in terms of worker precision and management oversight to achieve the same pass rates on inspection; as a result, inspection pass rates within an order generally fall when there is sudden jump in steel prices. To demonstrate this within-order performance effect of steel prices, I run the following regression:

$$y_r = \alpha_{\text{order}(r)} + \beta_{\text{steel}} p_{\text{steel},m(r)} + \varepsilon_r$$

where y_r is the inspection outcome for part r, $\alpha_{\text{order}(r)}$ is an order-level fixed effect, and $p_{\text{steel},l(r)}$ is the price of steel in the month $m(r)$ that the steel used for part r was purchased.[8] The results are presented in Table 4, translating to a 3.5% drop in performance for a 10% increase in steel price within an order. This within-supplier response of performance to effort cost shocks is an important source of identification for the effort-related model parameters.

A demonstration of the steel price effect can also be seen in Figure 8a, which plots the performance impact of a 25% increase in steel price from Dec 2020 to Jan 2021. For all orders that had steel purchased in both Dec 2020 and Jan 2021, I compare the performance on lots with steel purchased in Dec 2020 and the performance on lots with steel purchased in Jan 2021. The figure demonstrates that the average performance dropped significantly, with all but one order seeing a performance drop across the two months.

Less direct evidence of moral hazard can be found in the performance response to the incentives induced by the buyer's policy. Given that the buyer rewards high performance by moving suppliers to higher tiers, and assuming that higher tiers face lower risk of relationship termination, a supplier who performs badly on the first order should be more likely to reach a lower tier and hence face higher incentives to perform well on the next order. This is one interpretation of the trend demonstrated in Figure 8b, which shows that suppliers with a below mean first order performance have a significantly higher performance in the second order than their first order.

Motivation for the structural model These empirical patterns jointly motivate the core features of the structural model. The presence of performance heterogeneity suggests a role for persistent supplier types. The dynamic nature of the tier system and its interaction with performance motivate the modeling of a dynamic business allocation policy with tier-contingent rules. Finally, the fact that suppliers respond to both cost shocks and to the buyer's policy necessitates a model where suppliers make an endogenous effort choice that responds to induced incentives.

4 Model

This section details the structural model that I take to the data. There are five broad categories of model objects that need to be specified—the order and supplier environments, the prefer-

[8] Because parts are manufactured, inspected, and delivered in lots, I cluster the standard errors at the lot-level to account for this shared processing.

ences of the two parties, the buyer's order-assignment process, the supplier equilibrium, and the buyer's payoff in equilibrium. The first four of these are directly relevant to the estimation. Since I do not impose optimality of the buyer's policy in the estimation, the buyer's payoff will be relevant only during the counterfactual exercises.

Orders and suppliers The buyer sources orders from downstream customers. Orders arrive in an exogenously specified sequence with subscript t indexing the sequence of orders.[9] The set of active suppliers is divided into mutually exclusive clusters within the different regions that the buyer operates in. The set S in a cluster represents the suppliers that are capable of doing similar orders in a region, based on their observed certifications, assigned order characteristics, and capacity information. The set of available active suppliers for the order at month t is denoted by $S_t \subseteq S$. The buyer can either pick an available active supplier from S_t or pick a new supplier from the outside pool. Suppliers have a fully persistent performance-relevant type denoted by $\theta \in [0,1]$. A new supplier's type is drawn IID from a distribution F. Orders are composed of many individual "parts", with groups of parts being worked on and delivered in "lots" at different points in time. Binary quality outcomes $y \in \{0,1\}$ are observed at the part level.

Preferences The supplier gets a normalized flow benefit $\bar{b} = 1$ from being assigned an order; the supplier side objects are identified relative to this benefit.[10] The buyer obtains a normalized flow benefit equal to the fraction of parts in an order that pass the first inspection.[11] The supplier's production processes on an order are grouped together at the lot level, including steel purchase, fabrication, inspection, and dispatch. This implies that the lot is the level at which the supplier chooses his endogenous action, and the level at which parts can face correlated quality shocks.[12] Thus, the outcome of the inspection for each part r in a lot l is drawn IID from a Bernoulli distribution $P(y_r = 1|\theta_{a_t}, e_l)$, where a_t is the supplier assigned to the order and $e_l \in \mathbb{R}_+$ is the effort chosen by the supplier for lot l, and the outcome probability $P(y = 1|\theta, e)$

[9] For the purpose of exposition, assume that t denotes a month. In the data, there can be multiple orders assigned in the same month, which requires a straightforward extension of the setup described here.

[10] This normalization is without loss since the payments made to a supplier for an order do not depend on performance on the order or on past performance, which is a fact I observe in the contracts between the buyer and suppliers. I also present evidence of this fact from the data in Appendix E.

[11] The timing of the first inspection can be treated as exogenous since the production timeline for an order, including the timing of the first inspection, is dictated by the customer's downstream timeline. Almost all of the buyer's customers operate on a Just-In-Time schedule, which necessitates the arrival of parts from the supplier exactly when needed according to the timeline of the downstream project.

[12] For example, a worker strike, power cut, or the strictness of an inspector, can affect the processing of all parts a lot.

for each lot is drawn from a distribution $G(\theta, e)$ to account for lot-level correlated shocks. The supplier's cost of choosing effort e_l is given by $c(e_l; X_l)$, where X_l are cost shocks that realize after order assignment and before effort choice, drawn IID from a known distribution. The two parties have a monthly discount factor δ.

Tier-based allocation policy Suppliers are assigned tiers based on performance history. There are four tiers that suppliers can be assigned to as a function of realized performance: Gold (G), Silver (S), Bronze(B), Dropped (D). Suppliers drawn from the outside pool are treated as a separate tier denoted New.

Let $\mathcal{T} = \{G, S, B, D\}$ denote the set of possible internal tiers. Let n_t denote the count of active suppliers at time t, i.e., of suppliers with tiers G, S and B. The dropped tier D is an absorbing state, indicating that the supplier will receive no further orders. Let $\mathbf{T_t} = \left(T_{1,t}, T_{2,t}, \ldots, T_{n_t,t}\right)$ denote the vector of the n_t active suppliers' tiers at the start of period t, where $T_{j,t}$ is supplier j's tier. Let $\mathbf{N_t}$ denote the vector of the number of suppliers in each tier i.e. $N_{t,T}$ gives the number of suppliers in tier T at time t. The buyer observes the full vector $\mathbf{T_t}$, while each supplier j only observes $(T_{j,t}, n_t)$ i.e. his own tier and the size of the active supplier pool.

Given a history $h_t = (a_{t'}, y_{r,t'}, X_{l,t'}, T_{t'})_{t'<t}$ that tracks the order assignments, order outcomes, observed cost shocks, and tier assignments leading up to a time t, the buyer uses a probabilistic allocation rule A to pick a supplier $A(h_t) \in \Delta(S_t \cup \text{New})$. In practice, the buyer tracks tiers as a summary of the performance history of suppliers, and uses the tiers to assign orders. This motivates the use of the tier vector as a Markovian state for the buyer's policy, i.e., the buyer chooses a supplier for an order based on tiers and then updates the tiers based on realized performance on the order. Formally:

Assumption 1. *The buyer uses the following tiered rule to determine A: for each history h_t, the buyer commits to $A_j(h_t) = p(T_{j,t}, \mathbf{N_t})$ for some function p, with $\sum_{k\in\mathcal{T}} N_{t,k}\, p(k, \mathbf{N_t}) \leq 1$ and $p(D, \mathbf{N_t}) = 0$ for all $\mathbf{N_t}$. With residual probability, the buyer chooses a new supplier.*

Moreover, the buyer commits to $\tau(T_{j,t}, n_t, y) \in \Delta\mathcal{T}$, i.e., the transition probabilities to a new tier for supplier j depends on his tier, the number of active suppliers, and the outcome that is produced if the supplier is selected.

The specific form of dependence on the tiers for p and τ are motivated by the context. The form for p captures the idea that suppliers within a tier are promised to be treated similarly, which implies that a given supplier's choice probability depends on his own tier and the number

of active suppliers in each tier. The form for τ captures the idea that the buyer promises to update a supplier's tier based only on commonly observed tier information.

Supplier equilibrium Fix a choice of p and τ by the buyer. Then, the suppliers' effort choices $e^*(T_{j,t}, n_t, X_t; \theta_j, p, \tau)$ form the equilibrium of a dynamic game among suppliers induced by the buyer's chosen policy. I make the following assumption about how suppliers form beliefs about the relevant objects in this induced incomplete information game:

Assumption 2. *Each supplier*

1. *treats the number of suppliers in each tier, given by the vector* $\mathbf{N_t}$*, conditional on the count of active suppliers,* n_t*, as given by a distribution* $H_1(n_t) \in \Delta\mathbb{N}^3$*, and*
2. *treats the transition of the count of active suppliers,* n_t*, as given by a process* $H_2(n_t) \in \Delta\mathbb{N}$*, i.e.* n_{t+1} *realizes from* $H_2(n_t)$*.*

Both of these beliefs $H = (H_1, H_2)$ *are rationally consistent with the buyer's policy* p*,* τ *and the long-run distribution of types and efforts of suppliers i.e.* H_1 *is given by*

$$H_1(\mathbf{N}|n) = \frac{\sum_{\mathbf{T}:|\mathbf{T}|=n} \pi(\mathbf{T}) \prod_{k\in\{G,S,B\}} \mathbb{1}_{\{\sum_{i=1}^{n} \mathbb{1}_{\{T_i=k\}}=N_k\}}}{\sum_{\mathbf{T}':|\mathbf{T}'|=n} \pi(\mathbf{T}')}$$

where π *is the long-run distribution of tier vectors conditional on* (p, τ, F, e^*)*, and* H_2 *is then formed by averaging the true state transition probabilities over the supplier's belief about the current state,* H_1*:*

$$H_2(n'|n) = \sum_{\mathbf{N}:\sum_T N_T=n} P(n'|\mathbf{N}) H_1(\mathbf{N}|n)$$

where $P(n'|\mathbf{N})$ *is the probability, induced by* (p, τ, F, e^*)*, that the number of active suppliers transitions to* n' *given the current tier vector is* $\mathbf{N}$*.*

The key restriction in this assumption is that the suppliers have the correct beliefs about the long-run dynamics of the buyer's tier system, and that they hold these beliefs fixed regardless of the specific time period that they interact with the buyer and whether they are chosen at that time period. The motivation is that it captures a rational supplier that has a simple, static belief formation process, which is in line with the small and informal nature of these businesses in practice. A caveat is that the data need not be representative of the long-run, but this condition is hard to relax tractably.

Given this assumption, the game reduces to the case in which each supplier solves a dynamic programming problem to compute his optimal effort choices. A supplier's continuation value for the dynamic programming problem $W(T_{j,t}, n_t, X_t; \theta_j, H)$ depends on his tier, the number of active suppliers, and the realized cost shock. Suppressing the dependence on type θ_j and the beliefs H, this value function solves the following recursive Bellman equation:

$$\begin{aligned} W(T_{j,t}, n_t, X_t) = \mathbb{E}\Bigg[p(T_{j,t}, \mathbf{N_t}) \max_e \Bigg\{ \bar{b} - c(e; X_t) \\ + \delta \int_y W(T_{j,t+1}, n_{t+1}, X_{t+1}) dP(y|\theta_j, e) \Bigg\} \Bigg] \\ + \mathbb{E}[\big(1 - p(T_{j,t}, \mathbf{N_t})\big)\, \delta W(T_{j,t+1}, n_{t+1}, X_{t+1})] \end{aligned} \tag{1}$$

The first term in Equation 1 captures the continuation value when selected, for which the supplier must choose optimal effort. The second term captures the continuation value when not selected. The expectations are taken over the number of active suppliers in each tier, the probabilistic outcome realization process conditional of effort and type, the probabilistic tier-transition process conditional on outcome, the transition of the number of active suppliers into the next period, and the stochastic cost shock process.

Buyer's payoff Given the prior F about supplier types, the buyer evaluates the present discounted value of a policy as an expectation over the path of realizations of supplier types, given the induced effort by the suppliers:

$$\begin{aligned} V(p, \tau) = \mathbb{E}_{\{(\theta_1, \theta_2, \ldots) | \theta_k \in F\}} \Big[P(y = 1 | \theta_1, e(\text{New}, n_1, X_1; \theta_1, p, \tau)) \\ + \delta\, p(T_2, \mathbf{N}_2) (P(y = 1 | \theta_1, e(T_2, n_2, X_2; \theta_1, p, \tau))) \\ + \delta\, (1 - p(T_2, \mathbf{N}_2)) (P(y = 1 | \theta_2, e(\text{New}, n_2, X_2; \theta_2, p, \tau))) + \ldots \Big] \end{aligned} \tag{2}$$

In Equation 2, the first line denotes the flow payoff in the first period, when the first supplier is picked from the outside pool. Then, the second and third lines denote the value contingent on the two possible choices in the second period - either pick the first supplier again, or pick the second supplier from the outside pool. The full expected PDV can be written as an expectation over the entire tree of contingent paths.

5 Estimation and identification

This section details the estimation and identification of the model. I begin by presenting the two-stage estimation procedure used to recover the model's parameters, which separates the estimation of the buyer's policy from the estimation of supplier-side components. Following this, I describe the specific functional forms used to parametrize the model objects. Then, I discuss the implementation of the estimation, including the computational algorithms used to handle latent state variables and solve the supplier's dynamic programming problem. Finally, I discuss how the model is identified.

5.1 Estimation procedure

The estimation of the model's structural objects is carried out in two sequential stages. This method is designed to overcome two major challenges inherent in the analysis of dynamic contracting models: specifying the buyer's choice problem, and the complexity that arises from the interdependence of the buyer's policy and supplier behavior. Specifying the buyer's choice problem in the estimation requires taking a stand on the feasible policy space and the constraints under which the buyer operates, including whether she can commit to future actions, which is especially challenging in developing country contexts. Moreover, incorporating the buyer-supplier equilibrium into the estimation necessitates explicit assumptions about information structure and strategic behavior—such as whether the revelation of private information creates dynamic incentive problems like the ratchet effect—and joint estimation is computationally demanding since it requires solving the full equilibrium of the dynamic game at each iteration of the estimation procedure.

Let $\{\mathbf{T_t}\}$ denote a realized path of tier vectors over time. Denote the observed tiers as $\{\mathbf{T}_t\}_{t^{\text{snap}}}$ at the tier snapshot times t^{snap}, and ζ as the suppliers side objects. The key idea is that the policy uses only the realized tiers and performance in the present period to determine supplier choice and tier transitions. This means that the joint likelihood of the data factors into

two:

$$\mathscr{L}\big((p,\tau),\zeta \,\big|\, \{a_t, y_t\}, \{\mathbf{T}_t\}^{\text{snap}}\big) = \int_{\boldsymbol{\theta}} \sum_{\{\mathbf{T}_t | t \notin \mathcal{T}^{\text{snap}}\}} \left[\prod_t \underbrace{\Pr(\mathbf{T}_t \mid \mathbf{T}_{t-1}, n_{t-1}, y_{t-1})}_{\text{transition}} \underbrace{\Pr(a_t \mid \mathbf{T}_t)}_{\text{choice}} \underbrace{\Pr\big(y_t \mid a_t, \mathbf{T}_t; p, \tau, \zeta, \boldsymbol{\theta}\big)}_{\text{supplier outcome}} \right] dF(\boldsymbol{\theta})$$

$$= \underbrace{\sum_{\{\mathbf{T}_t | t \notin \mathcal{T}^{\text{snap}}\}} \prod_t \Pr(a_t \mid \mathbf{T}_t)\ \Pr(\mathbf{T}_t \mid \mathbf{T}_{t-1}, n_{t-1}, y_{t-1})}_{\text{Stage 1}}$$

$$\times \underbrace{\sum_{\{\mathbf{T}_t | t \notin \mathcal{T}^{\text{snap}}\}} \underbrace{\frac{\prod_t \Pr(a_t \mid \mathbf{T}_t)\ \Pr(\mathbf{T}_t \mid \mathbf{T}_{t-1}, n_{t-1}, y_{t-1})}{\sum_{\{\tilde{\mathbf{T}}_t | t \notin \mathcal{T}^{\text{snap}}\}} \prod_t \Pr(a_t \mid \tilde{\mathbf{T}}_t)\ \Pr(\tilde{\mathbf{T}}_t \mid \tilde{\mathbf{T}}_{t-1}, n_{t-1}, y_{t-1})}}_{\text{policy-implied distribution over tier paths}} \int_{\boldsymbol{\theta}} \left[\prod_t \Pr\big(y_t \mid a_t, \mathbf{T}_t; p, \tau, \zeta, \boldsymbol{\theta}\big) \right] dF(\boldsymbol{\theta})}_{\text{Stage 2}} \tag{3}$$

where the second step multiplies and divides by the term labeled Stage 1. Thus, the buyer's chosen policy can be recovered first from the data using the order assignments, the realized outcomes, and snapshot tiers. Then, the second stage estimates the supplier-side objects, taking the estimated buyer policy as given. This greatly simplifies the estimation of the model, but the downside to this approach is that the estimators are less efficient than a joint MLE, since the first stage does not take into account the information about the policy that is contained in the supplier problems.

5.2 Parametrization

The estimation imposes a parametric structure for the model objects. Parametrization is necessitated in part by model complexity. The theoretical model in its full generality defines a policy over a high-dimensional state space, which presents challenges for estimation using a finite sample. At the same time, parametrization allows the model's core mechanisms to be represented by a small set of interpretable coefficients. The specific functional forms, chosen to be parsimonious, flexible, and faithful to the context, are detailed below.

Choice probabilities Active suppliers face an IID availability shock each period. The availability probability of an active supplier s for a particular order o is captured by a parametrized logistic link, $\frac{1}{1+\exp(-\zeta Z_{s,o})}$, where $Z_{s,o}$ represents observed characteristics. These characteristics include the log total capacity of the supplier and the log size of pending orders. This captures the idea that larger suppliers may be more likely to be available for an order and hence more likely to be assigned an order, and that suppliers with a higher existing workload may be less likely to have free capacity to take up new orders. Note that the scale of the availability shock

is normalized.[13]

The buyer then chooses from the available suppliers, or picks a new supplier, for an order. In practice, higher tier suppliers get higher consideration for an order, with approximate parity within a tier. This choice process can be represented as the buyer assigning each supplier s a choice weight that is indexed by his tier T_s, and then choosing from all available suppliers in proportion to the weight that they have been assigned. I use tier-specific weights ϕ_k to capture this process:

$$p(T_s, \mathbf{N_t}) = \frac{\exp \phi_{T_s}}{1 + \sum_{s' \in S_t} \exp \phi_{T'_s}} \tag{4}$$

with the weight for picking a new supplier being normalized to 0.

Supplier transitions Three simplifications govern tier transitions for suppliers. First, I assume that tier assignments are updated on a monthly basis, which is motivated by how tier reviews are implemented in practice. Second, I specify that existing suppliers can either be upgraded, remain in their current tier, be downgraded, or be dropped entirely.[14] This means that the tier movements have a ladder feature, where suppliers can move $B \to S \to G$ or $G \to S \to B$, if not dropped. This is designed to capture the essence of the tier system, where performance dictates moving up or down the ladder rungs of the system. Finally, suppliers that are not selected for an order in a given month retain their current tier going into the next month. This is a feature of the rule implemented by the buyer in practice, primarily to assure suppliers that their tiers are not being changed arbitrarily.[15]

Any supplier can drop out of the pool permanently in a given month with probability ψ, which is a parameter in the model, for exogenous reasons. This is meant to represent various relationship-ending shocks that the parties may face that are unrelated to performance, for example a supplier finding another regular customer or a supplier going out of business. These shocks are especially prevalent among informal suppliers in developing economies, characteristic of this setting, who face greater business instability.

I use a logit formulation to specify endogenous tier-transition probabilities for selected suppliers. Let the movement $h \in \{+1, 0, -1\}$ denote an upgrade, retention, or a downgrade, respectively, and let $\mathcal{H}(T, h)$ map a tier and a movement into the next tier. The function τ

[13] The scale of the availability shock is not identified separately from the parameters that govern choice probability.

[14] Given that Gold is the highest tier, being upgraded is the same as retaining their tier. Being downgraded for a Bronze supplier is treated as being dropped.

[15] This is also backed by the observation in the data that suppliers that do not receive an order between two snapshots also do not see a change in their tiers.

for suppliers in G, S, B tiers is parametrized by λ_h, the baseline parameter for a transition h; κ, a performance reward parameter that increases the likelihood of being upgraded; and ν, a parameter that shifts the probability of a supplier being dropped as a function of the number of active suppliers. Then, the probability of transitioning to tier $T' \neq D$ for a selected supplier in active tier T who had a pass rate of y across all lots in that month is given by:

$$\begin{aligned} \tau(T' \mid T, n, y) &= \frac{\sum_{h:\mathscr{H}(T,h)=T'} U(h; T, y)}{\sum_{h':\mathscr{H}(T,h')\neq D} U(h'; T, y) + U(D; T, n, y)} \quad \text{where} \\ U(h; T, y) &= \exp\left(\lambda_h + \kappa y \cdot \mathbb{1}_{\{h=+1\}}\right) \quad \text{and} \\ U(D; T, n, y) &= \exp(\nu n)\, I(T, y) \end{aligned} \tag{5}$$

where the adjustment term $I(T, y)$ for the Dropped utility is constructed such that the probability of being dropped by the buyer goes smoothly to zero as y goes to 1, implying that suppliers with a perfect outcome of $y = 1$ are never dropped for performance reasons, giving:[16]

$$I(T, y) = \left(1 - \frac{\sum_{h:\mathscr{H}(T,h)\neq D} U(h; T, y)}{\sum_{h':\mathscr{H}(T,h')\neq D} U(h'; T, y = 1)}\right)$$

Constructing the adjustment term $I(T, y)$ using the active tier utilities $U(h; T, y)$ ensures that the performance-based likelihood of remaining in the active pool is also controlled by the performance reward parameter κ, along with the likelihood of moving to a higher tier.

