Customer-Driven Product Scope

Shresth Garg* Gaurav Khanna† Sagar Saxena‡

January 25, 2026

Abstract

We study how customer preferences shape firms' product introduction decisions using administrative records of 54 million transactions from a large Indian state. By tracking buyer-seller interactions over time, we document three patterns: over half of initial new product sales go to buyers with prior relationships; product introductions increase sales of complementary products; and firms are more likely to introduce products that their existing buyers purchase from other suppliers. We develop and estimate a structural model where buyers prefer purchasing from familiar sellers and value sourcing multiple products from a single supplier. These demand-side mechanisms favor incumbents with established relationships and larger product portfolios. We quantify their importance with two counterfactual exercises: reducing the preference for established sellers by 25% increases product introductions by 50%, while reducing the preference for single-sourcing by 50% increases them by 5% on average but has larger effects in markets where single-sourcing is prevalent.

*University of Pennsylvania, Department of Economics. Contact: shresth@sas.upenn.edu
†UC San Diego, School of Global Policy and Strategy. Contact: gakhanna@ucsd.edu
‡University of Pennsylvania, Wharton School. Contact: sagarsxn@wharton.upenn.edu

1 Introduction

Multi-product firms account for the vast majority of economic activity: in US manufacturing, they represent 41% of firms but 91% of output (Bernard, Redding, & Schott, 2010). Product introduction is therefore a key margin of firm growth. Much of the existing literature on product scope emphasizes supply-side factors: production capabilities, input relationships, and technology. In this paper, we focus on demand-side determinants of product scope. Specifically, we study how customer preferences over whom to buy from and whether to buy multiple products from one seller shape which firms can profitably expand and what products they introduce. If buyers prefer familiar sellers, a firm's existing clients become natural customers for new products; if they prefer bundling purchases, adding a complement raises demand for both the new and existing products. Thus, a firm's customer base and product mix may determine its expansion opportunities.

Studying these demand-side factors requires data that links firms to their specific buyers. We leverage administrative transaction records from India's electronic waybill (e-way bill) system to observe these links with rare precision. The system mandates electronic documentation for commercial shipments above a value threshold; each record lists the specific items, quantities, and values in the shipment, along with the identities and locations of both buyer and seller. Our sample covers approximately 54 million shipments over two years from a large Indian state. These records let us identify when firms introduce new products, track whether those products sell to existing clients or new buyers, and observe whether buyers source multiple products from a single seller.

We use these data to document descriptive evidence on determinants and consequences of new product introductions. We find three patterns. First, firms sell their new products primarily to existing clients. In the first month after introduction, 55 percent of sales go to buyers who had previously purchased from the firm, and this share remains above 50 percent after six months. Second, introducing a product increases sales of the firm's other products. Comparing firms that introduce products to those that never do, total firm sales rise by 95 percent around introduction, and three-quarters of this increase comes from products the firm was already selling. These gains are largest for complementary products: those frequently purchased together with the new product see bigger sales increases than other products in the firm's portfolio. Third, potential demand from existing clients predicts which products firms introduce. We measure access to buyers for each product and find that firms are roughly fifteen times more likely to introduce a product if their existing clients already buy it elsewhere than if only non-clients do. Together, these patterns suggest two demand-side mechanisms: buyers prefer familiar sellers, and buyers value sourcing multiple products from a single supplier.

Next, we estimate a structural model to quantify these mechanisms. The model treats markets as specific product bundles. For example, buyers who need both cement and steel form a distinct market from those who need only cement. In each market, buyers choose a combination of sellers. They can source the entire bundle from one firm or mix and match across multiple suppliers. This structure implies that a firm selling only cement competes in multiple markets simultaneously: it vies for buyers who want only cement, but also for those building a bundle of cement and steel. Introducing steel allows the firm to offer the complete bundle itself, which may be valued by buyers. Two parameters in the buyer demand model capture the mechanisms of interest: a single-seller bonus, which captures the value that buyers place on sourcing multiple products from the same firm, and a repeat-seller bonus, which captures the value of transacting with existing sellers. Buyers weigh these factors against price and distance. On the supply side, firms engage in Bertrand-Nash price competition and decide whether to introduce new products by comparing expected profit gains to a fixed cost. Expected profit gain has two components: direct sales of the new product and indirect effects on existing products.

We estimate the model in two stages. The first stage recovers demand parameters using a nested fixed-point procedure that matches observed market shares and maximizes the likelihood of observed purchase decisions; we then invert firms' first-order conditions to recover marginal costs. Our estimates show that both the repeat-seller and single-seller bonuses are positive: buyers prefer existing sellers and value bundling from a single source. The second stage estimates the fixed cost distribution from observed entry decisions. For each potential product introduction, we simulate expected profit gain by drawing quality and cost realizations and solving for post-entry equilibrium prices. We then recover the distribution of fixed costs by matching the entry probabilities implied by these profit gains to the actual entry rates in the data.

We use the estimated model to run counterfactual experiments which examine how product introduction would change if buyers cared less about purchasing from existing sellers or about sourcing multiple products from the same firm. In the first experiment, we reduce the repeat-seller bonus by 25 percent, representing an environment where existing relationships confer a smaller advantage. This raises the mean product introduction probability by 50 percent. The mechanism is straightforward. Incumbents have been around longer and have larger existing customer bases than potential entrants. Reducing the repeat-seller bonus lowers the return on these established relationships. This makes it easier for entrants to compete for buyers and increases the expected profits of entry. The second experiment reduces the single-seller bonus by 50 percent. This also increases aggregate entry, though the effect is smaller. However, the probability of entry rises significantly in markets where incumbents derive a greater share of their revenue from single sourced transactions. In these markets, reducing the bonus erodes the portfolio advantage of multi-product firms and increases the incentives for new firms to enter.