The specification above is applied for suppliers at the end of the month from their second month onward, when they already have an active tier assigned. For new suppliers at the end of their first month, I use parameters μ_G, μ_S, and μ_B for their initial tier assignments, giving:

$$\begin{aligned} \tau(T' \mid \text{New}, n, y) &= \frac{U(T'; \text{New}, n, y)}{\sum_{T'' \in \mathscr{T}} U(T''; \text{New}, n, y)} \quad \text{where} \\ U(T'; \text{New}, n, y) &= \begin{cases} \exp\left(\mu_{T'} + \kappa y \cdot \mathbb{1}_{\{T'=G\}}\right) & \text{if } T' \neq D \\ \exp(\nu n)\left(1 - \frac{\sum_{T''\neq D} U(T''; \text{New}, n, y)}{\sum_{T'''\neq D} U(T'''; \text{New}, n, y=1)}\right) & \text{o.w.} \end{cases} \end{aligned} \tag{6}$$

This formulation enables a minimal parametrization, but comes at the cost of restrictions on the tier transitions. In particular, the ladder feature implies that multiple tier jumps at one go (for example, $G \to B$) are disallowed;[17] the baseline transition tendencies being independent

[16] This normalization—that suppliers with a performance of $y = 1$ are never Dropped—enables the exogenous relationship-ending shock ψ to be identified separately from the endogenous transition parameters.

[17] In the data, I do not observe such multiple tier jumps for suppliers that have only one order between any two

of the starting tier impose uniformity across the tiers (for example, $G \to S$ and $S \to B$ have the same baseline tendency λ_{-1}); and the performance reward has the same shifter across the tiers.

Outcome and effort cost functions The outcome distribution G is modeled as a Beta distribution with mean $= \mu(\theta, e)$ and precision $= \varphi$. This means that supplier type and effort fix the mean of the distribution from which $P(y = 1)$ is drawn for a lot. The variability of this distribution, i.e., how far away $P(y = 1)$ could be from the mean, is controlled by the precision parameter φ, which captures the extent of lot-level shocks. With a large φ, most lots will have $P(y = 1)$ quite close to the mean μ, i.e., there will not be significant variability in average pass rates across lots with the same type and chosen effort. The choice of the Beta distribution to model the variability in outcomes due to lot-level shocks also enables a closed form for the lot-level likelihood, which resolves to draws from a BetaBinomial distribution for lot-level successes.

The mean is parametrized as $\mu(\theta, e) = \theta + e^{\alpha} K(\theta)$, where $K(\theta) = [1 + \exp(-\rho\theta)]^{-1}$. This form is chosen to incorporate the two primary economic factors affecting outcomes: how sensitive outcomes are to incentives, and how effort and type interact in the outcome production process. The parameter α captures sensitivity of outcomes to incentives, and ρ captures the complementarity between effort and type. The additive form for this function is intended to capture the idea that θ is a measure of the supplier's fundamental capability, with effort raising outcomes beyond the supplier's innate type. The buyer's prior over supplier types F is specified as a Beta distribution with parameters (η, ι), since types are assumed to be in $[0, 1]$.

The supplier effort cost function is parametrized as $c(e) = \gamma\left(\frac{\tilde{p}_t}{\tilde{p}_{t-1}}\right)e$, where $\tilde{p}_t$ is the steel price at t, realizing after order assignment and assumed to follow a geometric random walk: $\log \tilde{p}_t = \log \tilde{p}_{t-1} + \xi_t$ where $\xi_t \sim N(0, \sigma_\xi)$. Since effort is an unobserved index, it is without loss of generality for cost to be linear in effort, with the shape of the performance response entering via the α parameter in the outcome function. The steel price ratio captures an unanticipated IID cost shock, rationalized by the following economic mechanism. Suppliers are paid a fixed amount based on expected steel price at order allocation, $\tilde{p}_{t-1}$ (say, $(1 + \epsilon) l_{\text{steel}} \tilde{p}_{t-1}$, where l_{steel} is the steel needed based on engineering drawings and ϵ is a margin for wastage). When the realized price $\tilde{p}_t$ is higher, suppliers reduce the amount of steel purchased, shrinking the available margin (say, to $\epsilon' < \epsilon$). This means that higher organizational effort cost, in the form of worker precision and manager supervision, must be incurred to achieve the same pass rate.

snapshots.

5.3 Implementation

Stage 1: Policy parameters estimation In the first stage, I estimate the buyer's chosen policy, conditional on the observed supplier choice decisions, outcomes, and snapshot tier assignments. The Stage 1 likelihood is specified as:

$$\begin{aligned}\mathscr{L}\Big(p,\tau\,\Big|\,\{a_t,y_t\},\{\mathbf{T}_t\}_{t^{\text{snap}}}\Big) &= \sum_{\{\mathbf{T}_t|t\notin\mathscr{T}^{\text{snap}}\}}\prod_t \Pr(a_t\mid\mathbf{T}_t)\ \Pr(\mathbf{T}_t\mid\mathbf{T}_{t-1},n_{t-1},y_{t-1})\\ &= \sum_{\{\mathbf{T_t}|t\notin\mathscr{T}^{\text{snap}}\}}\prod_t p\Big(T_{a_t,t},\mathbf{N_t}\Big)\prod_j \tau\Big(T_{j,t};\,T_{j,t-1},n_{t-1},y_{t-1}\Big)\end{aligned} \tag{7}$$

I use a Monte Carlo Expectation-Maximization (MCEM) algorithm (see Dempster, Laird, and Rubin, 1977; Wei and Tanner, 1990) to implement this estimator. The fundamental premise of the EM algorithm is to replace the single, difficult problem of maximizing the likelihood in Equation 7 with a sequence of two simpler, iterative steps. In the Expectation (E) step, given the current best guess of the policy parameters, I "fill in" the latent information by calculating the expected log-likelihood of the complete data. Since this expectation is analytically intractable, I approximate it by simulating a large number of latent tier paths from this posterior and replacing the expectation with their sample average—this is the Monte Carlo component of the algorithm. In the Maximization (M) step, I treat this simulated average log-likelihood as the new objective function and find the parameter values that maximize it. This gives an updated, better guess for the parameters.

Formally, within each iteration k, the E step and the M step are applied:

E step. Given the policy (p_{k-1},τ_{k-1}) from the previous iteration, I first simulate $Q=2000$ complete latent tier paths recursively, as follows. Using the transition rule τ_{k-1}, the predictive distribution of the latent tier vectors at each time t between two snapshot times t_1 and t_2, i.e. for $t_1<t<t_2$, is recursively derived as:

$$\begin{aligned}&P\Big(\mathbf{T}_t\,\Big|\,\{a_{t'},y_{t'}\}_{t'\le t-1};\,\tau,\mathbf{T}_{t_1}\Big)\\ &\quad=\sum_{\mathbf{T}_{t-1}} P\Big(\mathbf{T}_{t-1}\,\Big|\,\{a_{t'},y_{t'},\}_{t'\le t-2};\,\tau,\mathbf{T}_{t_1}\Big)P\Big(T_t\,\Big|\,T_{t-1},a_{t-1},y_{t-1};\,\tau\Big)\\ &\quad=\sum_{\mathbf{T}_{t-1}} P\Big(\mathbf{T}_{t-1}\,\Big|\,\{a_{t'},y_{t'}\}_{t'\le t-2};\,\tau,\mathbf{T}_{t_1}\Big)\prod_j \tau\Big(T_{j,t};\,T_{j,t-1},n_{t-1},y_{t-1}\Big)\end{aligned} \tag{8}$$

For the case where $t = t_1$, the probability is defined as:

$$P\Big(\mathbf{T}_t \,\Big|\, \{a_{t'}, y_{t'}\}_{t' \le t-1};\, \tau, \mathbf{T}_{t_1}\Big) = \begin{cases} 1 & \text{if } \mathbf{T_t} = \mathbf{T}_{t_1} \\ 0 & \text{o.w.} \end{cases}$$

This framework extends to the period preceding the initial snapshot, given that my data starts at the buyer's first order, and that suppliers begin in the "New" tier before their first order.

After drawing the simulated paths, for each path q, I assign an importance-weight proportional to the product of the observed assignment likelihoods in that interval given the pre-decision tiers, and normalize across paths:

$$\tilde{w}_q^{(k)} = \prod_t \Big[\prod_{o \in \mathcal{O}_t} p_{k-1}(a_t^{(o)} \mid \mathbf{T}_t^q)\Big], \qquad w_q^{(k)} = \frac{\tilde{w}_q^{(k)}}{\sum_{q'=1}^{Q} \tilde{w}_{q'}^{(k)}}.$$

Using these weights, I form a Monte Carlo approximation to the expected complete-data log-likelihood:

$$\mathcal{Q}_k(p, \tau) \approx \sum_{q=1}^{Q} w_q^{(k)} \Bigg\{ \sum_t \Big[\sum_{o \in \mathcal{O}_t} \log p(a_t^{(o)} \mid \mathbf{T}_t^q)\Big] + \sum_t \sum_j \log \tau\Big(T_{j,t}^q \mid T_{j,t-1}^q, n_{t-1}^q, y_{t-1}\Big) \Bigg\}.$$

M step. Holding $\{w_q^{(k)}\}_{r=1}^R$ fixed, I update the policy by implementing:

$$(p_k, \tau_k) = \arg\max_{p,\tau} \mathcal{Q}_k(p, \tau).$$

This procedure is assumed to have converged when the update to each policy parameter is smaller than a pre-specified tolerance.

Stage 2: Supplier parameters estimation Given the estimated policy p^*, τ^* from Stage 1, the second stage estimates the supplier side objects, namely the supplier cost function c, the distribution of supplier types F, and the outcome function G. Let $P_l(y_r|\theta, e^*)$ denote the probability of observing outcome y_r for part r, given supplier type θ and the supplier's optimal effort e^* (which is given by the solution to the supplier problems specified in Equation 1, which in turn depend on the estimated policy and the tier path). Then the Stage 2 likelihood is specified

as:

$$
\begin{aligned}
&\mathcal{L}(c, F, G \mid \{a_t, y_t\}, \{\mathbf{T}_t\}_{t^{\text{snap}}}; p^*, \tau^*) \\
&\quad = \sum_{\{\mathbf{T}_t | t \notin \mathcal{T}^{\text{snap}}\}} P(\mathbf{T} \mid \{a_t, y_t\}, \{\mathbf{T}_t\}_{t^{\text{snap}}}; p^*, \tau^*) \int_{\boldsymbol{\theta}} \left[\prod_t \Pr\big(y_t \mid a_t, \mathbf{T}_t; p, \tau, \zeta, \boldsymbol{\theta}\big) \right] dF(\boldsymbol{\theta}) \\
&\quad = \sum_{\{\mathbf{T}_t | t \notin \mathcal{T}^{\text{snap}}\}} P(\mathbf{T} \mid \{a_t, y_t\}, \{\mathbf{T}_t\}_{t^{\text{snap}}}; p^*, \tau^*) \prod_j \left\{ \int_{\theta_j} \left[\prod_l \int_{P_l} \prod_{r \in l} P_l\big(y_r \mid \theta, e^*\big)\, dG \right] dF(\theta_j) \right\}
\end{aligned}
\tag{9}
$$

where $P(\mathbf{T} \mid \{a_t, y_t\}, \{\mathbf{T}_t\}_{t^{\text{snap}}}; p^*, \tau^*)$ is the policy-implied posterior over tier paths. I implement a Simulated Maximum Likelihood (SML) procedure for the estimation, as the likelihood function in Equation 9 involves an integral over both the latent tier paths from Stage 1 and the supplier types.[18] I first describe the construction of the supplier beliefs that are needed in order to solve the supplier problems, and then detail the construction of the simulated likelihood.

Derivation of supplier beliefs. Deriving the supplier optimal effort e^* within Equation 9 requires solving the supplier problems W as specified in Equation 1. This, in turn, depends on the beliefs H held by the suppliers. Given Assumption 2, the object H_2, which represents the belief regarding changes in the number of active suppliers, is constructed from the data:

$$
H_2(n'; n) = \frac{\sum_t \mathbb{1}_{\{|S_t|=n \ \& \ |S_{t+1}|=n'\}}}{\sum_t \mathbb{1}_{\{|S_t|=n\}}}
$$

where I infer the total number of active suppliers in each month in each cluster as the number of suppliers who have been observed to get at least one order preceding that month, and at least one order during or after that month.

To derive the belief H_1, which captures the distribution of the number of suppliers within each tier conditional on the total number of active suppliers, I use the distribution over tiers implied by Equation 8. Specifically, I draw $Q = 2000$ paths of tier assignments for the suppliers and then calculate the fraction of times a particular configuration of tier assignments was

[18] It is well-known that while the simulated likelihood provides an unbiased estimate of the true likelihood, maximizing the simulated log-likelihood can introduce a bias into the parameter estimates for finite simulations, with the bias disappearing in the limit as the number of draws increase (Train, 2009). To mitigate simulation bias, I use a large number of simulation draws: $Q = 2000$ for the latent tier paths and $M = 200$ for the latent supplier types for each path.

realized relative to the times the corresponding number of active suppliers was realized:

$$H_1(N;n) = \frac{\sum_{q,t} \mathbb{1}_{N_{q,t}=N}}{\sum_{q,t} \mathbb{1}_{n_{q,t}=n}}$$

Likelihood construction. With a fixed set of parameters, the simulated likelihood can then be formed as follows. First, I simulate $Q = 2000$ complete latent tier paths $\{\mathbf{T}^{(q)}\}_{q=1}^{Q}$ from the policy–implied posterior over tier trajectories $P(\mathbf{T} \mid \{a_t, y_t\}, \{\mathbf{T}_t\}_{t^{\text{snap}}}; p^*, \tau^*)$, and then calculate the normalized importance-weights $\{w_q\}_{q=1}^{Q}$ for the paths, using the same procedure as the E step in Stage-1 .

Conditional on a simulated tier path q, for each supplier j, I draw types $\theta_j^{(m)} \overset{\text{i.i.d.}}{\sim} F$ for $m = 1, \ldots, M = 200$. For each draw, the optimal effort e_{jl}^* at each lot l is part of the solution to the supplier's dynamic problem in Equation 1. Given that z_{jl}^1 parts pass inspection out of z_{jl} total parts in lot l, integrating over the lot-level pass probability using the Beta distribution with precision φ and mean $\mu(\theta, e)$ (the lot-level pass probability is drawn from a Beta distribution, as specified in Subsection 5.2) yields a Beta–Binomial marginal for $(z_{jl}^1 \mid z_{jl})$. Thus, the simulated supplier-level likelihood under path q is:

$$\hat{\mathscr{L}}_j^{(q)} = \frac{1}{M} \sum_{m=1}^{M} \prod_{l \in \mathscr{L}_j} \text{BB}\Big(z_{jl}^1;\, z_{jl},\, \alpha_{jlm}^{(q)},\, \beta_{jlm}^{(q)}\Big), \qquad \begin{aligned} \alpha_{jlm}^{(q)} &= \varphi\, \mu_{jlm}^{(q)}, \\ \beta_{jlm}^{(q)} &= \varphi\big(1 - \mu_{jlm}^{(q)}\big), \\ \mu_{jlm}^{(q)} &= \mu\Big(\theta_j^{(m)},\, e_{jl}^*\big(T_{j,t(l)}^{(q)},\, n_{t(l)}^{(q)},\, X_{jl}\big)\Big) \end{aligned}$$

where $\text{BB}(z^1; z, \alpha, \beta)$ denotes the Beta–Binomial PMF, $\text{BB}(z^1; z, \alpha, \beta) = \binom{z}{z^1} \dfrac{B(z^1 + \alpha,\, z - z^1 + \beta)}{B(\alpha, \beta)}$.

Finally, I aggregate across suppliers within each path and then average across paths using the policy-implied weights:

$$\hat{\mathscr{L}} = \sum_{q=1}^{Q} w_q \prod_{j=1}^{J} \hat{\mathscr{L}}_j^{(q)}, \qquad \hat{\ell} \equiv \log \hat{\mathscr{L}} = \log\left(\sum_{q=1}^{Q} w_q \exp\left\{ \sum_{j=1}^{J} \log \hat{\mathscr{L}}_j^{(q)} \right\} \right)$$

Solving the suppliers' dynamic programming problems in Equation 1 for each guess of the structural parameters imposes significant computational cost. To mitigate this cost, I use a neural network to approximate the supplier's optimal effort function during the estimation. The details of this procedure are in Appendix B.

5.4 Identification

The formal identification results for the model are presented in Appendix C. Here, I discuss the statistical arguments and economic intuition on how the model is identified.

Identification of the buyer's policy (p, τ) The main hurdle in identifying the buyer's policy is that supplier tiers are latent for most of the sample; if tiers were always observed, the policy could be read off the data. In the observed-tier benchmark, two adjacent periods suffice: period-t choices, indexed by the current tier $\mathbf{T}_t$, recover the allocation rule $p(\mathbf{T}_t)$, and the adjacency $(\mathbf{T}_t, \mathbf{T}_{t+1})$ recovers the transition rule $\tau(\mathbf{T}_{t+1} \mid \mathbf{T}_t, \cdot)$.

The standard approach for identification of Markov processes with latent states in the econometrics literature is to use observables in nearby periods, leveraging the first-order Markov property. For example, the argument in Hu and Shum (2012) would translate as follows into this paper's setting. Given the latent state $\mathbf{T}_t$ and the observed vector $\mathbf{A}_t$, construct two functions of nearby data, $A^- \equiv g_-(\mathbf{A}_{t-2})$ and $A^+ \equiv g_+(\mathbf{A}_{t+1})$. Due to the first-order Markov property, A^+ is conditionally independent of A^- given $\mathbf{T}_t$. With sufficient variation in A^+ and A^-, the joint distribution of observables over a short run of adjacent periods (four under stationarity, five without it) can recover the choice and transition rules.

However, these short-panel results hinge on short-run observables being sufficiently rich relative to the latent state space (e.g., many choices or covariates with wide support). In my data, the number of possible tier configurations can be much larger than the number of suppliers, so I instead apply a related approach from the Hidden Markov Model (HMM) literature (see Allman, Matias, and Rhodes (2009)) that provides a way to leverage the first-order Markov property by grouping longer sequences of observations. Fix a snapshot month t in which the tier vector $\mathbf{T}_t$ is observed, and let A^0 denote the realized within–snapshot order allocations at t. Form a *future block* A^+ that collects the observed order allocations in the k periods after t, and a symmetric *past block* A^- for the k periods before t. The policy's first–order Markov property implies the following exclusion restriction: the past and future blocks are independent conditional on the snapshot observables, i.e.,

$$A^- \perp\!\!\!\perp A^+ \,\big|\, \left(\mathbf{T}_t, A^0, y^0\right)$$

The distribution of A^+ can be represented as a finite mixture over the possible next-period tiers, with the mixture weights being determined by $\Pr(\mathbf{T}_{t+1} \mid \mathbf{T}_t, \cdot)$ and choice rule $p(\mathbf{T}_{t+1})$, and symmetrically for A^-. Increasing the size of these blocks provides more observable patterns

in A^- and A^+, restricting the choice and transition rules more tightly. Thus, with sufficiently large blocks of observations, the past and future blocks provide sufficiently rich conditionally-independent sources of variation that are needed to identify the policy.

Economic intuition. For economic intuition, consider the two primary set of policy parameters of interest: the tier-specific selection weights, the ϕ_k parameters, which govern how the buyer chooses among suppliers in different tiers, and the performance reward, κ, which governs how much performance influences a supplier's tier transition.

A natural starting point is to look at suppliers of lower tiers (say, Silver) who produce a high pass rate in a snapshot month. If we observe that such suppliers are selected often in the months after the snapshot, it suggests that good performance is being rewarded. However, the high selection frequency is a compound effect: a high performance leads to a probable tier upgrade, which in turn leads to a higher chance of being selected. A given pattern of high-performing lower-tier suppliers being chosen frequently could be explained by a large performance reward κ and small differences between the tier choice weights ϕ_k, or a small κ and large differences between the ϕ_k values. So the two sets of parameters are confounded.

This is why we leverage both past and future selection probabilities. The policy's Markov property implies that, conditional on the supplier tiers at the snapshot, the selection history before that point is statistically independent of future selections. Thus, the relative difference in future versus past selection probabilities, around a high pass rate snapshot, provides a statistically valid indicator of the performance reward κ. This allows κ and ϕ_k to be recovered jointly from the selection probabilities.

Identification of the supplier side (c, G, F) The challenge on the supplier side is to disentangle three unobserved components that simultaneously affect performance: a supplier's persistent type, the chosen effort, and random lot-level shocks. The economic intuition of the identification approach is that the simultaneity can be resolved by leveraging sequentially nested variation in the data, starting at the most granular level. Variation in performance within an order recovers the distribution of lot-level random shocks; variation within a supplier then allows the identification of parameters related to effort choices at the supplier level; finally, variation across suppliers allows for the recovery of the type-related parameters and the distribution of types.

Notation and equations. For supplier i, in policy state $s \equiv (T, n)$,[19] and lot ℓ, recall that the

[19]The identification of the cost-side parameters requires the policy state, i.e. the tiers, to be observed; this leverages the snapshot periods. The rest of the arguments go through using the distribution over tiers that is

probability that a part passes inspection is given by $y_{isℓ} = \text{Beta}(\text{mean} = \theta_i + e_{isl}^{\alpha} K(\theta_i), \text{precision} = \varphi)$, where effort e_{isl} solves the supplier FOC, and K captures effort-type complementarity. Recall that the cost shock is denoted $X_t = p_t / p_{t-1}$. We can combine the outcome map and the FOC,

$$\mathbb{E}[y_{isl}|i,s] = \theta_i + e^{\alpha} K(\theta_i), \qquad \underbrace{c'(e;X)}_{=\gamma X} = \underbrace{\frac{\partial \mu}{\partial e}}_{=\alpha e^{\alpha-1} K(\theta_i)} \cdot \underbrace{\overline{\mathscr{W}}(s,\theta_i)}_{\text{marginal continuation value}},$$

into a single equation that underpins the identification arguments:

$$\mathbb{E}[y_{isl}|i,s] - \theta_i = \tilde{A}(s,\theta_i) X_l^{\frac{\alpha}{\alpha-1}}, \qquad \tilde{A}(s,\theta_i) \equiv \alpha^{-\frac{\alpha}{\alpha-1}} \gamma^{\frac{\alpha}{\alpha-1}} K(\theta_i)^{\frac{1}{1-\alpha}} \overline{\mathscr{W}}(s,\theta_i)^{-\frac{\alpha}{\alpha-1}}. \tag{10}$$

Within-order variation. Consider orders with at least two lots that face the same cost shock X_l (e.g., orders with two lots in the same month). Type θ_i and incentives are fixed, so optimal effort e_{isl} is common across those lots. Any residual dispersion in performance must be due to the lot-level shock; the within-order dispersion in performance therefore identifies the precision (variance) of the lot-level shock.

Within-supplier variation. The next step uses exogenous shifts to marginal cost, while keeping marginal benefit fixed, to recover the elasticity of performance with respect to incentives. Using Equation 10,

$$\mathbb{E}[y_{isl}|i,s] - \theta_i = \tilde{A}(s,\theta_i) \cdot (X_l)^{\frac{\alpha}{\alpha-1}}$$

Given three distinct cost shocks in the same tier for the same supplier (e.g., three lots in the same order, or three lots across two different snapshots with the same tier, with different steel prices), we can take differences of this equation and then take ratios to eliminate A and θ_i, giving an equation that can be solved for α,

$$\frac{\mathbb{E}[y_{i,t}] - \mathbb{E}[y_{i,t-1}]}{\mathbb{E}[y_{i,t-1}] - \mathbb{E}[y_{i,t-2}]} = \frac{(X_t)^{\frac{\alpha}{\alpha-1}} - (X_{t-1})^{\frac{\alpha}{\alpha-1}}}{(X_{t-1})^{\frac{\alpha}{\alpha-1}} - (X_{t-2})^{\frac{\alpha}{\alpha-1}}}$$

Moreover, given α, we can then recover θ_i for all suppliers that have at least two lots in the same tier with different shocks,

$$\theta_i = \frac{\mathbb{E}[y_{i,t}] \cdot (X_{t-1})^{\frac{\alpha}{\alpha-1}} - \mathbb{E}[y_{i,t-1}] \cdot (X_t)^{\frac{\alpha}{\alpha-1}}}{(X_{t-1})^{\frac{\alpha}{\alpha-1}} - (X_t)^{\frac{\alpha}{\alpha-1}}}$$

Thus, this step identifies the elasticity of performance with respect to incentives and the types for a subset of suppliers (the "anchor" types).

recovered in the first stage.

Across-supplier variation. Next, the variation in performance across the recovered anchor types enables recovery of effort-type complementarity and the scale of cost of effort. Define a type- and cost-adjusted performance measure $Y^*_{is} = \ln(\mathbb{E}[y_{isl}|i,s] - \theta_i) - \frac{\alpha}{1-\alpha}\ln(X_s)$. Rewriting Equation 10 in logs using this performance measure,

$$Y^*_{is} = B \ln \widetilde{\mathscr{W}}(s, \theta_i) + C \ln K(\theta) + D \tag{11}$$

where B, C, D are known constants. Given the policy (p, τ) recovered in the first stage, the marginal continuation value $\widetilde{\mathscr{W}}$ decomposes into a type-specific part and a part computable from the data:

$$\begin{aligned}\widetilde{\mathscr{W}}(s, \theta_i) &= \underbrace{\delta \left(\nabla_y \tau(s,y)^\top\right) V_{\theta_i}}_{\text{gradient of tier transitions } \times \text{ tier-specific continuation value}} \\ &= \delta \, \nabla_y \tau(s,y)^\top (I - \delta P)^{-1} q \times u_{\theta_i}\end{aligned}$$

where the second step decomposes the the tier-specific continuation value vector V_{θ_i} into the type-specific net flow value of being assigned an order u_{θ_i} (which is a known function of K and γ) and the discounted stream of future selection opportunities from each tier $(I - \delta P)^{-1} q$. The term $\nabla_y \tau(s,y)^\top (I - \delta P)^{-1} q$ is computed using objects recovered in the first stage: $\nabla_y \tau(s,y)$ is the gradient of the tier transition probabilities wrt y at state s, P is the tier transition matrix, and q is the per-tier assignment vector. Substituting this into Equation 11 and comparing across different anchor types identifies the effort-type complementarity and cost scale parameters.

Deconvolution. With the parameters of supplier behavior identified, the observed distribution of performance is a mixture over performance by different types. On first orders, all suppliers are in the same observed New tier and their types are drawn from the underlying distribution F, so

$$f_{Y^{(1)}}(y) = \int f_{Y^{(1)}|\theta}(y \mid \theta) \, dF(\theta).$$

Given that performance is strictly increasing in type for sufficiently large cost shocks,[20] the distribution of types can be recovered from the distribution of performance produced on first orders by using the following deconvolution argument. The latent mean performance on the first order is a known increasing function of type, $\mu(\theta)$, fixing a sufficiently large cost shock. Conditional on $\mu(\theta)$, the observed number of successful parts in a lot, Z^+, follows a Beta-Binomial distribution. This implies that the moments of Z^+ are proportional to the moments of the pass probability, so the observed moments of Z^+ nonparametrically identify the full

[20] When cost shocks are very low, all types will choose very large effort. With a sufficiently negative complementarity between type and effort, this can theoretically lead to lower types performing the same or better than high types.

distribution of the latent mean performance measure $\mu(\theta)$. The final step is to invert this known mapping μ to recover the underlying distribution of types, $F(\theta)$.

6 Results

Estimates for the buyer's policy parameters The policy parameter estimates from the first stage of the estimation procedure are presented in Table 5. I discuss the process of calculating standard errors for the estimates in Appendix D. We can gain an intuitive understanding of the estimates by translating them into choice and transition probabilities, presented in Table 6 and Table 7.

The estimates for the choice rule p imply that in a pool with the snapshot average of 2 Gold suppliers, 2 Silver suppliers and 2 Bronze suppliers, the Gold supplier has a 23% chance of being selected when available, a Silver supplier has a 20% chance of being selected, and a Bronze supplier has 15% chance of being selected. As the supplier mix changes over time, the selection probabilities also change, as illustrated in Figure 9. To illustrate how availability enters the model, a doubling of the workload for an average supplier reduces the probability that the supplier is available for an order from 54% to 45%, while doubling his capacity instead increases it to 62%.

The estimates for the transition rule τ show that tier transitions are heterogeneous by tier. Gold suppliers are likely to retain their tier - at the mean dataset performance of $y = 0.75$, they retain their tier with 64% probability. Silver and Bronze suppliers have a more equalized distribution across their different transition paths, with the latter having a higher likelihood of exiting the active pool. New suppliers are significantly more likely to be dropped. The estimates also imply that tier movements are strongly tied to performance, as Figure 10 shows. For example, as performance shifts from $y = 0$ to $y = 1$ for a Silver supplier in a five supplier pool, the supplier goes from close to 0% likelihood of being upgraded to about 80% likelihood of being upgraded. Alternatively, shifting performance from the sample average of $y = 0.75$ to $y = 1$, the likelihood of being upgraded goes from 38% to 80%.

Estimates for the supplier-side parameters The second stage estimates can be viewed in Table 8. I discuss the process of obtaining standard errors for the estimates, which accounts for the uncertainty in Stage 1 estimates, in Appendix D. To aid the interpretation of the estimates, I present some key implied statistics in Table 9, by simulating the full model for 1000

simulations.[21]

The estimated external type distribution, plotted in Figure 11, implies that the mean external type is 0.70. Recall that lot-level mean outcomes are given by $\theta + e^{\alpha}[1+\exp(-\rho\theta)]^{-1}$, where I refer to the term $e^{\alpha}[1+\exp(-\rho\theta)]^{-1}$ as the 'effort contribution'. The average effort contribution in the simulations is 0.03, and its standard deviation is 0.03. Since the average pass rate per order in the data is 0.79, this gives us an approximate sense that incentivization is responsible for a smaller share of the improvement in performance, relative to selection; a more systematic decomposition of value into the two mechanisms is deferred to the next section. The split of average type by tier over time, plotted in Figure 12, makes it clear that the buyer's chosen policy leads to an expected sorting of types to tiers, with higher types reaching higher tiers.

Interestingly, the same plot also makes it clear that the dispersion of outcomes across tiers is less than that of types across tiers. This primarily comes down to the disparity in effort, presented in Figure 13. This is partly due to the fact that lower tier suppliers have a higher incentive to perform well in order to be upgraded to safer, more stable tiers. But another major reason that lower tier suppliers choose higher effort is the fact that the estimates also imply that effort and type are substitutes in the outcome production function, since the estimate ρ is negative (see Table 8). This means that, given the same level of incentives, lower types choose higher effort. Of course, different types have different induced incentives via the policy even at the same tier, by virtue of facing different transition distributions and different future pass rates. Nonetheless, in equilibrium, the lower types of the lower tier suppliers contributes to higher effort choice.

The final takeaway from the second stage estimates is that supplier effort generally follows an inverted U-shape as a function of the number of competitors, as demonstrated in Figure 14.[22] There are two, opposing driving forces behind this result. The first force is the increase in the likelihood of being dropped due to performance as the supplier pool size increases. When a supplier is the monopolist in his cluster, the supplier doesn't have to worry too much about being replaced. As the pool size increases, the supplier finds it valuable to increase her effort to try to stay in the active pool. But there is a second force at play. Being in a competitive pool is associated with the pool staying competitive in the future as well, which

[21] The details of the simulation for this exercise are as follows. I treat the order arrivals and characteristics as given. For each simulation, a sequence of suppliers is drawn from the prior distribution. This is the sequence in which they become available when the buyer decides to choose a new supplier.

[22] The figure plots the optimal effort $e^*(T, n, \tilde{X}; \theta, p^*, \tau^*)$ as a function of n by fixing $T = \text{Silver}$, $\tilde{X} = 1$, $\theta = 0.7$. Note that the optimal effort solves the dynamic programming problem in Equation 1, taking the estimated supplier beliefs as given.

implies a lower likelihood of being chosen in the future. So the supplier's overall value from the relationship also reduces as the pool size increases. Past a certain point, the latter effect dominates, and hence the supplier's effort reduces as well. This observation implies that there is an optimal level of competition in the pool, with respect to inducing supplier effort.

Model fit In Figure 15, I compare the predicted distribution of tiers in the three snapshots (labeled p1, p2, p3) with the observed distribution, aggregating across all clusters. The predicted distribution is obtained by simulating tiers many times, under the estimated policy parameters. The distributions are broadly similar, indicating a fit that is reasonable, especially given the restricted formulation and parametrization that I have used. In Figure 16, I compare the observed CDF of supplier-level pass rates to the one predicted by simulating the model 1000 times.[23] The predicted distribution closely tracks the observed distribution.

7 Policy analysis and counterfactuals

The estimated structural model enables a rigorous evaluation of the buyer's policy design. In this section, I first quantify the policy's impact. I measure the total performance gain relative to a baseline of random order allocation and then decompose this value into its two constituent channels: the firm's ability to select better suppliers versus its capacity to incentivize their effort.

With the policy's performance improvement quantified, I then characterize the optimal policy to establish a performance frontier for the estimated supplier environment. This comparison reveals a significant performance gap between the estimated and optimal policies, raising a critical question: what rationalizes the firm's deployment of a seemingly suboptimal strategy? I show that the firm's chosen policy is more robust to economic uncertainty, outperforming the optimal policy when faced with a less capable pool of suppliers.

Finally, I contextualize the policy's economic effect by benchmarking it against two alternative arrangements. To translate the performance improvement into monetary terms, I calculate the monetary bonus required for the same performance improvement in a spot-market setting. To determine how much the policy's effectiveness relies on the scale of operations, I quantify the performance loss in a counterfactual that replaces the buyer's intermediation with direct customer-supplier contracting.

[23] This simulation treats both the order details and the assigned supplier details as given.

7.1 Value added by the estimated policy

I quantify the value added by implementing the selected policy using the metric of expected discounted measure of performance, as defined in Equation 2. Since the buyer's policy leverages the two interconnected mechanisms of incentivization and selection to create value, I also decompose this metric into these components. The methodology for this decomposition is as follows:

1. The baseline: I calculate the baseline level of outcomes using a policy that picks a new supplier for each order and then drops them after the order. This gives V_{baseline} by plugging in this policy into Equation 2. Note that the suppliers' optimal response to a such a policy is to put in no effort.

2. The value of selection: Then, I calculate the value associated with selection by evaluating Equation 2 for the estimated policy with assumption that suppliers choose $e = 0$ always. This gives $V_{\text{selection}}$.

3. The combined value: Finally, I evaluate Equation 2 for the estimated policy under the full model with both selection and incentivization, giving V_{combined}.

The decomposition into the selection and incentivization effects can then be derived as

$$\left(\frac{V_{\text{selection}} - V_{\text{baseline}}}{V_{\text{combined}} - V_{\text{baseline}}}, \frac{V_{\text{combined}} - V_{\text{selection}}}{V_{\text{combined}} - V_{\text{baseline}}}\right) \tag{12}$$

A visual representation of the value and its decomposition can be found in Figure 17. The estimated policy increases value by around 12% above the baseline, with around 66% of the improvement being attributable to selection. This means that the majority of the additional value comes via transitioning better types to higher tiers, and then selecting the higher types more often due to a preferencing of the higher tiers.

When the present discounted value is calculated starting one year into the sample, this split becomes more lopsided, with around 75% of the value added being attributable to selection. This is expected, given the fact that this is enough time for the supplier types to sort more cleanly across tiers. Moreover, because the estimates imply that type and effort are substitutes in the outcome function, as the average type in the active pool increases to a steady state value, the average effort decreases.

7.2 The optimal policy

What policy would the buyer have chosen to implement had she wanted to maximize the pass rate, assuming that the suppliers' cost of effort and type distribution were known? Since the estimation did not impose optimality of the buyer's chosen policy, the performance-optimal policy could potentially be different from the estimated one. Given the supplier-side estimates, I can computationally solve the buyer's problem for any candidate policy, hence making it possible to compare different policies in terms of the expected discounted performance measure in Equation 2.

The general space of policies that the buyer could choose from is intractably large, so I instead consider the class of policies indexed by the specific parametrization of p, τ specified in Subsection 5.2. In particular, the policies are indexed by the base transition parameters λ and ζ, the performance reward κ, and the tier choice weights ϕ. This reduces the exercise to an optimization problem that selects the optimal policy:

$$p^*, \tau^* = \arg\max V(p, \tau) \tag{13}$$

over a policy space represented by 10 parameters, taking as given the observed order arrival sequence and estimated supplier parameters. For each candidate policy, the supplier problems in Equation 1 are solved to get the optimal supplier effort responses to the policy.

The optimal policy parameters are presented in Table 10. It is approximately a two tier policy - the tier choice weights ϕ indicate a very high preference for Gold suppliers, and in their absence, a very high preference for New suppliers relative to both Silver and Bronze. So the policy creates a highly preferred set of suppliers, and goes to the outer pool with large probability if there are no highly preferred suppliers available. The transition dynamics of the policy, displayed in Figure 18, involves an extreme performance bonus: the probability of being upgraded to a Gold supplier goes from close to 0 at a performance of 0.7, to being around 1 at the highest possible performance of 1. This leads to the Gold tier being filled with very high types, which complements the policy's high preference for Gold suppliers.

I perform the incentive-selection decomposition exercise for the optimal policy as well, and compare the result to the estimated policy in Figure 19. The optimal policy significantly improves the overall measure of discounted performance, almost exclusively by increasing the selection effect. This largely comes down to the extreme performance bonus in the optimal policy, which leads to extreme sorting of types by tier. Since type and effort are substitutes in the production process, the very high types selected by the optimal policy choose significantly

lower effort, hence reducing the incentive effect of the larger performance bonus.

A noteworthy aspect here is that the optimal policy does not take the logic of this sorting to its natural limit, i.e. a cutoff policy. As we can see in Figure 18, lower types still have a non zero probability of transitioning to Gold sometimes, and higher types have a non zero probability of being Dropped for performance. For the sake of comparison, I also solve for the optimal cutoff policy within the estimated supplier environment, which keeps suppliers in the active pool only when their monthly performance exceeds the cutoff of $y_c = 0.88$, and selects randomly from this active pool when available or selects a new supplier when the pool is empty. The PDV of performance from this optimal cutoff policy is included in Figure 19.

There are two reasons why the general optimal policy is not a cutoff policy. The first is that a cutoff policy provides very little incentives to the types above or sufficiently below the chosen cutoff. This follows due to the estimated effort cost being high enough to discourage consistently large effort choices (the cutoff policy would require a supplier to consistently perform above the cutoff every time). As a result, the incentive effect of a cutoff policy is close to zero. The second reason is more subtle. The buyer takes an expectation over possible paths of realized type draws of suppliers over time when evaluating her objective. Given the estimated distribution of types being somewhat dispersed, a reasonable portion of these paths could start with type draws that are in some band above the average type. Since the optimal cutoff turns out to be quite a bit higher than the average type, for these paths, the optimal cutoff policy will keep dropping these initial suppliers. This means that the cutoff policy will keep getting the average type as the outcome for these paths. The optimal policy improve on this by retaining and incentivizing at least some of the suppliers with types between the average and the optimal cutoff.

7.3 Robustness of the estimated policy

The optimal policy differs from the estimated one in some key ways. It is approximated by a policy that employs just two tiers and it has a stark performance bonus. To understand why the buyer might be justified in using the estimated policy instead of the optimal one, I perform two counterfactuals that exhibit the potential drawbacks of the simple structure of the optimal policy.

The first counterfactual is motivated by the idea that the buyer's draws from the outside pool may be evolving over time and that robustness to this evolution may be desirable. At the start of operations, she might have started with a smaller, weaker set of contacts. Over time,

an established reputation could have enabled access to larger and better set of suppliers. I capture this scenario by assuming that in the first third of the sample, the buyer draws types from Distribution 1 in Figure 21, and then transitions to Distribution 2 afterwards. These Beta distributions are chosen such that the weighted combination of their mean and variance (weighted by the relative time lengths that they would be operational in) are equal to the mean and variance of the estimated type distribution. In such an environment, the gap between the optimal and the estimated policies reduces considerably, as Figure 22 shows. This is a result of the optimal policy's simplistic tier structure. In the initial stage, since the optimal policy has approximately an "in or out" logic, it largely fails to build a good inner pool of suppliers. This significantly reduces its effectiveness, until the transition to the high type regime.

I take this exercise even further in the second counterfactual, by constructing a distribution of types for which the estimated policy actually performs better than the optimal policy. The distribution is plotted in Figure 23, and the value comparisons are in Figure 24. The reason why this happens is intuitive: the optimal policy leans aggressively into selecting very high types, with moderate and low types being dropped with high probabilities. As a result, when faced with a distribution of mostly moderate and low types, it achieves almost no selection, and achieves much less incentivization since it provides no marginal incentives on performance for those types.

The main takeaway from these exercises is that the buyer chosen policy is more robust to high levels of initial uncertainty about the supplier environment. One possible concern is that this would also justify continually updating the policy as the buyer's uncertainty about the environment reduces over time, but it can be logistically quite challenging to change business strategy regularly. It takes time for the implementation of the strategy to become stable, and even more time for its result to become clear. Moreover, a steady and stable policy is important to ensure that suppliers understand, trust, and buy-in to the buyer's system, so changing it regularly can in fact backfire.

7.4 Monetary incentives versus relational incentives

A unique feature of this setting is that the buyer has implemented a relational strategy using dynamic business allocation. This naturally leads to the question: what would change if monetary contracts were used instead?

To answer this question, I conduct a simple benchmark that compares the relationship-based approach to an approach that relies on one-off spot monetary contracts. Specifically,

suppose that the buyer draws a new supplier every time from the outside pool. The supplier is offered a contract that pays her the standard amount that we observe in the data, and additionally, has a linear monetary bonus on performance, i.e., the supplier will earn himself an additional payment of αy on achieving a pass rate of y.[24] For this benchmark, I ask: for what α does the buyer ensure herself the same expected PDV of performance as the estimated policy? The answer is that the buyer must choose $\alpha^* = 1.6\bar{b}$, where $\bar{b}$ is the average margin that they get from an order. Given that the average supplier margin in the data is around 14% of total payments,[25] this implies that buyer would need to increase payments to the supplier by 18%,[26] which is a relatively large amount. This simple exercise gives us a way to contextualize the economic value of the buyer's relationship-based approach.

7.5 The value of relational scale

The scale of the buyer's intermediated platform is central to a strategy of using business allocations to drive supplier performance. A larger number of orders can enable the buyer to learn more quickly by trying out more suppliers, and also provides a larger relational value budget with which to build relationships that incentivize suppliers. To quantitatively investigate the importance of scale, I perform two exercises.

First, I remove the buyer's intermediated services entirely and simulate a scenario where downstream customers contract directly with suppliers, using the estimated policy. In this world, each customer maintains their own internal supplier pool, and draws from the estimated external type distribution each time that they pick a new supplier. This strips out all sources of relational scale—it slows down learning, it reduces relational incentives, and it also reduces competition incentives by increasing supplier market power—and the magnitude of reduction in performance gives us an absolute value of relational scale.

A follow-up question is whether performance improvements grow proportionally with intermediary size. In a second, less extreme counterfactual, I simulate an environment where the buyer is split into two independent, smaller intermediaries, each handling half of the order flow and using the estimated policy. Both start with no suppliers initially, draw suppliers from the same estimated external type distribution, and each maintains its own separate supplier

[24] Multiple factors, including the extremely large amounts of money required to buy steel for an order and the highly credit-constrained nature of the suppliers, lead to the necessity of a base level of payments that must be made upfront to the supplier for an order. Additionally, limited liability implies that the suppliers cannot usually be made to make payments to the buyer.

[25] Calculated as total payments net of input costs and wages.