These counterfactual simulations show that buyer preferences for existing sellers and single-sourcing favor established firms by rewarding incumbents with deep customer networks and large portfolios. In contrast, entrants and smaller firms are constrained in their ability to scale. This creates a feedback loop where past success in building a customer base lowers the cost of future expansion. Consequently, market structure may remain concentrated not because incumbents are more productive, but because their demand-side assets (their relationships and portfolios) insulate them from competition.

Related Literature. This paper contributes to several literatures. First, it relates to work on multi-product firms and the determinants of product scope. Much of this work has focused on supply-side determinants: firms expand into products where they possess relevant production capabilities (Bernard, Redding, & Schott, 2010; Eckel & Neary, 2010), where they can exploit existing input relationships (Boehm, Dhingra, & Morrow, 2022; Flagge & Chaurey, 2014; Goldberg et al., 2010), or where competitive pressure is less intense (Mayer, Melitz, & Ottaviano, 2014). The demand side has received less attention. A notable exception is Bernard et al. (2019), who document that manufacturers frequently export products they do not produce, a pattern they attribute to demand complementarities favoring multi-product sellers. We develop a structural demand model that incorporates both this bundling channel and an additional force: buyer-seller relationships built through prior transactions. The model allows us to estimate the contribution of each channel from observed purchase and entry decisions.

Second, we contribute to research on buyer-seller relationships. Prior work shows that building a customer base requires costly investment (Foster, Haltiwanger, & Syverson, 2016; Gourio & Rudanko, 2014), that switching suppliers is expensive (Monarch, 2022), and that established relationships are therefore valuable assets (Eaton et al., 2025). This literature has focused on how relationships affect trade volumes. We show they also affect product scope: firms rely on existing relationships when introducing new products.

Finally, our findings speak to firm growth in developing economies. A large literature examining why firms remain small in these settings emphasizes supply-side constraints: lack of capital, limited managerial capacity, regulatory barriers, or technological constraints. Recent work highlights that demand-side factors also matter: market size and trade access affect firms' incentives to adopt efficient technologies and enter new markets (Goldberg & Reed, 2023; Leone, Macchiavello, & Reed, 2025). We identify a specific demand-side barrier: firms without established customer networks face difficulty introducing new products. This barrier may be especially relevant in developing countries, where search, contracting, and information frictions could make buyers more hesitant to transact with unfamiliar sellers and thereby impede firm growth through product expansion.

The remainder of the paper proceeds as follows. Section 2 describes the institutional setting and data. Section 3 presents descriptive evidence on the role of existing clients and complementarity in product introduction. Section 4 develops the structural model. Section 5 describes the estimation procedure and presents results. Section 6 conducts counterfactual experiments. Section 7 concludes.

2 Setting and Data

To study how firms expand their product portfolios, we use administrative transaction records from India's e-way bill system. These data track the universe of formal sector shipments above a value threshold, recording the identities and locations of buyers and sellers, the products shipped, and the transaction value. We observe each firm's complete shipment history and can therefore identify when firms begin selling new products and to whom.

2.1 The E-Way Bill System

In 2017, India introduced a Goods and Services Tax (GST) administered through a centralized digital platform. The reform mandated electronic waybills (e-way bills) for commercial shipments, creating comprehensive real-time records of goods movements. Any shipment above Rs. 50,000 (approximately $600) requires an electronic document recording buyer, seller, product, and value. These records form the basis of our data.

Three features of this system generate high-quality data. First, the legal mandate ensures that all formal sector shipments above the threshold are documented. The Rs. 50,000 cutoff is low enough to capture the bulk of business-to-business transactions while filtering out small retail purchases. Second, penalties and physical enforcement at highway checkpoints create compliance incentives: officers can intercept vehicles and detain goods lacking valid documentation. Third, the tax structure makes the system self-policing. Buyers can claim credits for taxes paid on their purchases, but only with valid documentation. This creates pressure on sellers to report accurately.

2.2 Data and Sample

We obtain the universe of e-way bills involving sellers or buyers in a large Indian state for the period April 2018 through March 2020.[1] Each e-way bill records the tax identifiers and locations of both buyer and seller, the value and quantity of goods, and the HSN (Harmonized System of Nomenclature) code classifying the product.

We define firms at the establishment-level; in our sample, a firm is a combination of tax identifier (GSTIN) and location identifier (PIN code). The same company operating from two locations appears as two distinct firms. This approach treats

[1] We end the sample in March 2020 to avoid the disruptions of the COVID-19 pandemic. See Appendix A for details on raw data characteristics and sample construction.

each establishment as a separate economic unit, which aligns with our interest in understanding location-specific product decisions.

2.3 Defining Product Introductions

A product introduction is a firm's first shipment of a product category that was not part of its initial portfolio. To operationalize this concept, we impose three conditions. First, the introduction must occur at least six months after the firm's initial appearance in the data. This lag ensures we observe expansion rather than the firm's starting product mix. Second, the firm must ship the product in more than half of the months following introduction. This requirement distinguishes genuine additions from experimental shipments that the firm quickly abandoned. Third, the introduction must happen at least six months before the end of our sample, so we have adequate time to observe post-introduction outcomes. We provide details on how we construct our sample for reduced-form analysis and structural estimation in Appendix A.