[26] Calculated as the required bonus payment $\alpha y = (1.6\bar{b})y = (1.6 \times 0.14) \times 0.79 = 0.17696 \approx 18\%$.

pool. By comparing the performance improvement generated by these two smaller entities to the performance increase for the single integrated buyer, we can assess whether relational scale exhibits increasing or decreasing returns.

The results of these exercises are demonstrated in Figure 20. Splitting the buyer into two only marginally reduces value, from around 12% value added to 11%. However, eliminating the intermediation entirely leads to a significant drop - to around 3%. A large part of this drop is driven by a drop in the selection effect, i.e. the one-on-one interactions take much longer to find the top suppliers. Taken together, these results indicate that a larger relational scale does enable higher supply chain performance, but the marginal improvement in performance is significantly smaller at larger scales.

8 Conclusion

This paper has shown that firms can create significant value by governing their supply relationships with a dynamic policy of selective business allocation. The buyer leverages the promise of future business to design a tier system that effectively addresses the twin challenges of selecting and incentivizing suppliers that are inherent in the process of building and maintaining a supply chain. The economic significance of this relational approach is underscored by the finding that achieving the same performance boost as the buyer's policy through simple monetary incentives would be expensive, requiring bonuses that dwarf typical profit margins.

An important insight is that competition, often viewed as a force that erodes cooperation, can be a productive tool when deployed internally. By managing the size and composition of the supplier pool, the buyer calibrates competitive pressure to improve overall performance. While a counterfactual optimal policy performs significantly better by leaning heavily on selection, the buyer's observed choice of a more graduated system suggests a preference for robustness—a strategy that performs better when the underlying type distribution is worse. Thus, the paper's quantitative analysis also provides insights into why seemingly sub-optimal policies and contracting may be observed in practice.

The scope of this analysis is, by design, centered on a single buyer's policy. Even though I have argued that some of the economic mechanisms are generalizable, the specifics of the model are context-dependent. Future research could build on this foundation in several directions. A natural extension would be to endogenize the supplier's persistent type, allowing suppliers to make investments in their own capabilities in response to the incentives provided by the buyer's system. Another fruitful avenue would be to embed the firm's policy problem

within a market equilibrium model, exploring the interaction between a firm's internal governance structure and external competition among buyers for the suppliers. Such work would further deepen our understanding of how firms architect and manage the relationships that form the backbone of modern supply chains.

References

Ackerberg, Daniel A and Maristella Botticini (2002). "Endogenous matching and the empirical determinants of contract form". In: *Journal of Political Economy* 110.3, pp. 564–591.

Aguirregabiria, Victor and Pedro Mira (2007). "Sequential estimation of dynamic discrete games". In: *Econometrica* 75.1, pp. 1–53.

Allman, Elizabeth S, Catherine Matias, and John A Rhodes (2009). "Identifiability of parameters in latent structure models with many observed variables". In.

Andrews, Isaiah and Daniel Barron (2016). "The allocation of future business: Dynamic relational contracts with multiple agents". In: *American Economic Review* 106.9, pp. 2742–59.

Ashraf, Nava et al. (2020). "Losing prosociality in the quest for talent? Sorting, selection, and productivity in the delivery of public services". In: *American Economic Review* 110.5, pp. 1355–1394.

Athey, Susan and Guido W Imbens (2019). "Machine learning methods that economists should know about". In: *Annual Review of Economics* 11.1, pp. 685–725.

Bajari, Patrick, C Lanier Benkard, and Jonathan Levin (2007). "Estimating dynamic models of imperfect competition". In: *Econometrica* 75.5, pp. 1331–1370.

Bajari, Patrick, Christina Dalton, et al. (2014). "Moral hazard, adverse selection, and health expenditures: A semiparametric analysis". In: *The RAND Journal of Economics* 45.4, pp. 747–763.

Billingsley, Patrick (1995). *Probability and Measure*. 3rd. Wiley Series in Probability and Mathematical Statistics. New York: John Wiley & Sons. ISBN: 978-0-471-00710-4. DOI: 10.1002/9780470316658.

Board, Simon (2011). "Relational contracts and the value of loyalty". In: *American Economic Review* 101.7, pp. 3349–3367.

Booth, James G and James P Hobert (1999). "Maximizing generalized linear mixed model likelihoods with an automated Monte Carlo EM algorithm". In: *Journal of the Royal Statistical Society Series B: Statistical Methodology* 61.1, pp. 265–285.

Cajal-Grossi, Julia, Rocco Macchiavello, and Guillermo Noguera (2023). "Buyers' sourcing strategies and suppliers' markups in Bangladeshi garments". In: *The Quarterly Journal of Economics* 138.4, pp. 2391–2450.

Carroll, Raymond J and Peter Hall (1988). "Optimal rates of convergence for deconvolving a density". In: *Journal of the American Statistical Association* 83.404, pp. 1184–1186.

David, Etienne et al. (2024). "Identifiability of discrete input–output hidden Markov models with external signals". In: *Statistics and Computing* 34.1, p. 54.

Dempster, Arthur P, Nan M Laird, and Donald B Rubin (1977). "Maximum likelihood from incomplete data via the EM algorithm". In: *Journal of the royal statistical society: series B (methodological)* 39.1, pp. 1–22.

Einav, Liran et al. (2013). "Selection on moral hazard in health insurance". In: *American Economic Review* 103.1, pp. 178–219.

Fan, Jianqing (1991). "On the optimal rates of convergence for nonparametric deconvolution problems". In: *The Annals of Statistics*, pp. 1257–1272.

Gartner, Inc. (2025). *Supplier Relationship Management: A Complete Guide*. https://www.gartner.com/en/supply-chain/topics/supplier-relationship-management. Accessed 2025-10-11.

Harris, Adam and Thi Mai Anh Nguyen (2025). "Long-term relationships and the spot market: Evidence from us trucking". Working paper.

Helper, Susan and Abdul Munasib (Apr. 2022). *Economies of Scope and Relational Contracts: Exploring Global Value Chains in the Automotive Industry*. BEA Working Paper WP2022-5. U.S. Bureau of Economic Analysis (BEA).

Hornik, Kurt, Maxwell Stinchcombe, and Halbert White (1989). "Multilayer feedforward networks are universal approximators". In: *Neural networks* 2.5, pp. 359–366.

Hotz, V Joseph and Robert A Miller (1993). "Conditional choice probabilities and the estimation of dynamic models". In: *The Review of Economic Studies* 60.3, pp. 497–529.

Hu, Yingyao and Matthew Shum (2012). "Nonparametric identification of dynamic models with unobserved state variables". In: *Journal of Econometrics* 171.1, pp. 32–44.

Jääskeläinen, Aki (2021). "The relational outcomes of performance management in buyer-supplier relationships". In: *International Journal of Production Economics* 232, p. 107933.

Jeziorski, Przemyslaw, Elena Krasnokutskaya, and Olivia Ceccarini (2017). "Adverse selection and moral hazard in the dynamic model of auto insurance". UC Berkeley, Haas School of Business working paper.

Lazear, Edward P (2000). "Performance pay and productivity". In: *American Economic Review* 90.5, pp. 1346–1361.

Lazear, Edward P and Sherwin Rosen (1981). "Rank-order tournaments as optimum labor contracts". In: *Journal of political Economy* 89.5, pp. 841–864.

Li, Hongmin, Hao Zhang, and Charles H Fine (2013). "Dynamic business share allocation in a supply chain with competing suppliers". In: *Operations Research* 61.2, pp. 280–297.

Liang, Liping and Derek Atkins (2021). "Rewarding suppliers' performance via allocation of business". In: *Manufacturing & Service Operations Management* 23.2, pp. 331–345.

Louis, Thomas A (1982). “Finding the observed information matrix when using the EM algorithm”. In: *Journal of the Royal Statistical Society Series B: Statistical Methodology* 44.2, pp. 226–233.

Macchiavello, Rocco and Ameet Morjaria (2015). “The value of relationships: evidence from a supply shock to Kenyan rose exports”. In: *American Economic Review* 105.9, pp. 2911–45.

— (2021). “Competition and relational contracts in the Rwanda coffee chain”. In: *The Quarterly Journal of Economics* 136.2, pp. 1089–1143.

Maestrini, Vieri et al. (2018). “Effects of monitoring and incentives on supplier performance: An agency theory perspective”. In: *International Journal of Production Economics* 203, pp. 322–332.

Matzkin, Rosa L (2003). “Nonparametric estimation of nonadditive random functions”. In: *Econometrica* 71.5, pp. 1339–1375.

— (2007). “Nonparametric identification”. In: *Handbook of econometrics* 6, pp. 5307–5368.

Murphy, Kevin M and Robert H Topel (2002). “Estimation and inference in two-step econometric models”. In: *Journal of Business & Economic Statistics* 20.1, pp. 88–97.

Newey, Whitney K and Daniel McFadden (1994). “Large sample estimation and hypothesis testing”. In: *Handbook of econometrics* 4, pp. 2111–2245.

Pakes, Ariel, Michael Ostrovsky, and Steven Berry (2007). “Simple estimators for the parameters of discrete dynamic games (with entry/exit examples)”. In: *the RAND Journal of Economics* 38.2, pp. 373–399.

Pech, Martin and Drahoš Vaněček (2020). “Supplier performance management in context of size and sector characteristics of enterprises”. In: *Quality Innovation Prosperity* 24.1, pp. 88–103.

Schervish, Mark J (2012). *Theory of statistics*. Springer Science & Business Media.

Shapiro, Carl and Joseph E Stiglitz (1984). “Equilibrium unemployment as a worker discipline device”. In: *The American economic review* 74.3, pp. 433–444.

Syverson, Chad (2011). “What determines productivity?” In: *Journal of Economic literature* 49.2, pp. 326–365.

Train, Kenneth E (2009). *Discrete choice methods with simulation*. Cambridge university press.

Wei, Greg CG and Martin A Tanner (1990). “A Monte Carlo implementation of the EM algorithm and the poor man’s data augmentation algorithms”. In: *Journal of the American statistical Association* 85.411, pp. 699–704.

Zhou, Rui et al. (2022). “A stochastic programming model with endogenous uncertainty for selecting supplier development programs to proactively mitigate supplier risk”. In: *Omega* 107, p. 102542.

A Exhibits

A.1 Tables from Section 2

Table 1: Summary of orders dataset

Metric	Mean	Min	Max	SD
Order-level stats (N=896)				
Pass rate	0.79	0.00	1.00	0.32
Weight (MT)	99	0.1	1000	140
Days to complete order	53	1	199	58
Supplier-level stats (N=274)				
Number of orders	3	1	38	4
Pass rate	0.75	0.00	1.00	0.27
Capacity (MT)	99	10	500	78
Days to complete order	52	1	199	51

Notes: Pass rate is defined as the fraction of parts that passed the first inspection. MT is short-form for Metric Tonnes.

Table 2: Supplier tier composition across snapshots

Metric	Mean	Min	Max
Snapshot 1			
Total Active (G+S+B)	9	2	16
Gold	3	1	5
Silver	4	0	8
Bronze	3	0	6
Dropped	5	1	13
Snapshot 2			
Total Active (G+S+B)	8	2	12
Gold	3	1	6
Silver	3	0	7
Bronze	2	0	4
Dropped	14	8	25
Snapshot 3			
Total Active (G+S+B)	4	1	9
Gold	2	1	5
Silver	1	0	4
Bronze	1	0	3
Dropped	19	10	34

Notes: This table reports summary statistics on the number of suppliers per cluster for each tier at the three snapshot months. The Mean column reports the nearest round number. G denotes Gold, S denotes Silver, and B denotes Bronze.

A.2 Tables from Section 3

Table 3: Buyer responsiveness to supplier performance

	Dependent variable:		
	Total lifetime orders	Orders per month	Prob. of getting an order
		OLS	*Multinomial Logit*
	(1)	(2)	(3)
Lifetime performance	2.97***		
	(0.98)		
Past average performance		0.25***	2.85***
		(0.04)	(0.39)
Constant	1.39**	-0.16***	-6.92***
	(0.66)	(0.05)	(1.05)
Month FE		Yes	Yes
Observations	274	3,424	3,424

Notes: This table shows how the buyer rewards better-performing suppliers. **Column (1)** presents an OLS regression where the unit of observation is a supplier (N=274), showing that a higher lifetime average pass rate is associated with more total orders. **Columns (2) and (3)** examine the monthly order assignment decision (N=3,424 supplier-months). Column (2) is an OLS regression and Column (3) is a multinomial logit, both showing that higher cumulative performance up to a given month increases the likelihood of receiving an order in that month. The second and third regressions include month fixed effects. Standard errors are in parentheses. *p<0.1; **p<0.05; ***p<0.01.

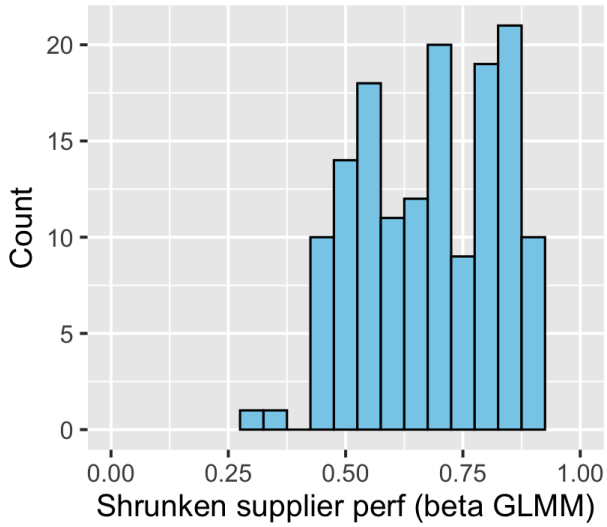
Table 4: Supplier performance drops when steel prices are higher

Dependent Variable:	Quality check passed (0/1)	
Model:	(1) OLS	(2) Logit
Steel price (in 10k Rs/MT)	-0.042**	-0.413**
	(0.019)	(0.199)
Order Fixed Effects	Yes	Yes
Observations	73,844	52,293
R^2 / Pseudo R^2	0.75	0.32

Notes: This table presents regressions where the unit of observation is an individual manufactured part. The dependent variable is an indicator for whether the part passed its first quality inspection. “Steel Price" is the price of steel in the month that the steel for the part was purchased, in 10k Rupees per Metric Tonne (Rs/MT). Both models include order-level fixed effects to control for time-invariant order characteristics. Standard errors, clustered at the lot level, are in parentheses. **p<0.05, ***p<0.01.

A.3 Figures from Section 3

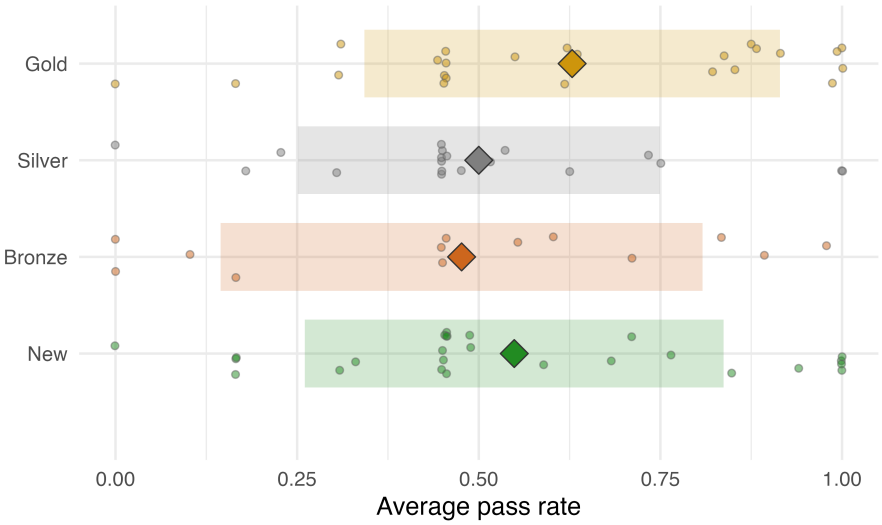
Figure 1: There is large variation in supplier-level performance



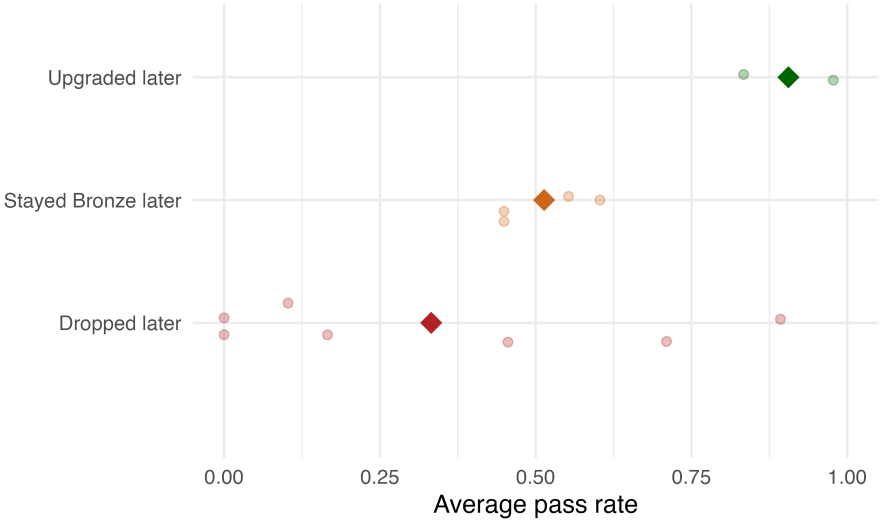
Notes: This figure plots the histogram of supplier-level mean performance on orders after applying Bayes shrinkage using a beta-GLMM (Generalized Linear Mixed Model). The GLMM is an extension of the standard Gaussian LMM (Linear Mixed Models) that allows the dependent variable to be a proportion with a non-Gaussian distribution, as is the case with order performance.

Figure 2: Supplier performance by tier and subsequent transition

(a) Supplier performance is heterogeneous by tier.

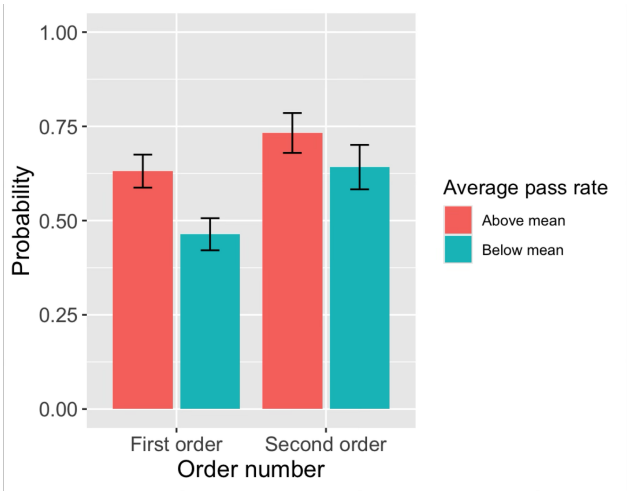


(b) Performance of Bronze suppliers, by subsequent tier transition.



Notes: In these plots, performance is defined as the fraction of parts that pass the first inspection. Figure (a) shows the distribution of supplier performance during months corresponding to tier snapshots. Figure (b) shows the distribution of pass rates specifically for suppliers who were in the Bronze tier during a snapshot month, partitioned by their tier in the *next* snapshot. In (b), each point represents a supplier and the diamond marks the mean pass rate for each group, showing that Bronze suppliers who were subsequently upgraded performed significantly better than those later dropped from the pool.

Figure 3: The buyer is more discriminating of performance in the first order.



Notes: This figure shows the probability that a supplier receives at least one more order, conditional on performance on first and second orders. Performance is classified as "Above mean" or "Below mean" relative to the sample average pass rate for that order number.

Figure 4: The mix of supplier tiers changes over time.

(a) Mix of tiers including dropped suppliers.

(b) Mix of tiers for active suppliers only.