3 Descriptive Evidence

Using this sample of product introductions, we explore two questions. First, what happens after introduction? Does the firm rely on existing clients to purchase the new product, and does the introduction affect sales of existing products? Second, what predicts which products firms introduce? Do geographic proximity, client relationships, and complementarity shape these decisions? To guide our empirical analysis, in Appendix B, we develop a stylized model that formalizes these intuitions and generates testable predictions.

We find three patterns. Firms rely heavily on existing clients when selling new products, with over half of initial sales going to buyers with prior relationships. Product introductions increase sales of complementary products already in the firm's portfolio. And market access to existing clients, particularly for complementary products, predicts which products firms introduce. These patterns inform the structural model we develop in the next section.

3.1 Sales following new product introduction

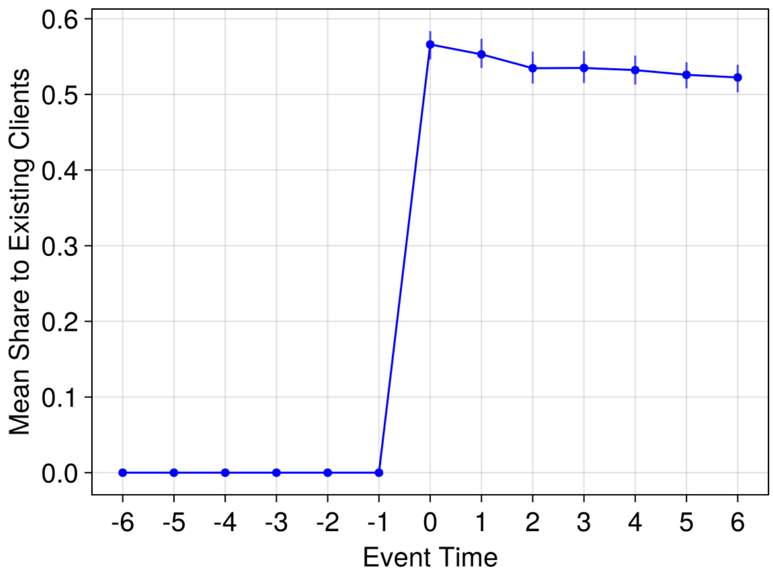
We denote sellers by i, products by p, and time by t. The month of a product introduction is normalized to $t = 0$.

3.1.1 Client Composition

Product introduction differs from entry into a new market because introducing firms have an existing customer base. We begin by quantifying how important these relationships are. If existing clients derive utility from purchasing from familiar sellers, they should account for a large share of new product sales.

We define existing clients as those who purchased from the firm in any month prior to the product introduction. The share of sales to existing clients, for product p by firm i in month t, is the fraction of that product's monthly sales going to buyers who previously purchased from the firm. We set this share to zero for pre-event periods ($t < 0$). Figure 1 plots the average of this share across all product introductions. Firms rely heavily on their existing customer base when launching new products. In the first month after introduction, 55% of sales go to existing clients. This share declines only modestly (to roughly 50%) over the following six months, suggesting that existing relationships remain an important source of demand well beyond the

Figure 1: *Share of Sales to Existing Clients*



Notes. This figure plots the share of new product sales going to buyers who previously purchased from the firm. Month 0 marks product introduction; pre-introduction shares are set to zero by construction. 95% bootstrap confidence intervals are shown.

initial launch.

3.1.2 Sales of existing products

Existing clients purchase new products, but does product introduction affect the firm's existing business? If buyers prefer to source bundles from a single seller, introducing a complementary product should increase sales of existing products.

We consider two outcome variables: *overall sales*, the sum of sales across all products in a given firm-month, and *prior sales*, sales of products that the firm was selling prior to introduction. We use an event study design to estimate how outcomes evolve around product introduction. For firm i that introduces a product, we estimate:

$$y_{i,t} = \alpha_i + \alpha_t + \sum_{\tau=-k}^{k} \beta_\tau D_{i,t}^{\tau} + \varepsilon_{i,t} \tag{1}$$

where $D_{i,t}^{\tau} = \mathbb{1}[t - G_i = \tau]$ and G_i is the month of product introduction. We use the Callaway and Sant'Anna (2021) estimator, comparing treated firms to control firms that never introduce new products during our sample period. We transform both outcome variables using the inverse hyperbolic sine.

Figure 2 presents the results. Product introduction increases overall sales by 95% and prior sales by 75%. The gap between these two figures implies that roughly 20 percentage points comes from the new product itself, while 75 percentage points represents spillovers to existing products. This analysis suggests that product introduction is not simply adding a new revenue stream; it amplifies the firm's existing business.

The theory predicts that this increase in prior sales should concentrate among complements. To test this, we need a measure of product complementarity. We use our transaction data to identify which products are purchased together in practice. Specifically, we measure complementarity using lift scores, which capture how much more likely two products are to be purchased together than would be expected if purchases were independent. For products p and p', the lift score is:

$$\text{lift}(p, p') = \frac{\Pr(p \cap p')}{\Pr(p) \times \Pr(p')} \tag{2}$$

where $\Pr(p)$ denotes the probability that product p is purchased in a given transaction, and $\Pr(p \cap p')$ denotes the probability that both products p and p' are purchased together in the same transaction. Higher lift scores indicate stronger complementarity.

To avoid mechanical relationships, we compute lift scores excluding all shipments from firms that introduce products during our sample period.[2] For each new product,

[2]If we included these firms, a mechanical bias would arise: firms that introduce a new product and sell it

we use the 90th percentile of its lift distribution as the threshold to classify other products as high complementarity (above threshold) or low complementarity (below threshold).[3]

We run separate event studies for sales of high and low complementarity products, restricting to treated firms that sell both types. Figure 2 shows that sales of both types increase, but the increase is substantially larger for high complementarity products. This differential effect by complementarity is consistent with bundled purchasing: buyers who value sourcing both products from the same seller increase purchases of both when the firm offers both.

3.1.3 Prices

Having established that product introduction increases sales of existing products, we now ask whether firms adjust prices. Bundling creates opposing incentives: lowering prices attracts more buyers to the bundle, but raising prices extracts surplus from buyers who value bundle convenience. The net effect is an empirical question.

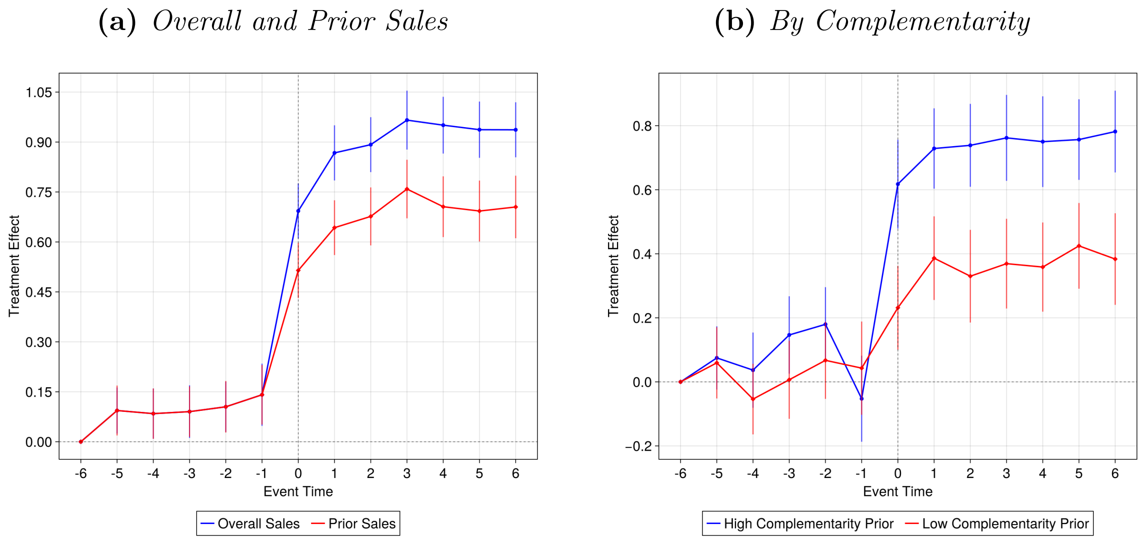
We define *bundled price* as the price of an existing product when sold together with the newly introduced product in the same transaction. We compare bundled prices to prices of the same product when sold separately, isolating the effect of bundling on prices while holding product characteristics fixed. The sample includes all firm-products with at least one sale every month, whether bundled or unbundled.

Figure 3 presents the event study results. Prices of existing products decrease

alongside existing products would generate co-purchase patterns reflecting their product portfolio decisions rather than underlying demand complementarity. This would inflate lift scores for product pairs sold by introducing firms, biasing our complementarity measure toward finding stronger effects precisely for the products we study.

[3] A significant proportion of lift scores are zero because many product pairs are never purchased together. In cases where the 90th percentile of lift scores for product p is zero, we classify all products with positive lift scores as having high complementarity with p.

Figure 2: *Treatment Effects on Sales Around New Product Introduction*



Notes. Panel (a) shows treatment effects on overall sales and sales of products the firm sold prior to introduction. Panel (b) decomposes prior sales by complementarity with the newly introduced product, where high complementarity denotes products with lift scores above the 90th percentile. Both panels report Callaway and Sant'Anna (2021) estimates with month 0 marking introduction. All outcomes are transformed using the inverse hyperbolic sine.

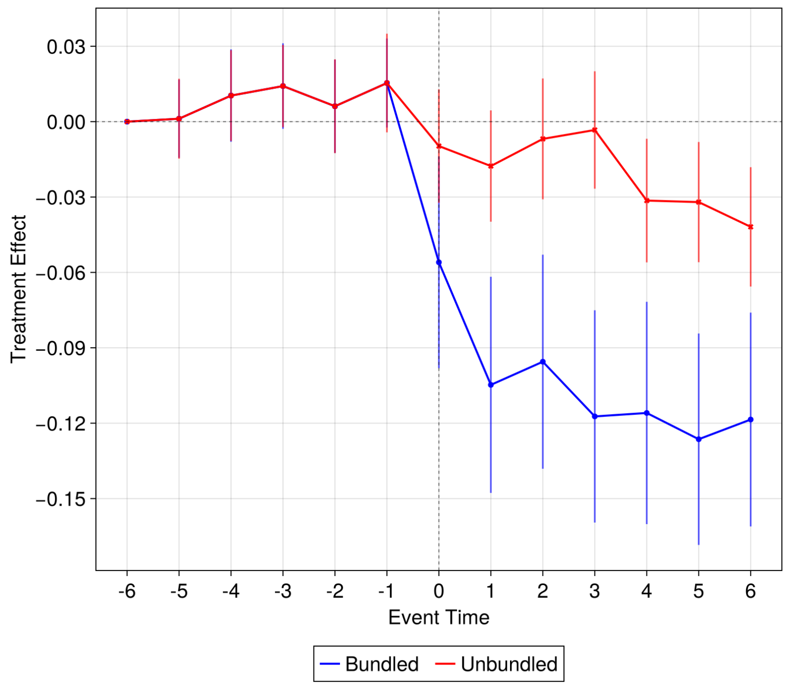
when bundled with the newly introduced product. This suggests that the demand expansion force dominates: firms lower prices on existing products to increase sales of both products via the bundle.