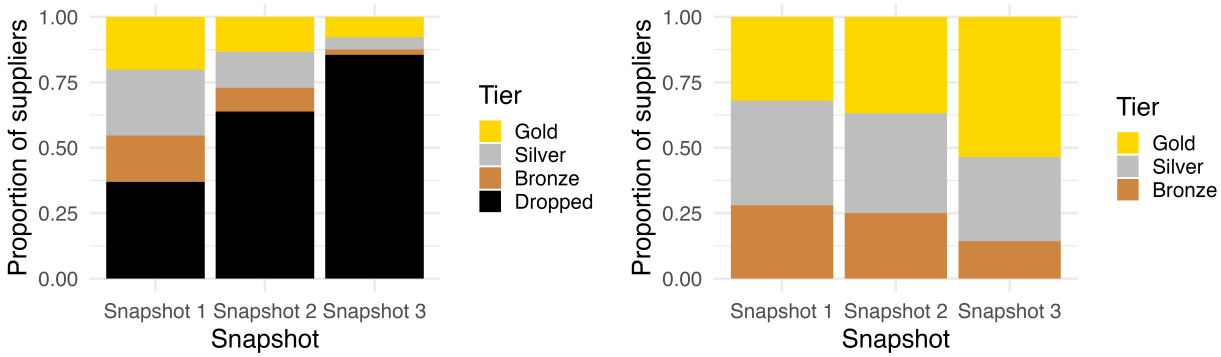
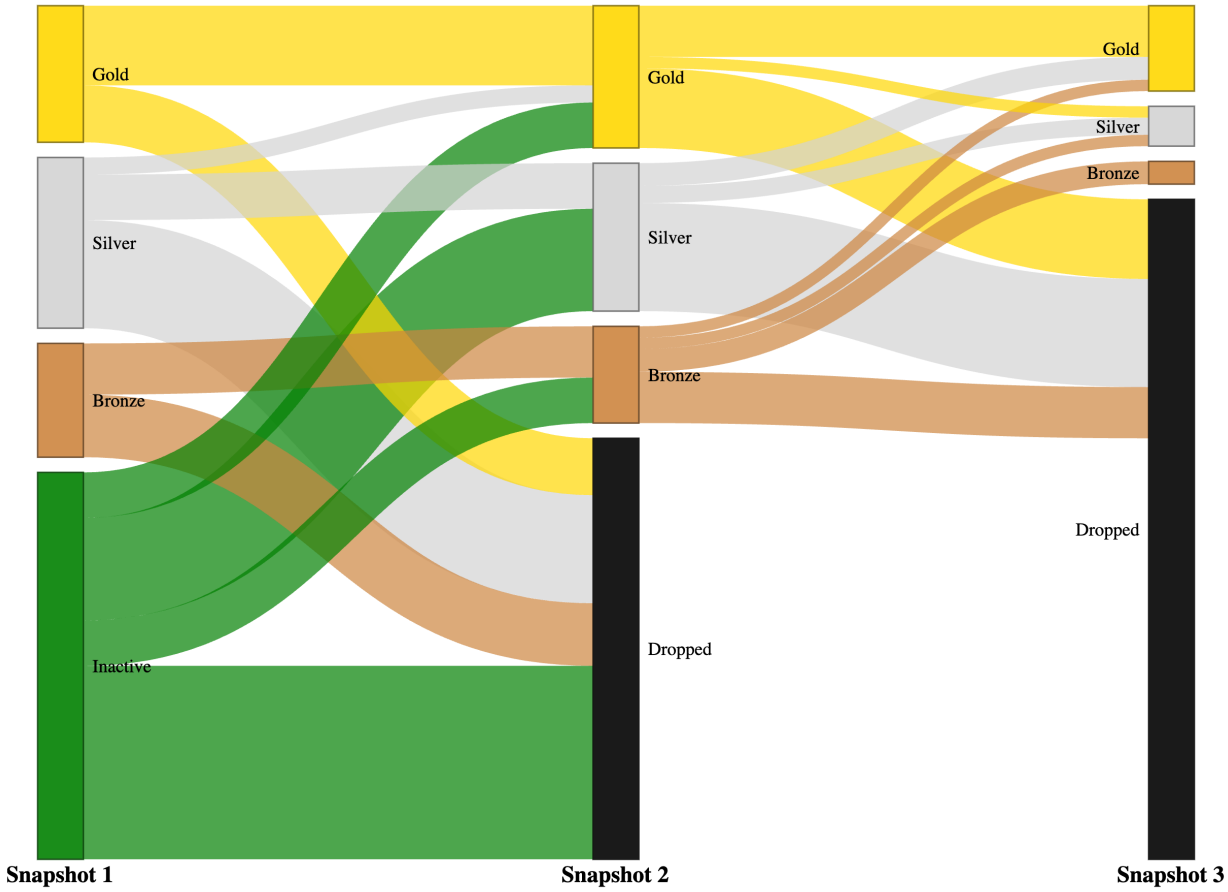
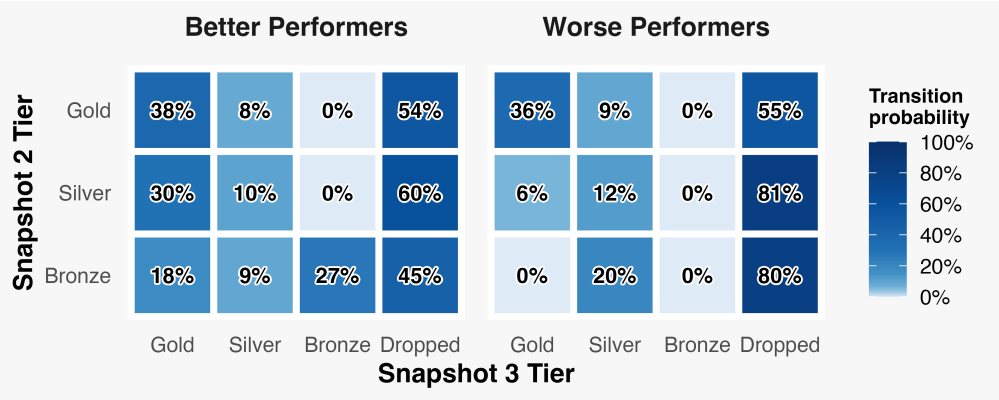


Figure 5: Movement of suppliers across tiers over the snapshots.



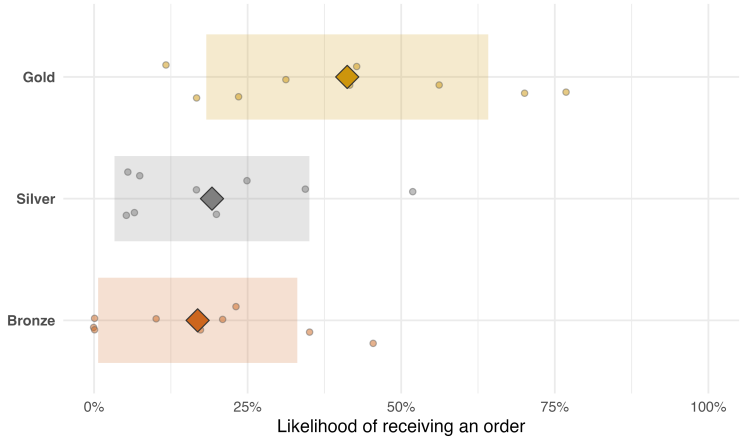
Notes: This diagram tracks suppliers who were observed to get an order between the first two snapshots. The Inactive category contains suppliers who had not yet entered the supplier pool at Snapshot 1 i.e. new suppliers.

Figure 6: Heatmap of transitions from Snapshot 2 to Snapshot 3, demonstrating the effect of performance by starting tier.



Notes: The suppliers that were observed to get an order between Snapshots 2 and 3 were divided into Better Performers and Worse Performers using the median performance in this set. The transition probabilities for Snapshot 2 to Snapshot 3 are calculated from the tier movements of these suppliers.

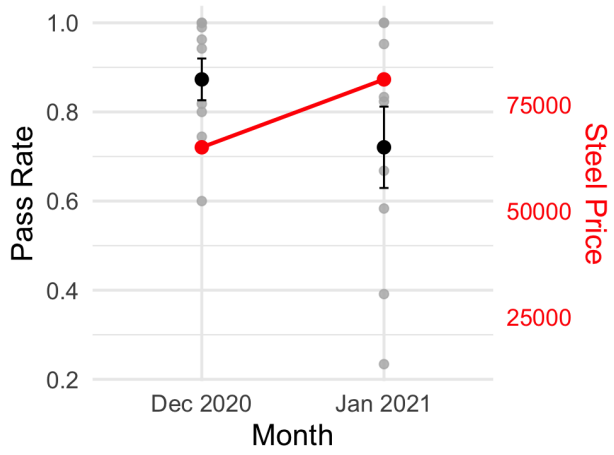
Figure 7: Gold suppliers are more likely to get an order.



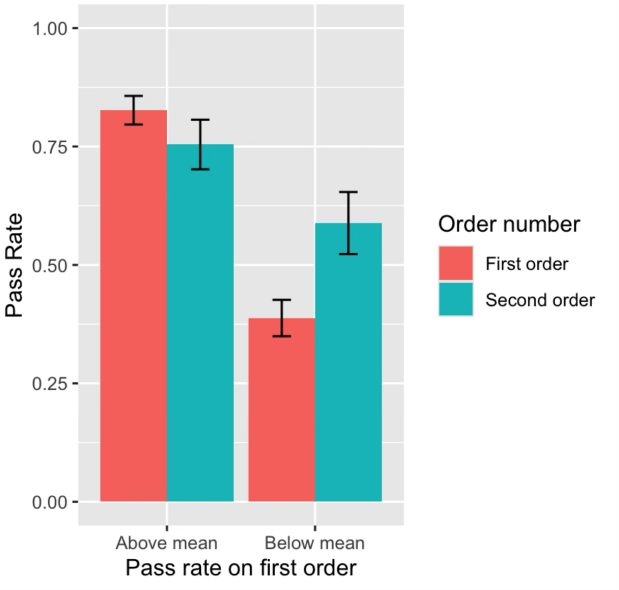
Notes: The observations in this plot are at the cluster times snapshot level. I calculate the probability a specific tier (say, Gold) being chosen by calculating the fraction of the number of suppliers in that tier that received an order divided by the number of suppliers of that tier that were active.

Figure 8: Suggestive evidence of moral hazard

(a) Performance reduces when steel price increases.



(b) Performance response to past outcomes.



Notes: Both plots show how supplier performance responds to changing incentives. Figure (a) demonstrates that pass rates fall following a large steel price shock, suggesting suppliers may reduce costly effort when inputs become more expensive. This plot only includes orders with steel purchased in both Dec 2020 and Jan 2021. Figure (b) shows that suppliers who perform below average on their first order tend to improve on their second, consistent with responding to the buyer's relational incentives. This plot includes all suppliers with at least two orders in the dataset.

A.4 Tables from Section 6

Table 5: First stage parameters estimates: The buyer's policy

Object	Parameter	Description	Estimate	SE
τ	λ_u	Base upgrade weight	-3.53	1.19
	λ_s	Base retain weight	0.79	0.67
	λ_d	Base downgrade weight	-0.30	0.66
	κ	Performance bonus	6.06	0.90
	μ_G	Base G weight (new supp)	-5.26	0.89
	μ_S	Base S weight (new supp)	-2.09	0.38
	μ_B	Base B weight (new supp)	-1.98	0.39
	ν	Drop probability shift wrt active supplier count	0.30	0.09
	ψ	Exogenous drop shock (logistic-scale)	-3.22	0.21
p	ϕ_G	Choice weight for G	-0.15	0.07
	ϕ_S	Choice weight for S	-0.24	0.17
	ϕ_B	Choice weight for B	-0.79	0.33
	ζ_1	Availability shifter (pending orders in logMT)	-1.22	0.25
	ζ_2	Availability shifter (capacity in logMT)	0.46	0.14

Notes: This table presents the maximum likelihood estimates of the buyer's policy parameters from the Stage 1 estimation, using the MCEM algorithm described in the text. The parameters control tier transitions (τ) and order allocation (p). Standard errors are computed using Louis's method (see Appendix D).

Table 6: Choice probabilities implied by Stage 1 parameters

Model Object	Value
P(Choose Gold | available)	0.23
P(Choose Silver | available)	0.20
P(Choose Bronze | available)	0.15
P(Choose New)	0.35
Baseline availability (σ)	0.53
$\Delta\sigma$ (Workload: 50 $\rightarrow$ 100)	-0.09
$\Delta\sigma$ (Capacity: 99 $\rightarrow$ 198)	0.08

Notes: This table shows the implied order allocation probabilities based on the Stage 1 parameter estimates. Probabilities are calculated for the median supplier pool with two Gold, two Silver, and two Bronze competitors. The top panel shows the probability of a particular supplier of a given tier being chosen, conditional on being available. The bottom panel shows how the baseline availability (σ) is affected by a doubling of a supplier's pending workload or total capacity.

Table 7: Transition probabilities implied by Stage 1 parameters

From \ To	Gold	Silver	Bronze	Dropped
Gold	0.64 0.55 / 0.81	0.16 0.20 / 0.10	-	0.20 0.25 / 0.08
Silver	0.36 0.21 / 0.64	0.27 0.34 / 0.18	0.16 0.20 / 0.10	0.20 0.25 / 0.08
Bronze	-	0.45 0.28 / 0.71	0.34 0.45 / 0.20	0.22 0.27 / 0.09
New	0.49 0.31 / 0.77	0.12 0.16 / 0.07	0.11 0.15 / 0.07	0.28 0.38 / 0.09

Note: Calculated using mean supplier performance $y = 0.75$. The values below in DarkRed / DarkGreen correspond to $y_{25} = 0.62$ / $y_{75} = 0.92$.

Table 8: Second stage parameter estimates: supplier side

Parameter	Description	Estimate	SE
η	Parameter 1 for Beta prior on type	1.91	0.62
ι	Parameter 2 for Beta prior on type	0.82	0.19
γ	Cost scale (relative to avg order benefit)	0.29	0.08
α	Exponent for effort	0.36	0.11
ρ	Effort, type complementarity	-2.89	0.46
φ	Lot-level precision parameter	0.93	0.07

Notes: This table presents the simulated maximum likelihood estimates of the supplier-side parameters from the Stage 2 estimation. These parameters are estimated conditional on the Stage 1 policy estimates. The parameters define the Beta prior on latent supplier type, the cost of effort function, and the technology that maps type and effort to outcomes. Standard errors account for Stage 1 estimation uncertainty (see Appendix D).

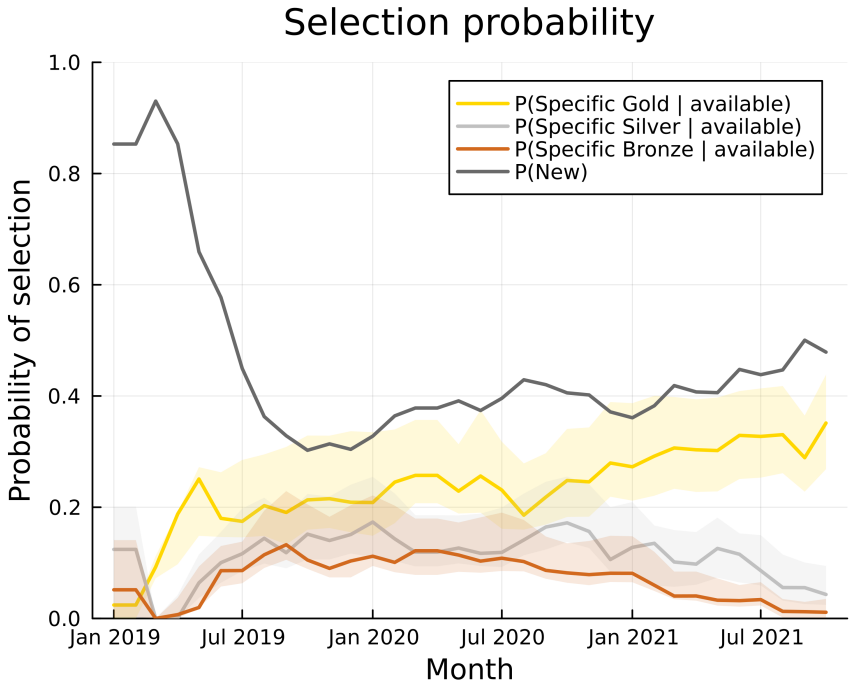
Table 9: Model object values implied by estimated Stage 2 parameters

Category	Mean	P25	P75	SD
Type (per supplier, from prior)	0.70	0.56	0.91	0.21
Effort contribution (per lot, from simulations)	0.03	0.02	0.05	0.03
Pass rate (per order, from data)	0.79	0.64	0.94	0.17

Notes: This table reports summary statistics for objects from the model, implied by the full set of estimated parameters. Type statistics are calculated from the estimated Beta prior distribution. Effort contribution is the mean of $\mu(\theta, e^*) - \theta$ across 1000 simulations of the full model. The pass rate row is the average per-order pass rate in the data.

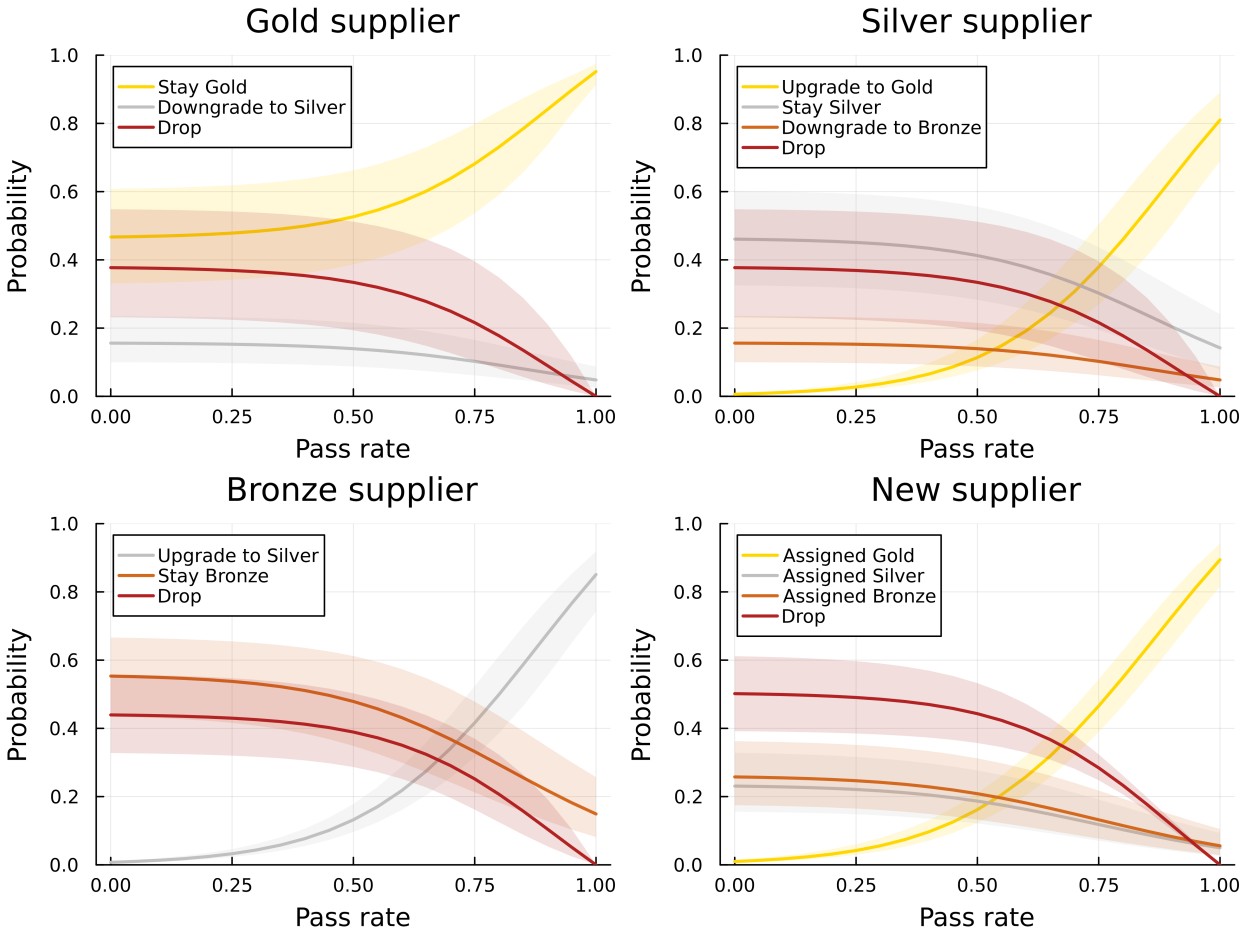
A.5 Figures from Section 6

Figure 9: Selection probabilities change significantly over time as the supplier mix changes.



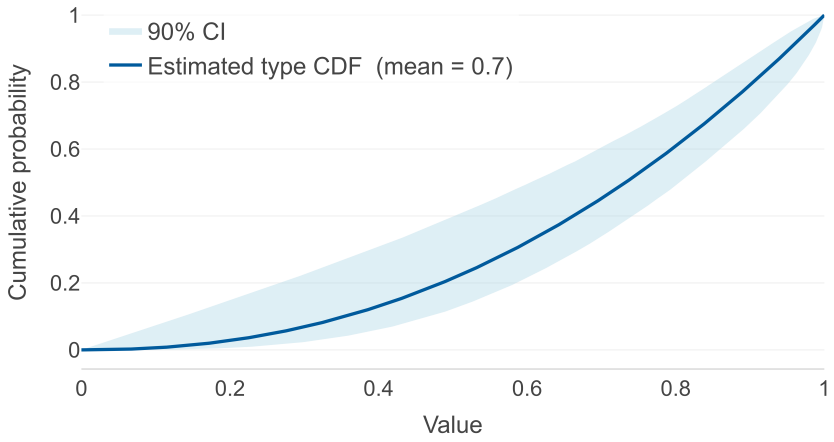
Notes: This graph plots the choice probability for each tier implied by Stage 1 estimates, given the set of orders, the assigned suppliers, the performance of the suppliers, and the snapshot tiers. P(Specific T | available) plots the probability of choosing an individual supplier of tier T conditional on being available. The bands represent 90% confidence intervals.

Figure 10: Higher performance leads to significantly better transitions.



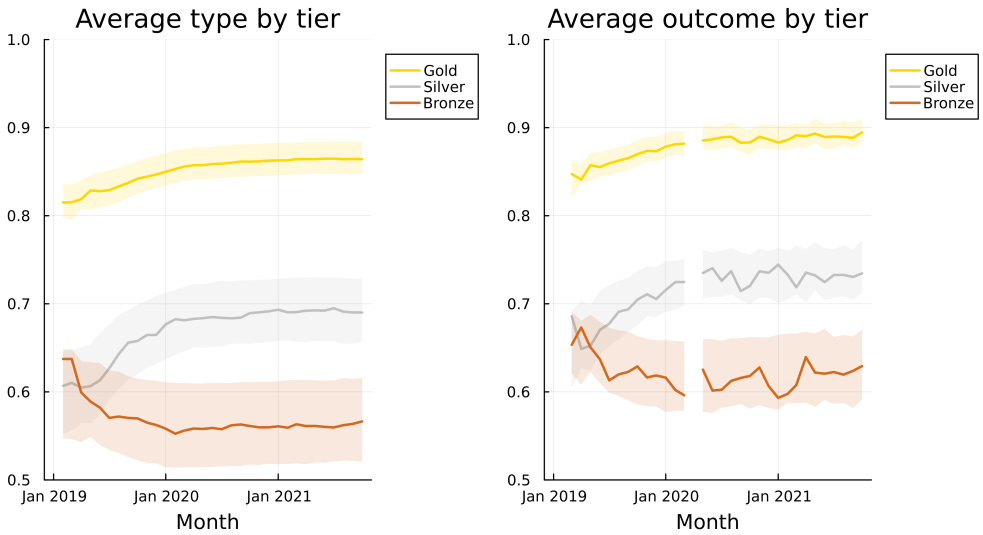
Notes: This graph plots the change in probability of transitions as performance changes, for each tier. It does not include the exogenous relationship-ending shocks. The bands represent 90% confidence intervals.

Figure 11: The estimated type distribution.



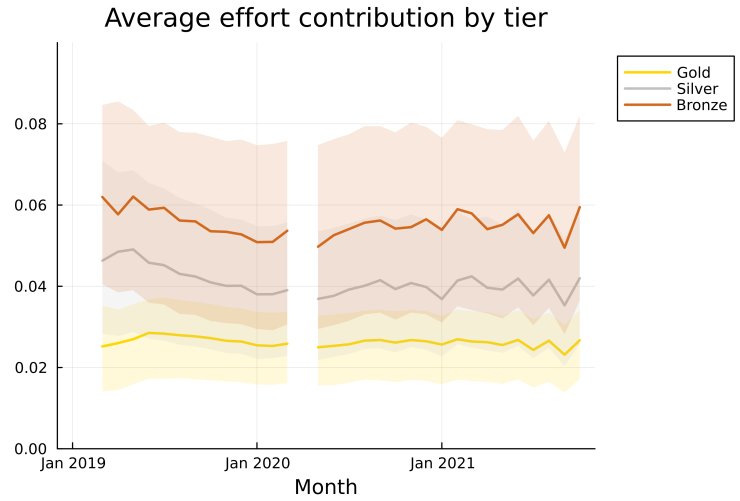
Notes: This graph plots the CDF of the parametrized type distribution, Beta(ι, η), at the estimated parameters from Stage 2 (see Table 8). The bands represent 90% confidence intervals.

Figure 12: Supplier outcomes are less disperse across tiers than supplier types.



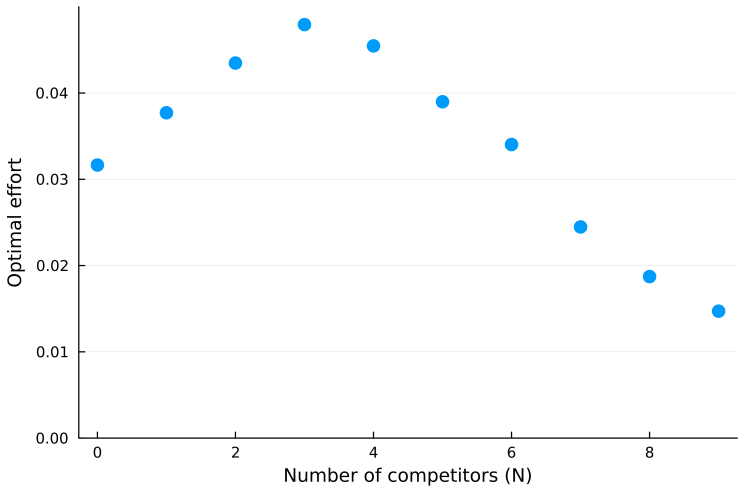
Notes: Averages calculated over 1000 simulations of the full model. The left graph calculates average type of all active suppliers in each tier each month, while the right graph calculates average outcomes by selected suppliers conditional on tier each month. The bands represent 90% confidence intervals.

Figure 13: Lower tier suppliers put in higher effort.



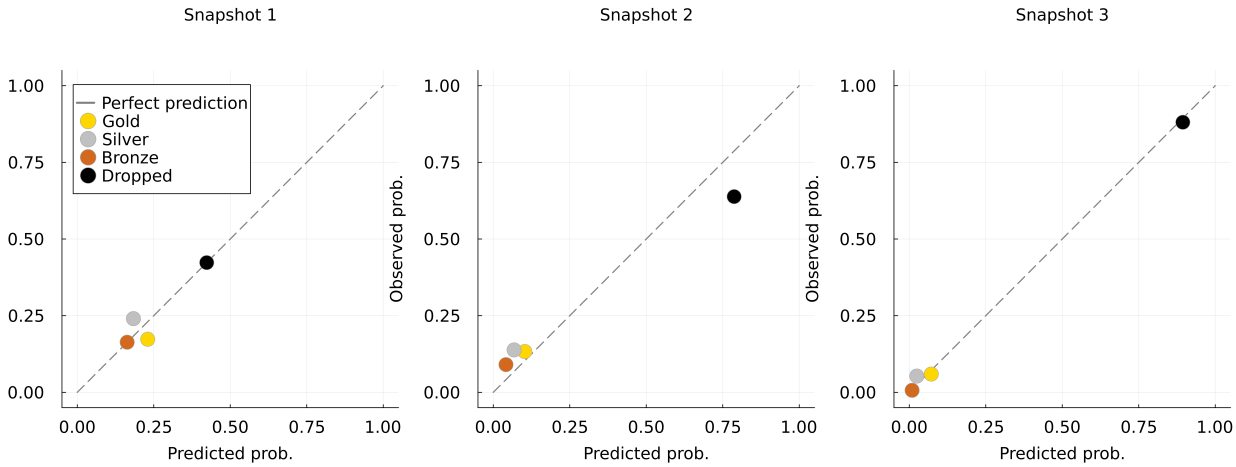
Notes: Averages calculated over 1000 simulations of the full model. The left graph calculates average type of all active suppliers in each tier each month, while the right graph calculates average outcomes by selected suppliers conditional on tier each month. The bands represent 90% confidence intervals.

Figure 14: Supplier effort follows an inverted U-shape in number of competitors.



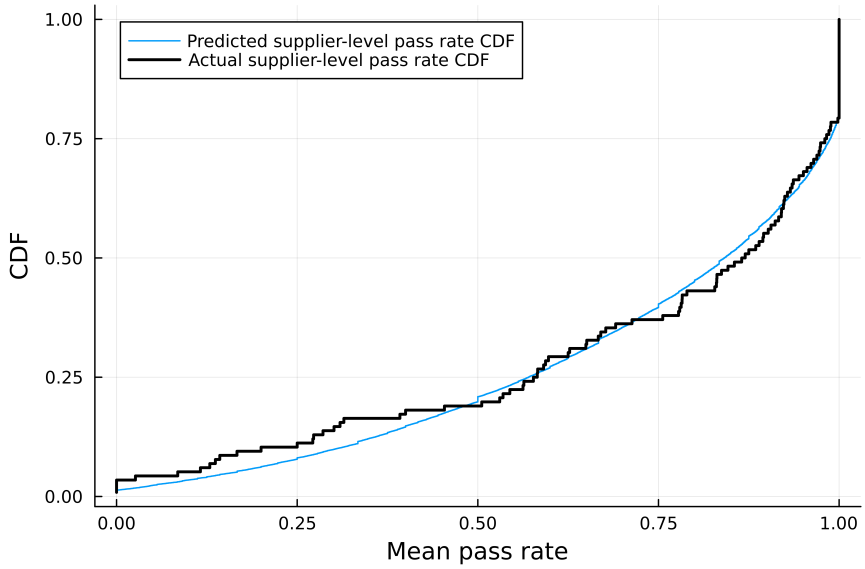
Notes: This graph plots the optimal effort $e^*(T, n, \tilde{X}; \theta, p^*, \tau^*)$ as a function of number of active suppliers n by fixing the tier of the supplier at Silver, i.e. $T =$ Silver, the cost shock at $\tilde{X} = 1$, and the type of the supplier at the mean of the type distribution, i.e. $\theta = 0.7$. Note that the optimal effort solves the dynamic programming problem in Equation 1, taking the estimated supplier beliefs as given.

Figure 15: Model fit of Stage 1 estimates.



Notes: The predicted distribution over tiers at each snapshot in this graph is constructed over many simulations of tier transitions using the Stage 1 estimates, taking the orders, assigned suppliers, and performance of suppliers as given.

Figure 16: Model fit of Stage 2 estimates.



Notes: This graph plots the CDF of supplier-level pass rates implied by the Stage 2 estimates along with the actual supplier-level pass rate CDF from the data.

A.6 Tables from Section 7

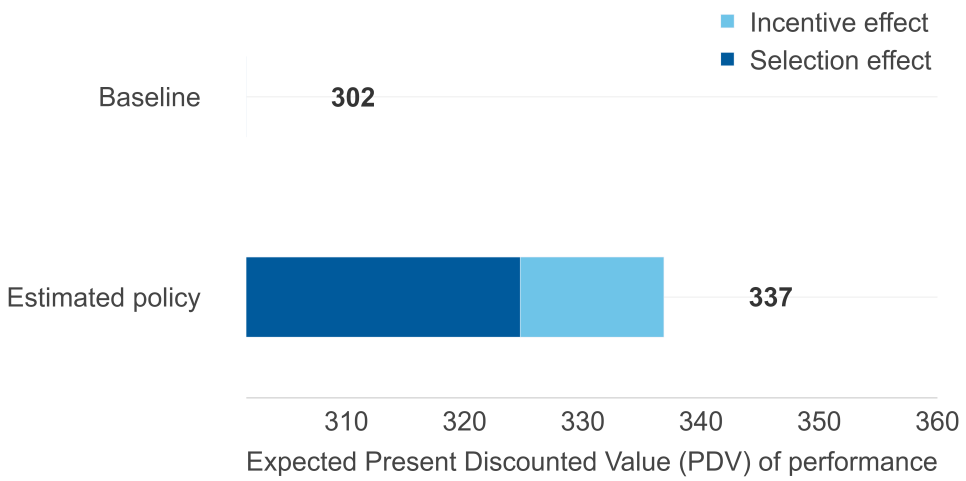
Table 10: Optimal policy parameters

Object	Parameter	Description	Value
τ	λ_u	Base upgrade weight	-15.42
	λ_s	Base retain weight	-12.47
	λ_d	Base downgrade weight	-12.57
	κ	Performance bonus	22.37
	μ_G	Base G weight (new supp)	-18.31
	μ_S	Base S weight (new supp)	-10.54
	μ_B	Base B weight (new supp)	-10.58
p	ϕ_G	Choice weight for G	7.24
	ϕ_S	Choice weight for S	-10.73
	ϕ_B	Choice weight for B	-10.74

Notes: This table presents the parameters of the buyer's optimal policy, found by numerically solving the optimization problem in Equation 13 using the estimated supplier-side parameters from Stage 2.

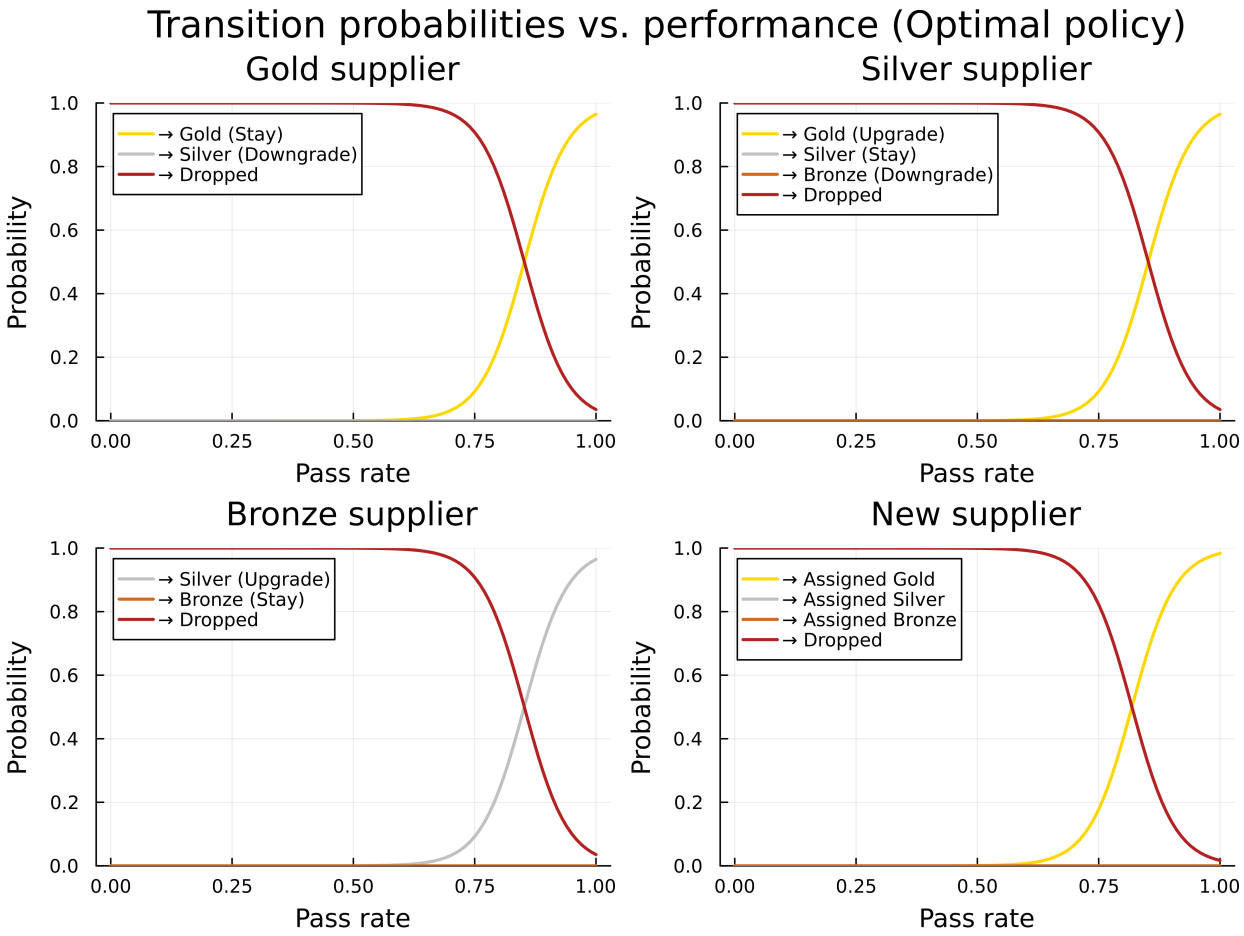
A.7 Figures from Section 7

Figure 17: The buyer's estimated policy improves expected PDV of performance by 12% over baseline, with around 66% of that improvement being attributable to selection.



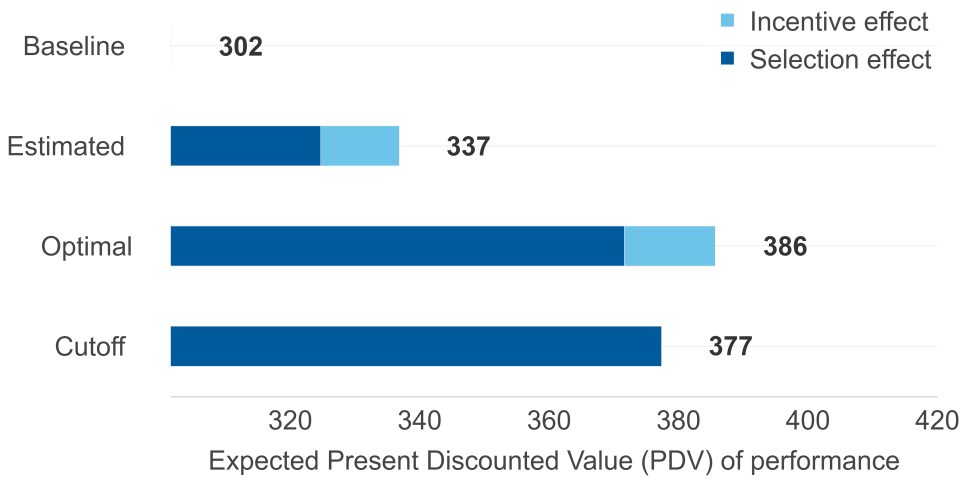
Notes: "Selection effect" is the value gained by allocating orders based on the estimated policy, assuming suppliers exert zero effort. "Incentivization effect" is the additional value gained from suppliers exerting their optimal effort in response to the policy.

Figure 18: The optimal policy has a more striking change in transition profile as performance changes.



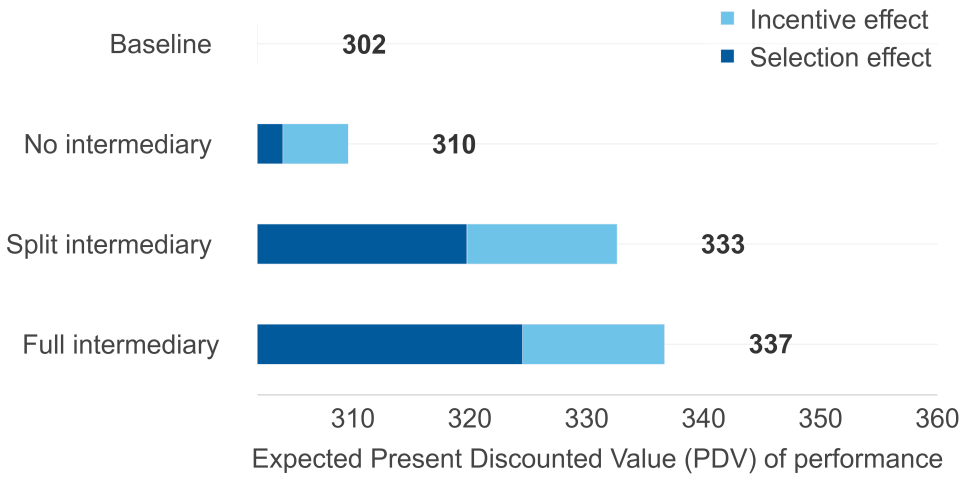
Notes: This figure shows the transition probabilities as a function of performance under the calculated optimal policy. Compared to the estimated policy (Figure 10), the optimal policy features a much stronger link between performance and outcomes, with a sharp increase in the probability of being upgraded (or avoiding being dropped) as pass rates approach 1.0.

Figure 19: The optimal policy significantly improves outcomes relative to the estimated policy.



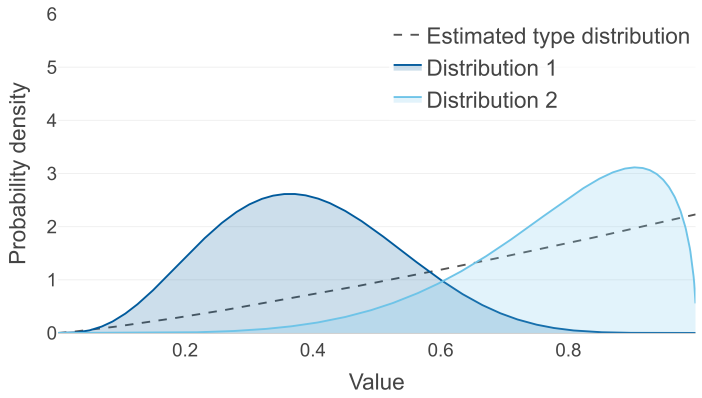
Notes: The optimal is calculated over the specific parametrization of the buyer's policy in Subsection 5.2.

Figure 20: Splitting the buyer into two marginally reduces value, while having only one-on-one relationships between the downstream customer and the suppliers leads to a significant drop in value.



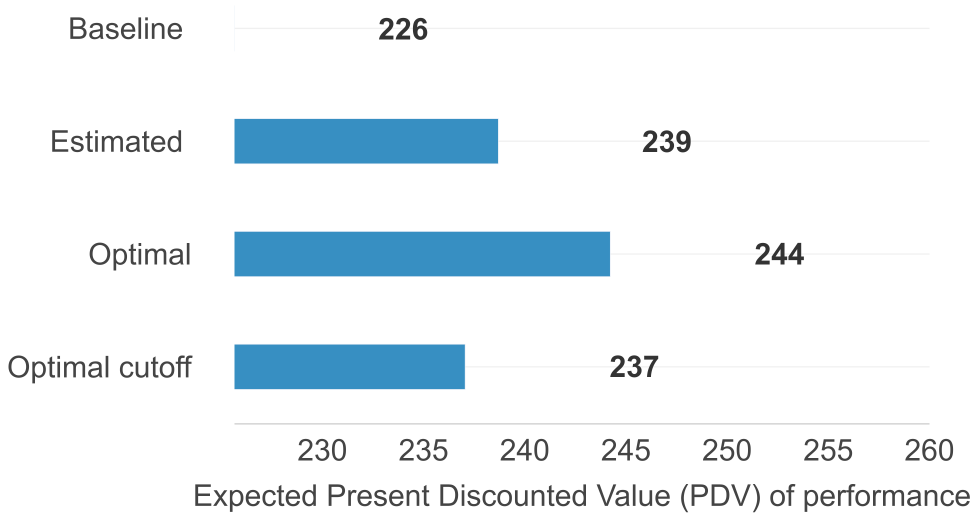
Notes: In all cases, the estimated policy is implemented.

Figure 21: Counterfactual distributions of types.



Notes: Under the assumption that types are drawn from Distribution 1 in the first third of the sample, and then from Distribution 2, the mean and variance of the implied distribution of types over the sample matches the mean and variance of the estimated distribution of types.

Figure 22: The optimal policy's improvement relative to the estimated policy shrinks significantly under the counterfactual staggered draw of types.



Notes: In this counterfactual scenario, types are drawn from Distribution 1 in Figure 21 in the first third of the sample, and then from Distribution 2.

Figure 23: Counterfactual distribution of types.