3.2 Market Access and Product Introduction

The previous results document what happens after introduction. We now ask whether firms anticipate these benefits when deciding which products to add. If firms understand that existing clients and complementarity raise returns to product introduction, these factors should predict introduction decisions.

Testing this requires measuring each firm's access to relevant demand. A firm's

Figure 3: *Treatment Effects on Bundling Prices*



Notes. This figure shows how prices of existing products change when sold together with the newly introduced product in the same transaction, relative to sales of those products sold separately. The sample includes firm-products with at least one observation in every month of the event window. Estimates use the Callaway and Sant'Anna (2021) estimator.

potential buyers are scattered geographically, and closer buyers are presumably easier to serve. We therefore construct a gravity-based measure, *market access*, that sums over buyers weighted by their purchasing volume and discounted by distance. A firm with high market access faces strong nearby demand; a firm with low market access faces either weak demand or distant buyers. For seller i and product p, we define:

$$\text{Market Access}_{i,p} = \sum_{j \in \mathcal{J}} \frac{\text{asinh}(\text{purchases}_{j,p})}{(\text{distance}_{ij})^2} \tag{3}$$

where asinh is the inverse hyperbolic sine, $\text{purchases}_{j,p}$ is the average monthly purchase of product p by buyer j, and distance_{ij} is the geographic distance between seller i and buyer j.

We compute market access using transaction data from each firm's first six months in the sample. This window precedes any product introduction, since our treatment definition requires introductions to occur at least six months after firm entry. Existing clients are buyers who purchased from the firm during this initial period; all other buyers are classified as new.[4] The set $\mathcal{J}$ varies by specification: all buyers, existing clients only, or new buyers only.

Each firm can potentially introduce thousands of products, so expanding the choice set generates a very large number of firm-product observations. To make estimation tractable, we retain all treated firms but randomly subsample control firms, keeping four control firms for each treated firm. We apply inverse probability weights to recover population-level estimates.

We estimate:

$$\text{introduced}_{i,p} = \beta \cdot \mathbb{1}[\text{High Market Access}_{i,p}] + \gamma \mathbf{X}_{i,p} + \lambda + \varepsilon_{i,p} \tag{4}$$

where $\text{introduced}_{i,p}$ is an indicator for whether seller i introduces product p. High market access is defined as above the median within each product category, so the comparison is between firms with high versus low market access for the same product. We progressively add fixed effects to isolate different sources of variation.

Table 1 shows that access to potential demand is positively associated with product introduction. All columns include product fixed effects, which compare firms with high versus low market access for the same product. The first column shows that high market access increases the probability of product introduction by approximately 0.3 basis points. With a baseline introduction probability of 0.5 basis points, this represents a 60% increase. The second column adds market access to all products

[4] This definition differs from the event study, where existing clients are defined relative to the introduction date. We cannot use that definition here because control firms—those that never introduce products—have no introduction date. Using the first six months after entry provides a consistent baseline across treated and control firms. This choice is conservative: clients acquired between month six and the introduction date are classified as new buyers, understating the role of existing relationships.

Table 1: *Market Access and Product Introduction*

	Introduced Product (Basis Points)			
	(1)	(2)	(3)	(4)
High Market Access (Product)	0.33394***	0.26880***	0.30656***	0.33847***
	(0.055)	(0.044)	(0.041)	(0.042)
High Market Access (All Products)		0.09067		
		(0.050)		
Product (HSN4) Fixed Effects	Yes	Yes	Yes	Yes
Pincode Fixed Effects			Yes	Yes
Firm Fixed Effects				Yes
N	4,038,457	4,038,457	4,038,457	4,038,457
R^2	0.000	0.000	0.002	0.006
Mean of Dep. Var.	0.506	0.506	0.506	0.506

*Notes. The dependent variable is an indicator for whether the firm introduces the product; coefficients represent changes in introduction probability in basis points. High market access is defined as above the median within each product category. Column 2 includes market access to all products as a control. Standard errors clustered by firm in parentheses. Inverse probability weights account for subsampling of control firms. *** p<0.01, ** p<0.05, * p<0.1.*

as a control, capturing total potential demand in the firm's location; this coefficient is small and insignificant, indicating that product-specific demand potential, rather than simply being located in a high-demand area, drives introduction decisions. The third column adds pincode fixed effects to absorb all location-specific differences and compare firms in the same location considering the same product. The coefficient remains statistically significant. The fourth column adds firm fixed effects, comparing products within the same firm. This is our preferred specification, with a coefficient of 0.34 basis points.

Table 2: *Market Access: Existing Clients vs New Buyers*

	Introduced Product (Basis Points)				
	(1)	(2)	(3)	(4)	(5)
High Market Access (Existing Clients)	2.62296***	2.80017***	2.88213***	3.21959***	3.50887***
	(0.199)	(0.208)	(0.212)	(0.312)	(0.366)
High Market Access (New Clients)	0.09696	0.17489***	0.21054***	-0.78079**	-0.43947
	(0.062)	(0.042)	(0.043)	(0.254)	(0.288)
Product (HSN4) Fixed Effects	Yes	Yes	Yes		
Pincode Fixed Effects		Yes	Yes		
Firm Fixed Effects			Yes		Yes
Product (HSN4) × Pincode Fixed Effects				Yes	Yes
N	4,038,457	4,038,457	4,038,457	3,155,505	4,038,457
R^2	0.000	0.002	0.006	0.105	0.236
Mean of Dep. Var.	0.506	0.506	0.506	0.506	0.506

*Notes. The dependent variable is an indicator for whether the firm introduces the product; coefficients represent changes in introduction probability in basis points. Market access is computed separately for existing clients (buyers who purchased during the firm's first six months) and new buyers. High market access is defined as above the median within each product category. Standard errors clustered by firm in parentheses. Inverse probability weights account for subsampling of control firms. *** $p<0.01$, ** $p<0.05$, * $p<0.1$.*

Table 2 decomposes market access by buyer type. We construct the same measure as before, but now separately summing over existing clients and new buyers. The first three columns progressively add product, pincode, and firm fixed effects. In our preferred specification with all three fixed effects, the coefficient on existing clients is 2.9 basis points, while the coefficient on new buyers is 0.2 basis points. Thus, firms appear to weight existing relationships heavily when choosing products.

The final two columns add product-by-pincode fixed effects, comparing firms in the same location considering the same product. Since market access is constructed from distance to buyers, firms in the same pincode have identical distances to all potential buyers. This means there is little variation in market access to new buyers across firms in the same location—they all face similar pools of potential new buyers. In contrast, existing client relationships vary substantially across firms even within the same location. The coefficient on existing clients remains large and precisely estimated (3.2–3.5 basis points).

Table 3 tests the role of complementarity in product introduction. We use the same lift-based measure of complementarity from the previous analysis looking at impact on sales. For each product p, we compute lift scores with all other products and classify those above the 90th percentile as high complementarity. We then construct market access to high complementarity products by summing over buyers who purchase products with high lift relative to p. The first column includes product fixed effects, while the second and third columns progressively add pincode and firm fixed effects. In our preferred specification with all three fixed effects, the coefficient on high complementarity market access is 0.31 basis points and the coefficient on product-specific market access is 0.36 basis points; both are statistically significant at the 1% level. Relative to the baseline introduction probability of 0.5 basis points, these represent increases of roughly 60% and 70%, respectively. Thus, complementarity provides additional explanatory power beyond access to buyers of the focal

Table 3: *Market Access with Complementarity*

	Introduced Product (Basis Points)		
	(1)	(2)	(3)
High Market Access (Product)	0.25580***	0.32421***	0.35993***
	(0.047)	(0.046)	(0.047)
High Market Access (High Complementarity)	0.15252**	0.24465**	0.30798***
	(0.053)	(0.088)	(0.084)
Product (HSN4) Fixed Effects	Yes	Yes	Yes
Pincode Fixed Effects		Yes	Yes
Firm Fixed Effects			Yes
N	3,603,740	3,603,740	3,603,740
R^2	0.000	0.002	0.006
Mean of Dep. Var.	0.506	0.506	0.506

*Notes. The dependent variable is an indicator for whether the firm introduces the product; coefficients represent changes in introduction probability in basis points. High complementarity products are those with lift scores above the 90th percentile relative to the focal product. High market access is defined as above the median within each product category. Standard errors clustered by firm in parentheses. Inverse probability weights account for subsampling of control firms. *** p<0.01, ** p<0.05, * p<0.1.*

product.

Table 4 interacts the buyer-type and complementarity decompositions. We construct four measures: market access to existing clients and new buyers for the focal product (as in Table 2), and market access to existing clients and new buyers for high-complementarity products. The latter two measures capture access to buyers of complementary products, further decomposed by whether those buyers have an existing relationship with the firm. The first three columns progressively add product, pincode, and firm fixed effects. In our preferred specification with all three fixed

Table 4: *Market Access: Buyer Type and Complementarity*

	Introduced Product (Basis Points)				
	(1)	(2)	(3)	(4)	(5)
High Market Access (Existing Clients)	2.59055***	2.77404***	2.91994***	3.14266***	3.57677***
	(0.207)	(0.211)	(0.217)	(0.302)	(0.372)
High Market Access (New Clients)	0.06523	0.18781***	0.22029***	-0.85230**	-0.55209
	(0.048)	(0.047)	(0.048)	(0.288)	(0.331)
High Market Access (High Comp., Existing Clients)	0.14901*	0.35825***	0.51229***	0.38331**	0.94372***
	(0.069)	(0.082)	(0.065)	(0.129)	(0.176)
High Market Access (High Comp., New Clients)	-0.02849	0.16558	0.25010**	-0.35830	0.02898
	(0.059)	(0.088)	(0.083)	(0.415)	(0.392)
Product (HSN4) Fixed Effects	Yes	Yes	Yes		
Pincode Fixed Effects		Yes	Yes		
Firm Fixed Effects			Yes		Yes
Product (HSN4) × Pincode Fixed Effects				Yes	Yes
N	3,603,740	3,603,740	3,603,740	2,818,733	3,603,740
R^2	0.000	0.002	0.007	0.104	0.236
Mean of Dep. Var.	0.506	0.506	0.506	0.506	0.506