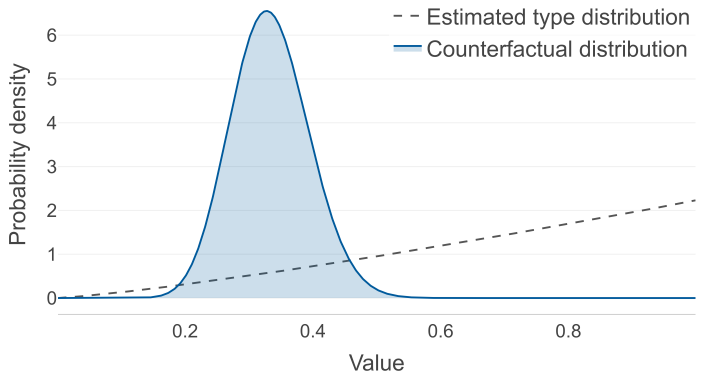
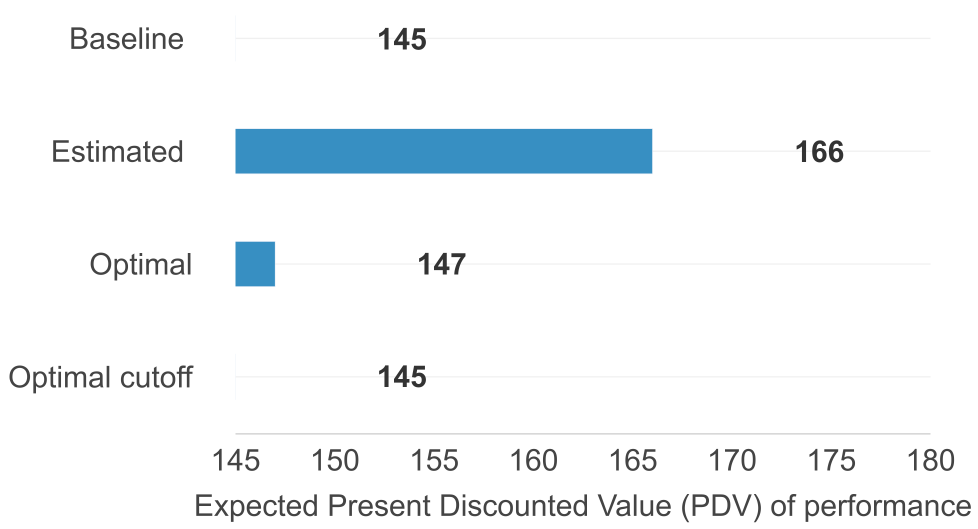


Figure 24: The optimal policy performs worse than the estimated policy if types were drawn from a significantly worse distribution.



Notes: In this counterfactual scenario, types are drawn from the distribution in Figure 23.

B Approximating suppliers' optimal effort

Repeatedly solving the suppliers' dynamic problem (Equation 1) for every guess of the structural parameters within the maximum likelihood routine can impose significant computational cost. Thus, I use a neural network to serve as a fast and flexible function approximator for the supplier's optimal effort policy, e^*. This approach, which leverages the universal approximation property of neural networks (Hornik, Stinchcombe, and White, 1989), is part of a growing literature applying machine learning methods to the estimation of economic models (Athey and Imbens, 2019). The procedure involves three key steps: generating training data, training the network, and deploying it within the estimation.

First, I generate a training dataset by solving the supplier's dynamic programming problem for the true optimal effort e^* across a wide range of scenarios. This is done by constructing a discrete grid of size 10 along each dimension over both the supplier's state variables (tier T, total number of active suppliers n, price state X, type θ) and the structural parameters (γ, α, ρ). At each point on this high-dimensional grid, the model is solved using value function iteration until convergence. The resulting dataset consists of input vectors (containing the state and parameter values) and their corresponding optimal effort e^* as the target output.

Second, I use this dataset to train a feed-forward neural network to learn the mapping from inputs to optimal effort. The network architecture consists of two hidden layers with 512 and 256 neurons respectively, using the GELU (Gaussian Error Linear Unit) activation function. To improve performance, standard machine learning practices are followed:

- Feature engineering: Categorical state variables, such as the supplier's tier T, are one-hot encoded. This process converts a categorical feature into a set of binary (0 or 1) variables, with one variable for each unique category. For instance, for the tier variable, a Gold supplier would be represented by a vector like [1,0,0] and a Silver supplier by [0,1,0]. This method allows machine learning models to process categorical data without incorrectly assuming a numerical or ordinal relationship between the categories.
- Data handling: The data is split randomly into 80% for training and 20% for validation. All input features and the target effort values are standardized to have zero mean and unit variance to improve training stability.
- Training: The network's weights are optimized using the standard AdamW algorithm to minimize the mean squared error (MSE) between the network's prediction, $\hat{e}$, and the true optimal effort, e^*. Training is stopped early if the performance on the validation set ceases to improve, ensuring the model generalizes well to unseen data.

The training stops at a validation loss of 0.003; recall that the target effort values are standardized to have unit variance, so this represents a 99.7% improvement in MSE over a naive predict-mean approach.

Finally, during the main Stage 2 estimation, the trained neural network acts as a surrogate for the DP solver. Whenever the maximum likelihood routine requires the optimal effort for a given set of parameters, it makes a near-instantaneous call to the neural network approximator, $\hat{e}$, instead of running the full value function iteration.

C Formal identification results

C.1 Stage 1: The buyer's policy

The buyer's policy consists of (i) an allocation rule p that maps the current tier vector to a distribution over suppliers, and (ii) a transition rule τ that updates next period's tier vector as a function of the current tier vector and the realized, observable outcome $z = (a, y)$. It is shown to be identified nonparametrically using a grouping of observations strategy from Allman, Matias, and Rhodes (2009) to satisfy rank conditions in large-t data. The idea involves treating the stream of orders as blocks so that the implied observed distributions are sufficiently informative, and then exploiting the fact that conditioning on the realized present block renders the past and the future independent.

Two assumptions are needed—one about what the data must contain, and one about generic regularity of the policy.

Assumption 3 (Data requirements)**.** *(i)* Snapshot. *The data include at least one snapshot period in which the tier vector T is observed.*

(ii) Three groups of orders. *There exist integers $k \geq 1$ and $r \geq 1$ such that, with positive probability, there are observed $2k + r$ consecutive orders that can be partitioned into three non-overlapping groups:*

$$Y_1 = (z_{t-k}, \ldots, z_{t-1}), \qquad Y_2 = (z_t, \ldots, z_{t+r-1}), \qquad Y_3 = (z_{t+r}, \ldots, z_{t+r+k-1}),$$

where each $z_s = (a_s, y_s)$ is the observed supplier choice and outcome. Let $\mathscr{Z}$ denote the support of z_s. The group length k is chosen so that the set of length-k sequences is large enough, generically,

to distinguish the K possible tier vectors:

$$|\mathscr{Z}|^k \geq K,$$

which ensures full-column-rank "grouped emission" matrices.

Assumption 4 (Policy regularity)**.** *(i)* Generic emissions. *The family of state-specific grouped emission distributions for Y_1 and for Y_3 is linearly independent across states.*
(ii) Full-rank outcome-sliced transitions. *For every realized present group value $\bar{z} \in \text{supp}(Y_2)$, the matrix*

$$\mathbf{A}^{(\bar{z})} := \left(\tau(T' \mid T, \bar{z})\right)_{T',T}$$

has full column rank.
(iii) Positivity. *All probabilities in the policy are strictly between 0 and 1 on their support.*

Assumption 3 is a "large-t" requirement: grouping makes the observations big enough to separate states without increasing the number of latent states. Assumption 4 is a genericity condition on the policy: grouped supplier choices are rich enough to pass linear-independence checks, and for each realized present block $\bar{z}$ the associated transition operator across tiers is of full column rank.

Theorem 1 (Joint identification of the policy (p, τ))**.** *Under Assumptions 3 and 4, the allocation rule p and the transition rule τ are identified.*

Proof. The proof proceeds in five steps.

Step 1: Fix a present block. Let the present state be the tier vector $\mathbf{T_t}$. Given $\mathbf{T_t}$, an observable sequence Y_2 of length r is generated where r is the number of orders in the present month; write its realized value as $\bar{z} \in \text{supp}(Y_2)$. The next latent state T_{t+r} is drawn using the outcome-contingent transition rule $\tau(\cdot \mid \mathbf{T_t}, \bar{z})$.

Step 2: Grouping to create "views". Form Y_1 (past group) and Y_3 (future group) of length k as in Assumption 3. Because the process is first-order Markov at the level of T, conditioning on the observed state $\mathbf{T_t}$ at the start of Y_2 and on the realized present block $\bar{z}$ renders Y_1 and Y_3 conditionally independent:

$$Y_1 \perp\!\!\!\perp Y_3 \,\Big|\, \left(\mathbf{T_t},\ Y_2 = \bar{z}\right).$$

Step 3: Three-view mixture conditional on $\mathbf{T_t}$ and $Y_2 = \bar{z}$. Consider the conditional joint law $\mathbb{P}(Y_1, Y_3 | Y_2 = \bar{z})$. The identification of the policy from this law follows from a result

for Input-Output Hidden Markov Models (IO-HMMs) from David et al. (2024). My model fits within their framework as a special case where the latent state is the tier vector $\mathbf{T_t}$, the output is the supplier choice a_t, and the input signal is the lagged performance outcome y_{t-1}. My Assumptions 3 and 4 ensure that the conditions for their main identification result (Theorem 1) are met. Their first assumption, H1, which defines the model's conditional independence structure, is satisfied by my policy specification. Their second assumption, H2, a technical condition ensuring that conditional laws are well-defined, is met as long as τ is a continuous function of y, which it is in the model's parametric construction. Their third assumption, H3, imposes rank conditions for identification. This is satisfied by my policy regularity requirements: the linear independence of grouped emissions is guaranteed by my Assumption 4(i), which is made possible by the large observation space created by grouping (Assumption 3(ii)), while the invertibility of transition matrices is guaranteed by my Assumption 4(ii). Finally, their fourth assumption, H4 on positivity, is guaranteed by my Assumption 4(iii). Given that these conditions are met, David et al. (2024) show that this conditional law admits a unique mixture-of-products representation (up to a common permutation of the latent states). Concretely, the following are recovered:

- the state-specific grouped emission distributions for Y_1, say $\{E^{-}_{T}(\cdot)\}_{T\in\mathscr{T}}$;
- the state-specific grouped emission distributions for Y_3 *after* the present block equals $\bar{z}$, which take the form

$$E^{+}_{T\to\cdot}(\cdot \mid \bar{z}) \;=\; \sum_{T'\in\mathscr{T}} \tau(T' \mid T,\bar{z})\, \widetilde{E}^{+}_{T'}(\cdot),$$

  where $\widetilde{E}^{+}_{T'}(\cdot)$ is the grouped emission distribution for Y_3 starting from state T'.

Stacking these objects across states produces (up to a permutation) two emission matrices and, for each realized $\bar{z}$, a linear function of the transition operator $\tau(\cdot \mid \cdot,\bar{z})$.

Step 4: Recovering the policy. Assumption 4(i) implies that the grouped emission matrix for Y_3, $\widetilde{\mathbf{E}}^{+} = [\widetilde{E}^{+}_{T'}(\cdot)]_{T'}$, has full column rank generically. Assumption 4(ii) then guarantees that, for each $\bar{z}$ in the support of Y_2, the map

$$\mathbf{A}^{(\bar{z})} \;:\; \Delta(\mathscr{T}) \longrightarrow \mathrm{Law}(Y_3 \mid \mathbf{T_t},\; Y_2 = \bar{z}) \quad \text{with} \quad \mathbf{A}^{(\bar{z})} \;=\; \big(\tau(T' \mid T,\bar{z})\big)_{T',T}\, \widetilde{\mathbf{E}}^{+\top}$$

is injective, so $\mathbf{A}^{(\bar{z})}$ and hence $\tau(\cdot \mid \cdot,\bar{z})$ are uniquely recovered (up to the same permutation of state labels). Varying $\bar{z}$ over its support identifies the full outcome-contingent transition rule $\tau(\cdot \mid T,z)$. The order allocation rule $p(a \mid \mathbf{T_t})$ is obtained directly from within-block orders by marginalizing to the appropriate events (e.g., the event that the first element in Y_2 equals (a,y)). Thus p is identified up to the same permutation of latent labels.

Step 5: Labeling the states using the snapshot. The previous steps identify (p, τ) up to relabeling. The snapshot period maps each statistical state to its realized tier vector, fixing labels in economic terms.

□

C.2 Stage 2: The supplier side

The identification of the supplier side leverages the structure of the model and the data to identify the parameters sequentially. First, I establish a lemma that identifies the lot-level outcome distribution, which then allows me to simplify the identification of the remaining objects.

Lemma 1 (Identification of lot-level outcome distribution)**.** *Assume that the data contain a large number of orders such that each has outcomes observed for at least two distinct lots that face the same cost shock and have a large number of parts. Then the mean and variance of the outcome distribution are identified.*

Proof. Let k_{il} be the number of successful parts out of N in lot l of order i. Let $p_i \in [0,1]$ be the unobserved, order-specific pass probability for any given part. By de Finetti's representation theorem for exchangeable sequences, the lot outcomes behave as if they are independent and identically distributed draws from a Binomial distribution, conditional on the latent order-specific parameter p_i, which is itself drawn from a mixing distribution $\tilde{F}(p)$ (Schervish, 2012). Let $\mu = \mathbb{E}[p_i]$ and $\sigma^2 = \text{Var}(p_i)$ be the first two moments of the mixing distribution $F\tilde{F}$. The mean of the latent distribution, μ, is identified from the expected number of successes in any single lot:

$$\mathbb{E}[k_{il}] = \mathbb{E}_i[\mathbb{E}[k_{il}|p_i]] = \mathbb{E}_i[Np_i] = N\mu \implies \mu = \frac{\mathbb{E}[k_{il}]}{N}.$$

The variance of the latent distribution, σ^2, is identified from the covariance of successes between two distinct lots, l and l', from the same order:

$$\text{Cov}(k_{il}, k_{il'}) = \mathbb{E}_i[\text{Cov}(k_{il}, k_{il'}|p_i)] + \text{Cov}_i(\mathbb{E}[k_{il}|p_i], \mathbb{E}[k_{il'}|p_i]).$$

Since lots are independent conditional on p_i, the first term is zero, so

$$\text{Cov}(k_{il}, k_{il'}) = \text{Cov}_i(Np_i, Np_i) = N^2\text{Var}_i(p_i) = N^2\sigma^2 \implies \sigma^2 = \frac{\text{Cov}(k_{il}, k_{il'})}{N^2}.$$

□

Simplifications and notation for latent tiers. Since the lot-level outcome distribution is identified, I work with mean outcomes μ_{it} for the rest of the results. Let the supplier's policy state be $S_{it} = (T_{it}, n_t)$. Since the econometrician does not observe the supplier's tier in between snapshot months, let the *econometrician's policy state* be

$$\widehat{S}'_{it} \equiv (I_{it}, n_t),$$

where n_t is observed and I_{it} indexes the information set used to filter the tier, inducing the posterior $\pi(T; I_{it}) \equiv \Pr(T_{it} = T \mid I_{it})$. Here, I_{it} includes the observed supplier choices, the observed outcomes, and the snapshot tiers, so the posterior is formed using the distribution over latent tiers in Equation 8 that is implied by the estimated policy conditional on the supplier choices, outcomes, and snapshot tiers.

Given the marginal continuation value $\mathscr{W}_y(y, S_{it}, \theta)$, define the ex-ante average

$$\overline{\mathscr{W}}(S_{it}, \theta) \equiv \mathbb{E}\big[\mathscr{W}_y(Y_{t+1}, S_{it}, \theta) \mid S_{it}, \theta\big],$$

where Y_{t+1} is the next-period outcome induced by current (S_{it}, θ) and the supplier's optimal e_{it}. The posterior π is used to specify two objects that will be used in the proofs:

$$\overline{\mathscr{W}}^{\text{obs}}(I, \theta) \equiv \mathbb{E}\big[\overline{\mathscr{W}}(S', \theta) \mid I\big],$$
$$\Lambda(I, \theta) \equiv \mathbb{E}\big[\ln \widetilde{\mathscr{W}}(S', \theta) \mid I\big],$$

i.e. a belief average of the level of the ex-ante marginal continuation value and a belief average of its log in the normalized representation.

The following assumptions provide the necessary variation and structure for the rest of the proof.

Assumption 5 (Sufficient within-supplier *belief*-state variation)**.** *Denote the observable cost shock $\tilde{X}_t = X_t / X_{t-1}$. There is a large set of suppliers such that each supplier i is observed for at least three periods where his* econometrician's policy state *is identical, i.e.*

$$\widehat{S}'_{it} = (I_{it}, n_t) = \widehat{S}'_{is} = (I_{is}, n_s) = \widehat{S}'_{ir} = (I_{ir}, n_r)$$

for some t, s, r, but the contemporaneous cost shocks differ, i.e. $\tilde{X}_t \neq \tilde{X}_s \neq \tilde{X}_r$. The types of these suppliers form a positive-measure set in the support of F.

Assumption 6. *The buyer's chosen policy is such that the following restrictions are satisfied:*

1. (Sufficient dynamics) *The tier transition policy τ induces an irreducible Markov chain over the set of active tiers $\mathcal{T}_{active} = \{G, S, B\}$.*

2. (Heterogeneous marginal incentives, observed-state version) *For any two* realized *supplier policy states $S'_1 \neq S'_2$,*

$$\theta \mapsto \frac{\overline{\mathscr{W}}(S'_1, \theta)}{\overline{\mathscr{W}}(S'_2, \theta)} \quad \text{is not constant on its domain, and } \overline{\mathscr{W}}(S', \theta) > 0.$$

3. (Heterogeneous marginal incentives, belief-state version) *There exist two econometrician policy states $I \neq I'$ such that*

$$\theta \mapsto \frac{\overline{\mathscr{W}}^{obs}(I, \theta)}{\overline{\mathscr{W}}^{obs}(I', \theta)} \quad \text{is not constant on its domain, and } \overline{\mathscr{W}}^{obs}(I, \theta) > 0.$$

Assumption 7 (Sufficient variation in cost shocks)**.** 1. *The cost shock $\tilde{X}_t$ varies over a continuous support for each state, and is independent of T_{it}.*

2. *There is a large enough cost shock $\tilde{X}_t = x$ such that the map $\theta \mapsto \mu(\theta, e^*(\theta; x, NEW))$ is strictly increasing in θ over the support of the type distribution F.*

Remark 1 (Snapshot months for cost-side steps)**.** *In months where tiers are directly observed ("snapshots"), the posterior π over tiers is degenerate and coincides with the realized tier. The cost-side arguments in Step 2.2 below are executed on snapshot months; all other steps use the belief-state formulation.*

Theorem 2. *Under Assumptions 5, 6, 7, the supplier cost function c, the outcome production function G, and the distribution of supplier types F are identified.*

Proof. The proof is sequential and constructive, proceeding in three steps.

Step 1: Identification of anchor types $\{\theta_i\}$ and the performance elasticity α. By Assumption 5, for many suppliers i there exist *three* periods t, s, r with the same belief state $\widehat{S}'_{it} = \widehat{S}'_{is} = \widehat{S}'_{ir}$ but different cost shocks $\tilde{X}_t, \tilde{X}_s, \tilde{X}_r > 0$. Combing the outcome equation with the supplier's FOC in any such period $k \in \{t, s, r\}$ implies

$$\mu_{ik} - \theta_i \;=\; C_i(\widehat{S}'_i, \theta_i)\,(\tilde{X}_k)^{\beta}, \qquad \beta \equiv \frac{\alpha}{\alpha - 1},$$

where $C_i(\widehat{S}'_i, \theta_i) > 0$ depends on $(\theta_i, \widehat{S}'_i)$ but is constant across k because $\widehat{S}'$ is fixed. (Formally, C_i equals the belief-average of the state-specific constant obtained by eliminating e^*_{ik} from the

FOC and outcome equation; separability in $(\tilde{X})^\beta$ survives taking the conditional expectation given $\widehat{S}'$.) Taking two differences removes θ_i and C_i:

$$\frac{\mu_{it}-\mu_{is}}{\mu_{ir}-\mu_{is}} \;=\; \frac{(\tilde{X}_t)^\beta-(\tilde{X}_s)^\beta}{(\tilde{X}_r)^\beta-(\tilde{X}_s)^\beta} \;\equiv\; \phi_{tsr}(\beta),$$

and the standard "between-endpoints" argument shows ϕ_{tsr} is strictly monotone, pinning down a unique β. Recover C_i and then θ_i from any two periods:

$$C_i \;=\; \frac{\mu_{it}-\mu_{is}}{(\tilde{X}_t)^\beta-(\tilde{X}_s)^\beta}, \qquad \theta_i \;=\; \mu_{it}-C_i(\tilde{X}_t)^\beta.$$

Step 2: Identification of outcome and cost parameters (ρ,γ). Take as given the anchor types $\{\theta_i\}$ and the elasticity α, and assume that $\rho \neq 0$.