*Notes. The dependent variable is an indicator for whether the firm introduces the product; coefficients represent changes in introduction probability in basis points. Market access is computed separately by buyer type (existing clients vs new buyers) and by complementarity (focal product vs high-complementarity products). High market access is defined as above the median within each product category. Standard errors clustered by firm in parentheses. Inverse probability weights account for subsampling of control firms. *** $p<0.01$, ** $p<0.05$, * $p<0.1$.*

effects, the coefficient on existing clients is 2.9 basis points and the coefficient on high-complementarity existing clients is 0.51 basis points.

The final two columns add product-by-pincode fixed effects, comparing firms in the same location considering the same product. Most of the variation in complementarity market access comes from existing clients. In our preferred specification with all fixed effects, the coefficient on existing clients is 3.6 basis points and the coefficient on high-complementarity existing clients is 0.94 basis points.

3.3 Summary

These results establish three patterns. First, firms rely on existing clients for over half of new product sales, suggesting that relationship-specific advantages shape returns to product introduction. Second, introducing a product increases sales of complementary products, so firms cannot evaluate new products in isolation. Third, market access to existing clients, particularly for complementary products, predicts which products firms choose to introduce.

These patterns can be rationalized by the following mechanisms: buyers prefer to purchase from familiar sellers and to consolidate purchases with a single supplier. The next section develops a structural model that formalizes these mechanisms.

4 Model

We develop a static model of product introduction that captures three features documented in the previous section: the advantage existing clients confer for new product sales, the demand complementarity between products in a firm's portfolio, and the role of geographic proximity in shaping which products firms introduce. The model incorporates these mechanisms through a relationship bonus β_r and a single-seller bonus γ. On the demand side, buyers purchase bundles of goods and choose which sellers to source from. These choices depend on geographic distance to sellers, existing client relationships, and a preference for sourcing the entire bundle from a single firm. On the supply side, firms engage in Bertrand-Nash price competition and decide which new varieties to introduce after weighing expected variable profits against fixed costs of introduction.

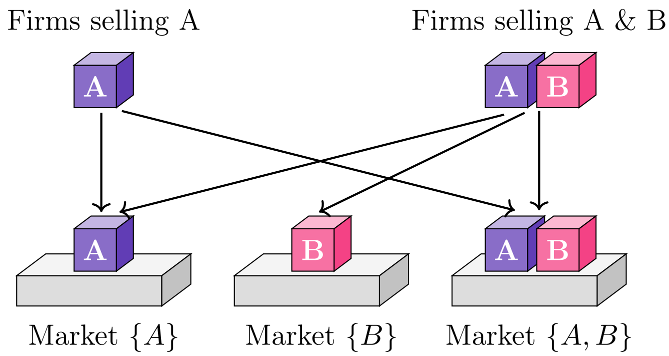
4.1 Setup

We denote HSN codes by h, firms by f, and time periods by t. A *good* is an HSN code while a *variety* is a firm-good pair (f, h), representing the particular offering that firm f provides for good h. Buyers may require a single good or multiple goods at once. A *bundle* $b = (h_1, \ldots, h_K)$ is an ordered collection of one or more goods that a buyer requires. Single-good bundles correspond to standalone purchases; multi-good bundles correspond to buyers who need multiple different goods in a single transaction. For each good in the bundle, the buyer must choose which firm's variety to purchase.

A *market* (b, t) is defined by a bundle b and a time period t. Figure 4 illustrates the market structure. Because markets are defined by bundles, the same product can appear in multiple markets. For example, a product purchased both alone and as part of a bundle participates in two distinct markets. Each market has a set of buyers $i \in I_b$, and each buyer demands specific quantities $\{q_{i,h_k}\}_{k=1}^{K}$ of the goods in the bundle. We take quantities as given; the buyer's decision is which variety to purchase for each good.

The model proceeds in two stages. In the first stage, firms decide which new goods to add to their portfolios. Each firm currently sells a set of goods $\mathcal{H}_f$, offering

Figure 4: *Markets Defined by Product Bundles*



Notes. Firms compete in all markets where their products are demanded. Single-product firms selling only A compete in markets $\{A\}$ and $\{A, B\}$. Multi-product firms selling both A and B compete in all three markets.

one variety per good. Firms consider expanding into goods they do not currently sell; adding good h means creating a new variety (f, h). In the second stage, buyers observe which varieties are available (including any new varieties that firms chose to introduce) and make purchase decisions. We begin with the second stage, characterizing demand and pricing for a given set of varieties, then turn to the entry decision that determines which varieties are offered.

4.2 Demand

We focus on a single market (b, t) and suppress b and t from notation where clear. For each good in a bundle, buyers choose which firm's variety to purchase. When a bundle contains multiple goods, buyers may source all goods from a single firm or divide purchases across multiple firms. We represent this choice as a *seller combination* $\mathbf{f} = (f_1, \ldots, f_K)$, where f_k is the firm supplying good h_k. The choice set $\mathcal{C}_b$ consists of all feasible seller combinations—those where each firm f_k sells good h_k.