Step 2.1: Construction of a structural equation and data-derived objects. The supplier marginal value function is homogeneous of degree 1 in $(\bar{b},\gamma)$, hence its average marginal value satisfies

$$\overline{\mathscr{W}}(S',\theta;\bar{b},\gamma)=\gamma\,\widetilde{\mathscr{W}}\Big(S',\theta;\tfrac{\bar{b}}{\gamma}\Big).$$

With $\bar{b}=1$, the FOC and outcome mapping imply, for $K(\theta)\equiv\sigma(\rho\theta)$ with $\sigma(x)=\frac{e^x}{1+e^x}$,

$$(\tilde{X}_t)=\widetilde{\mathscr{W}}(S'_{it},\theta_i)\cdot\alpha\,(e^*_{it})^{\alpha-1}K(\theta_i), \qquad \mu_{it}-\theta_i=(e^*_{it})^\alpha K(\theta_i).$$

Eliminating e^*_{it} and rearranging yields the central structural equation that holds pointwise in the realized S'_{it}:

$$y_{it} \;\equiv\; \beta\ln(\tilde{X}_t)-\ln(\mu_{it}-\theta_i) \;=\; \frac{\alpha}{\alpha-1}\Big[\ln\widetilde{\mathscr{W}}(S'_{it},\theta_i)+\ln\alpha\Big]+\frac{1}{\alpha-1}\ln K(\theta_i;\rho). \tag{14}$$

Taking conditional expectations given the belief state I_{it} produces the following *structural equation*

$$\mathbb{E}[y_{it}\mid I_{it},\theta_i]=\frac{\alpha}{\alpha-1}\Big[\Lambda(I_{it},\theta_i)+\ln\alpha\Big]+\frac{1}{\alpha-1}\ln K(\theta_i;\rho), \tag{15}$$

where the expectation is over tiers T (and next-period outcomes in the definition of $\widetilde{\mathscr{W}}$). I show that this structural equation can be rewritten using data-derived objects as:

$$\mathbb{E}[y\mid I,\theta] \;=\; A_\alpha\big\{\Lambda(I,\theta)+\ln\alpha\big\}+B_\alpha\ln K(\theta) \;=\; A_\alpha\big\{\ln u_\theta+\widehat{\Lambda}(I,\theta)+\ln\alpha\big\}+B_\alpha\ln K(\theta). \tag{16}$$

This is done as follows. Fix an anchor type θ. Let T be the current tier and $\tau(T'\mid T,z)$

the next–tier kernel with $z = (a, y)$. Define the y–sensitivity of next–tier probabilities

$$B_\tau \equiv \left(\tfrac{\partial}{\partial y}\,\tau(T' \mid T, z)\right)_{T'} \in \mathbb{R}^{|\mathscr{T}|}.$$

Let P_θ be the tier–tier transition matrix under the identified policy for type θ, and let $q \in \mathbb{R}^{|\mathscr{T}|}$ be the per–tier selection probabilities from the allocation rule p, both recovered nonparametrically from the data using snapshot transitions and selections for the known anchor types. Define the discounted order–occupancy vector

$$m_\theta \equiv (I - \delta P_\theta)^{-1} q = \sum_{h \geq 0} \delta^h P_\theta^h q.$$

Write the continuation value by tier as

$$V_\theta = (I - \delta P_\theta)^{-1} r_\theta, \qquad r_\theta = s_\theta\, q + \bar{r}_\theta\, \mathbf{1}.$$

Because $\mathbf{1}^\top B_\tau = 0$, the tier–invariant part $\bar{r}_\theta\, \mathbf{1}$ drops out of $B_\tau^\top V_\theta$, so

$$V_\theta = s_\theta\, m_\theta + \frac{\bar{r}_\theta}{1-\delta}\,\mathbf{1}, \qquad W_y(y, S', \theta) = \delta\, B_\tau^\top V_\theta.$$

Using the homogeneity normalization $W = \gamma\, f_W(\cdot; 1/\gamma)$,

$$f_W(S', \theta) = \frac{\delta}{\gamma} B_\tau^\top V_\theta = \underbrace{\left(\tfrac{s_\theta}{\gamma}\right)}_{u_\theta}\, \underbrace{\delta\, B_\tau^\top m_\theta}_{\widetilde{f}_W(S', \theta)}.$$

Given (P_θ, q) and B_τ, construct the following:

(C1) compute $m_\theta = (I - \delta P_\theta)^{-1} q$;

(C2) compute $\widetilde{f}_W(S', \theta) = \delta\, B_\tau^\top m_\theta$;

(C3) set $\widehat{\Lambda}(I, \theta) = \mathbb{E}[\ln \widetilde{f}_W(S', \theta) \mid I]$.

Then

$$\Lambda(I, \theta) \equiv \mathbb{E}[\ln f_W(S', \theta) \mid I] = \ln u_\theta + \widehat{\Lambda}(I, \theta), \tag{17}$$

with $u_\theta \equiv s_\theta / \gamma$ depending only on θ.

Finally, I show that u_θ can be recovered from the primitives. Define

$$A_\alpha \equiv \frac{\alpha}{\alpha - 1}, \qquad B_\alpha \equiv \frac{1}{\alpha - 1}, \qquad \beta \equiv A_\alpha.$$

On the interior $e^*(S',\theta,\tilde{X}) > 0$, the FOC and the cost function are

$$\tilde{X} \;=\; \alpha\, f_W(S',\theta)\, e^{\alpha-1} K(\theta), \qquad\qquad c(e) \;=\; \gamma\, \tilde{X}\, e.$$

Solving the FOC for e^* and substituting into $c(\cdot)$ yields

$$e^*(S',\theta,\tilde{X}) = \Big[\tfrac{\tilde{X}}{\alpha\, f_W(S',\theta) K(\theta)}\Big]^{B_\alpha},$$
$$c(e^*) = \gamma\, \alpha^{-B_\alpha}\, \tilde{X}^{\beta}\, f_W(S',\theta)^{-B_\alpha}\, K(\theta)^{-B_\alpha}.$$

Taking selection–conditional expectations,

$$s_\theta = \mathbb{E}[1 - c(e^*) \mid \text{selected}, \theta\,] = 1 - \gamma\, \alpha^{-B_\alpha}\, K(\theta)^{-B_\alpha}\, \mathbb{E}\Big[\tilde{X}^{\beta}\, f_W(S',\theta)^{-B_\alpha} \mid \text{selected}, \theta\,\Big].$$

Using $f_W = u_\theta\, \widetilde{f}_W$, define the data moment

$$M_\theta \;\equiv\; \mathbb{E}\Big[\tilde{X}^{\beta}\, \widetilde{f}_W(S',\theta)^{-B_\alpha} \mid \text{selected}, \theta\,\Big],$$

and obtain the scalar fixed point for $u_\theta \in (0, 1/\gamma)$:

$$\frac{1}{\gamma} - u_\theta \;=\; \alpha^{-B_\alpha}\, K(\theta)^{-B_\alpha}\, M_\theta\, u_\theta^{-B_\alpha}\,. \tag{18}$$

Step 2.2: Identification of (ρ,γ). Define the I–invariant composite

$$J_\theta \;\equiv\; \mathbb{E}[y \mid I,\theta\,] - A_\alpha\big\{\widehat{\Lambda}(I,\theta) + \ln\alpha\big\} \;=\; A_\alpha \ln u_\theta \;+\; B_\alpha \ln K(\theta),$$

where u_θ solves Equation 18. By construction, u_θ depends on (ρ,γ) through $K(\theta) = \sigma(\rho\theta)$ with $\sigma(x) = \exp(x)/(1+\exp(x))$. Hence, across anchors $\{\theta\}$,

$$J_\theta \;=\; A_\alpha \ln u_\theta(\rho,\gamma) \;+\; B_\alpha \ln \sigma(\rho\theta). \tag{19}$$

For each fixed (ρ,γ), the fixed-point equation Equation 18 admits a unique solution $u_\theta(\rho,\gamma) \in (0, 1/\gamma)$. Moreover, $u_\theta(\rho,\gamma)$ is strictly decreasing in γ for given ρ. Thus, the right-hand side of Equation 19 is strictly monotone in γ, while also strictly monotone in ρ through the logistic link $\sigma(\rho\theta)$ when $\rho \neq 0$.

Because the data deliver J_θ for each anchor θ, and given at least three distinct anchor values, the nonlinear system (19) admits at most one solution (ρ,γ). Hence (ρ,γ) are jointly identified.

Step 3: Identification of the type distribution F. Let $a \equiv (\tilde{X})^{\beta}$ and write, for any realized supplier policy state S',

$$m_{a,S'}(\theta) \;=\; \theta + a\,h(\theta, S'), \qquad h(\theta, S') \equiv \left(\frac{\alpha\, \overline{\mathscr{W}}(S', \theta)}{\gamma} \right)^{\frac{\alpha}{1-\alpha}} K(\theta)^{\frac{1}{1-\alpha}}.$$

The function $m_{a,\text{NEW}}(\cdot)$ is known for any realized a since all the supplier-side objects have been identified.

Fix first orders ($S' = \text{NEW}$) and condition on any realized a_0 with $\Pr\{a = a_0\} > 0$. Set $Z := m_{a_0,\text{NEW}}(\theta) \in [0,1]$. Let K denote the number of passes in a lot of size N. Conditional on Z and N,

$$K \,\big|\, Z, N \;\sim\; \text{BetaBinomial}\big(N, \phi Z, \phi(1-Z)\big),$$

with precision ϕ known from earlier steps.[27]

Now take factorial moments. Using the identity $\mathbb{E}[(K)_{(r)} \mid p, N] = (N)_{(r)} p^r$ and integrating over the mixture,

$$\mathbb{E}[(K)_{(r)} \mid N] = (N)_{(r)}\, \mu_r, \qquad \mu_r := \mathbb{E}[p^r].$$

Thus, for each $r \le N$ we obtain $\mu_r = \mathbb{E}[(K)_{(r)} \mid N]/(N)_{(r)}$. As N varies in the population, this delivers the full sequence $\{\mu_r\}_{r \ge 1}$ of moments of the mixing law on $[0,1]$; by standard results for the Hausdorff moment problem on a compact interval—namely, that a probability measure on $[0,1]$ is uniquely determined by its moments—these moments uniquely determine that law (see, e.g., Billingsley (1995)).[28] Finally, because $p \mid Z \sim \text{Beta}(\phi Z, \phi(1-Z))$ with known ϕ, the map $Z \mapsto \{\mathbb{E}[p^r \mid Z]\}_{r \ge 1}$ is polynomial and lower–triangular in $\{Z^j\}$ with nonzero diagonal, so the recovered $\{\mu_r\}$ recursively identify all moments of Z; since $Z \in [0,1]$, these moments pin down the distribution of Z. Denote this CDF by $G_{a_0,\text{NEW}}$.

Because $m_{a_0,\text{NEW}}(\cdot)$ is known and strictly increasing, it admits a well-defined inverse $m^{-1}_{a_0,\text{NEW}}$. Then the change-of-variables identity (see, e.g., the monotone transformation argument in Matzkin (2003) and Matzkin (2007)) yields, for every t in the support of θ,

$$F(t) \;=\; \Pr\{\theta \le t\} \;=\; \Pr\{Z \le m_{a_0,\text{NEW}}(t)\} \;=\; G_{a_0,\text{NEW}}\big(m_{a_0,\text{NEW}}(t)\big). \tag{20}$$

[27] Classical linear deconvolution results with known additive error (e.g., Carroll and Hall (1988), Fan (1991)) do not apply here because the lot-level uncertainty is generated by a Beta–Binomial mixture rather than by additive, θ-independent noise.

[28] If the support of lot sizes is bounded by $\bar{N} < \infty$, the procedure recovers $\{\mu_r\}_{r \le \bar{N}}$ only. This yields nonparametric partial identification; under a finite-dimensional parametric restriction on F, one obtains point identification of the parameters.

Thus F is identified.

□

D Calculation of standard errors

D.1 Stage 1: Contract parameters

The parameters of the buyer's policy in the first stage, denoted collectively by ϑ_1 in this section, are estimated using a Monte Carlo Expectation-Maximization (MCEM) algorithm. Standard errors for these estimates $\hat{\vartheta}_1$ are computed using the formula derived in Louis (1982), which relates the observed-data information matrix to quantities derived from the complete-data likelihood.

Let $L_c(Y, Z; \vartheta_1)$ be the complete-data log-likelihood, where Y represents the observed data (order assignments, outcomes, snapshot tiers) and Z represents the latent data (the unobserved paths of supplier tiers between snapshots). The observed-data log-likelihood is $L(Y; \vartheta_1) = \log \int \exp(L_c(Y, Z; \vartheta_1)) dZ$.

Louis' formula states that the observed-data information matrix, $I_{obs}(\vartheta_1)$, can be expressed as:

$$I_{obs}(\vartheta_1) = -E_{Z|Y;\vartheta_1}\left[\nabla^2_{\vartheta_1\vartheta_1} \log L_c(Y, Z; \vartheta_1)\right] - \mathrm{Var}_{Z|Y;\vartheta_1}\left[\nabla_{\vartheta_1} \log L_c(Y, Z; \vartheta_1)\right] \tag{21}$$

where ∇_{ϑ_1} and $\nabla^2_{\vartheta_1\vartheta_1}$ denote the gradient and Hessian matrix with respect to ϑ_1, respectively. The expectation $E_{Z|Y;\vartheta_1}[\cdot]$ and variance $\mathrm{Var}_{Z|Y;\vartheta_1}[\cdot]$ are taken with respect to the conditional distribution of the latent data Z given the observed data Y and parameters ϑ_1.

Monte Carlo approximation The terms in Equation 21 are approximated using Monte Carlo simulation. Given the MCEM parameter estimates $\hat{\vartheta}_1$:

1. Simulate R independent complete-data paths $Z^{(r)}$ for $r = 1, \ldots, R$, from the conditional distribution $p(Z|Y; \hat{\vartheta}_1)$. Each $Z^{(r)}$ consists of a set of tier trajectories.

2. For each path $Z^{(r)}$, compute the complete-data score vector $S^{(r)}(\hat{\vartheta}_1) = \nabla_{\vartheta_1} \log L_c(Y, Z^{(r)}; \hat{\vartheta}_1)$ and the complete-data Hessian matrix $H^{(r)}(\hat{\vartheta}_1) = \nabla^2_{\vartheta_1\vartheta_1} \log L_c(Y, Z^{(r)}; \hat{\vartheta}_1)$. These derivatives are obtained using automatic differentiation.

The components of Louis' formula are then estimated as:

- Expected complete-data Hessian:

$$\widehat{EH}(\hat{\vartheta}_1) = \frac{1}{R}\sum_{r=1}^{R} H^{(r)}(\hat{\vartheta}_1) \tag{22}$$

- Variance of the complete-data score: Let $\bar{S}(\hat{\vartheta}_1) = \frac{1}{R}\sum_{r=1}^{R} S^{(r)}(\hat{\vartheta}_1)$ be the mean score.

$$\widehat{\text{Var}(S)}(\hat{\vartheta}_1) = \frac{1}{R}\sum_{r=1}^{R} S^{(r)}(\hat{\vartheta}_1)\left(S^{(r)}(\hat{\vartheta}_1)\right)' - \bar{S}(\hat{\vartheta}_1)\left(\bar{S}(\hat{\vartheta}_1)\right)' \tag{23}$$

The estimated observed-data information matrix is:

$$\hat{I}_{obs}(\hat{\vartheta}_1) = -\widehat{EH}(\hat{\vartheta}_1) - \widehat{\text{Var}(S)}(\hat{\vartheta}_1) \tag{24}$$

Covariance matrix and standard errors The asymptotic covariance matrix of $\hat{\vartheta}_1$ is estimated by the inverse of the observed-data information matrix:

$$\widehat{\text{Var}}(\hat{\vartheta}_1) = \left[\hat{I}_{obs}(\hat{\vartheta}_1)\right]^{-1} \tag{25}$$

To account for the simulation error introduced by using a finite number of M_{EM} draws in the E-step of the MCEM algorithm that produced $\hat{\vartheta}_1$, the covariance matrix is adjusted following Booth and Hobert (1999):

$$\widehat{\text{Var}}_{adj}(\hat{\vartheta}_1) = \left(1 + \frac{1}{M_{EM}}\right)\widehat{\text{Var}}(\hat{\vartheta}_1) \tag{26}$$

The standard errors for the elements of $\hat{\vartheta}_1$ are the square roots of the diagonal elements of $\widehat{\text{Var}}_{adj}(\hat{\vartheta}_1)$.

D.2 Stage 2: Supplier parameters

The supplier-side parameters, denoted by ϑ_2 in this section, are estimated in a second stage, conditional on the first-stage estimates of the buyer's policy parameters, $\hat{\vartheta}_1$. To correctly account for the additional uncertainty introduced by using $\hat{\vartheta}_1$ in the second-stage estimation, I compute the standard errors for $\hat{\vartheta}_2$ using a two-step M-estimator adjustment procedure (see, e.g., Newey and McFadden (1994) or Murphy and Topel (2002)).

Let $\hat{\vartheta}_1$ be the vector of parameter estimates from the first stage, with an estimated covariance matrix $\hat{\mathbf{V}}_{\vartheta_1} = \widehat{\text{Var}}_{adj}(\hat{\vartheta}_1)$. The parameters ϑ_2 are estimated by maximizing a simulated log-likelihood function, denoted as the objective function $Q_2(\vartheta_2; \hat{\vartheta}_1)$. For numerical optimization, particularly when parameters are constrained (e.g., in the case of the parameters of the Beta prior on type - constrained to be positive), I work with an unconstrained vector of parameters $\mathbf{u}$, such that $\vartheta_2 = h(\mathbf{u})$, where $h(\cdot)$ is a known transformation function (e.g., $h(u_i) = \exp(u_i)$ for a positive parameter). The objective function in terms of these unconstrained parameters is $Q_2^*(\mathbf{u}; \hat{\vartheta}_1) \equiv Q_2(h(\mathbf{u}); \hat{\vartheta}_1)$. Let $\hat{\mathbf{u}}$ be the maximizer of $Q_2^*(\mathbf{u}; \hat{\vartheta}_1)$, so that $\hat{\vartheta}_2 = h(\hat{\mathbf{u}})$ are the second-stage parameter estimates.

The asymptotic variance of $\hat{\vartheta}_2$ that accounts for the first-stage estimation uncertainty is given by:

$$\text{Var}(\hat{\vartheta}_2) = \text{Var}(\hat{\vartheta}_2|\hat{\vartheta}_1) + D\hat{V}_{\vartheta_1}D' \tag{27}$$

where $\text{Var}(\hat{\vartheta}_2|\hat{\vartheta}_1)$ is the conditional variance of $\hat{\vartheta}_2$ (treating $\hat{\vartheta}_1$ as if it were the true parameter vector), and the second term, involving the matrix D and the variance of the first-stage estimates $\hat{V}_{\vartheta_1}$, is the adjustment for the first-stage uncertainty.

Conditional covariance matrix The conditional covariance of $\hat{\vartheta}_2$, $\text{Var}(\hat{\vartheta}_2|\hat{\vartheta}_1)$, is obtained using the Delta method. First, I compute the Hessian of the second-stage objective function (negative log-likelihood) with respect to the unconstrained parameters $\mathbf{u}$. This Hessian is evaluated at the optimum $\hat{\mathbf{u}}$ and uses the first-stage estimates $\hat{\vartheta}_1$:

$$H_u = \nabla^2_{uu} Q_2^*(u; \vartheta_1)\Big|_{u=\hat{u}, \vartheta_1=\hat{\vartheta}_1} \tag{28}$$

Under standard regularity conditions for MLE, H_u is a consistent estimator of the conditional information matrix for u. The conditional covariance of $\hat{u}$ is then $\text{Var}(\hat{u}|\hat{\vartheta}_1) = H_u^{-1}$. Let $J_h = \frac{\partial h(u)}{\partial u'}\Big|_{u=\hat{u}}$ be the Jacobian matrix of the transformation $h(\cdot)$ evaluated at $\hat{u}$. The conditional covariance of $\hat{\theta}_2$ is:

$$\text{Var}(\hat{\vartheta}_2|\hat{\vartheta}_1) = J_h \left(H_u^{-1}\right) J_h' \tag{29}$$

Adjustment for first-stage uncertainty The matrix D in Equation 27 captures the sensitivity of the second-stage estimates $\hat{\vartheta}_2$ to perturbations in the first-stage estimates $\hat{\vartheta}_1$. It is defined as:

$$D = -J_h H_u^{-1} G_{u\vartheta_1} \tag{30}$$

where $G_{u\vartheta_1}$ is the derivative of the score (gradient) of the second-stage objective function (with respect to u) with respect to the first-stage parameters ϑ_1. Let $s_u(u;\vartheta_1) = \nabla_u Q_2^*(u;\vartheta_1)$ be this score vector. Then $G_{u\vartheta_1}$ is the matrix of mixed partial derivatives:

$$G_{u\vartheta_1} = \left.\frac{\partial s_u(u;\vartheta_1)}{\partial \vartheta_1'}\right|_{u=\hat{u},\vartheta_1=\hat{\vartheta}_1} \tag{31}$$

This matrix $G_{u\vartheta_1}$ quantifies how the first-order conditions for $\hat{u}$ (i.e., $s_u(\hat{u};\hat{\vartheta}_1) = 0$) would shift if ϑ_1 were to change. The term $-H_u^{-1}G_{u\vartheta_1}$ represents $\frac{\partial \hat{u}}{\partial \vartheta_1'}$.

Final standard errors The final standard errors for the elements of $\hat{\vartheta}_2$ are the square roots of the diagonal elements of the total covariance matrix $\text{Var}(\hat{\vartheta}_2)$ computed using Equation 27.

E Robustness check for supplier payments

I use a context-driven assumption in the theory and the estimation: supplier gross benefit from an order, i.e. the payment they require to complete the order, does not vary with performance on the order. This is a fact that I observe from the contracts between the buyer and the suppliers. In this subsection, I present evidence from the data that further backs up the assumption.

In Table 11, we see that the payments made to the supplier for an order are not correlated with the performance in that order, conditional on order type and order value. This is in line with the observed contract terms, which do not index supplier payments on performance i.e. suppliers are not paid bonuses or imposed fines based on performance. There might still be concerns that higher types and bigger suppliers might be offered better terms. The data doesn't seem to support such a hypothesis, as shown in Table 12, using lifetime performance as a proxy for the supplier's fundamental type. The key takeaway is that the data seems in line with the context-driven assumption of gross flow benefit being independent of performance.

While this should address concerns about the gross benefits, it could still be the case that net benefits are heterogeneous. For example, if the better suppliers are also more cost efficient, they might be getting a larger net benefit out of it even when they are paid the same amount. Since I don't have access the supplier's bill of materials, this is not something I can empirically preclude. However, I know from field visits that equipment, machinery and wages are fairly standard across the suppliers, so this story is unlikely.

Dependent Variable:	Payments (log)	
Model:	(1)	(2)
Order performance (log)	0.0460	0.0369
	(0.0552)	(0.0656)
Order value FE	Yes	Yes
Order type FE	Yes	Yes
Supplier ID FE		Yes
Observations	805	805
Within R^2	0.00110	0.00089

Clustered (supplier ID) standard-errors in parentheses
*Signif. Codes: ***: 0.01, **: 0.05, *: 0.1*

Table 11: Order performance is not predictive of payments made to the supplier for an order.

Dependent Variable:	Payments (log)		
Model:	(1)	(2)	(3)
Lifetime performance (log)	-0.0454		
	(0.1147)		
Monthly capacity (log)		-0.1606*	
		(0.0880)	
Lifetime tonnage (log)			-0.0552
			(0.0455)
Order value FE	Yes	Yes	Yes
Order type FE	Yes	Yes	Yes
Observations	805	506	738
Within R^2	0.00018	0.01002	0.00472

Clustered (supplier ID) standard-errors in parentheses
*Signif. Codes: ***: 0.01, **: 0.05, *: 0.1*

Table 12: Higher types and bigger size/capacity don't increase payments.