A buyer's utility from a seller combination depends on the quality and price of each variety, the distance to each seller, existing relationships with each seller, and whether all goods come from the same firm.Since buyers demand different quantities of each good, we weight each good by its share of total quantity: $w_{ik} = q_{i,h_k} / \sum_{k'} q_{i,h_{k'}}$ is the quantity weight of good k in buyer i's bundle.

Buyer i's utility from seller combination $\mathbf{f}$ depends on the varieties purchased, the distance to each seller, whether the buyer has transacted with each seller before, and whether all goods come from the same firm. Formally, $U_{i\mathbf{f}} = V_{i\mathbf{f}} + \varepsilon_{i\mathbf{f}}$, where V is the deterministic component and ε is a taste shock. The deterministic component is:

$$V_{i\mathbf{f}} = \sum_{k=1}^{K} w_{ik} \cdot v_{i,f_k,h_k} + \gamma \cdot \mathbf{1}[\text{single}] + \beta_d \cdot \log\left(\sum_{k=1}^{K} w_{ik} \cdot d_{i,f_k}\right) \tag{5}$$

$$v_{i,f_k,h_k} = \delta_{f_k,h_k} + \beta_r \cdot \mathbf{1}[\text{repeat}_{i,f_k}]$$

$$\delta_{fh} = \alpha \cdot p_{fh} + \xi_{fh}$$

The first equation has three terms. The first term aggregates per-good utilities v_{i,f_k,h_k} across goods in the bundle, weighted by quantity shares w_{ik}. Per-good utility v_{i,f_k,h_k} is buyer i's utility from purchasing good h_k from firm f_k. It consists of two components: the mean utility δ_{f_k,h_k} of variety (f_k, h_k), and a relationship bonus $\beta_r \cdot \mathbf{1}[\text{repeat}_{i,f_k}]$ that adds utility when buyer i has previously purchased from firm f_k. Mean utility δ_{fh} decomposes into price p_{fh} and unobserved variety quality ξ_{fh}.

The second term $\gamma \cdot \mathbf{1}[\text{single}]$ is a single-seller bonus that adds utility when all goods in the bundle come from the same firm. The third term $\beta_d \cdot \log(\sum_k w_{ik} \cdot d_{i,f_k})$ captures weighted distance to sellers, where d_{i,f_k} is the distance between buyer i and firm f_k.

We assume taste shocks $\varepsilon_{i\mathbf{f}}$ are independent and identically distributed Type-I extreme value across buyers and seller combinations. The probability that buyer i chooses seller combination $\mathbf{f}$ is:

$$s_{i\mathbf{f}} = \frac{\exp(V_{i\mathbf{f}})}{\sum_{\mathbf{f}' \in \mathcal{C}_b} \exp(V_{i\mathbf{f}'})} \tag{6}$$

For estimation and counterfactuals, we also require the marginal probability that buyer i purchases good h from firm f. This marginal probability sums over all seller combinations where f supplies good h:

$$S_{ifh} = \sum_{\mathbf{f}' \in \mathcal{C}_b} s_{i\mathbf{f}'} \cdot \mathbf{1}[f'_h = f] \tag{7}$$

where f'_h denotes the firm supplying good h in seller combination $\mathbf{f}'$.

Two features of demand are central to our analysis. First, existing client relationships provide a demand-side advantage. Buyers who have previously transacted with a firm receive additional utility from any subsequent purchase, captured by the relationship bonus β_r. This creates a built-in customer base for new products: when a firm introduces a new good, its existing clients are predisposed to purchase from the firm rather than from competitors. Second, demand complementarities arise from bundled purchasing. The single-seller bonus γ captures the preference for sourcing multiple goods from the same firm, reflecting reduced transaction costs, simplified logistics, or relationship benefits. Consider a buyer requiring goods A and B. A firm selling only good A competes at a disadvantage against competitors offering both goods. Introducing good B allows the firm to offer this bundle convenience. This creates demand complementarity: adding B to the portfolio not only generates direct sales of B but also increases sales of A by making the firm's bundle more attractive.

4.3 Supply

We now turn to pricing. Firms engage in Bertrand-Nash competition and simultaneously choose prices to maximize profits. Each firm f sets a single price p_{fh} for good h, which applies to all sales regardless of the bundle in which the variety appears.

Firm f's total profit Π_f aggregates across bundle markets b where the firm sells at least one variety. Within each market, it sums over buyers i and seller combinations

$\mathbf{f}'$, weighting each combination by the probability $s_{i\mathbf{f}'}$ that buyer i chooses it:

$$\Pi_f = \sum_{b\in\mathcal{B}} \sum_{i\in I_b} \sum_{\mathbf{f}'\in\mathcal{C}_b} s_{i\mathbf{f}'} \left(\sum_{k=1}^{K} \mathbf{1}[f'_k = f] \cdot (p_{fh_k} - mc_{fh_k}) \cdot q_{i,h_k} \right) \tag{8}$$

The inner sum computes the contribution from each good in the bundle. The indicator $\mathbf{1}[f'_k = f]$ equals one when firm f supplies good k in seller combination $\mathbf{f}'$, so the firm earns a profit only on goods it actually sells. Each good contributes $(p_{fh_k} - mc_{fh_k}) \cdot q_{i,h_k}$: price minus marginal cost mc_{fh}, times the quantity q_{i,h_k} that buyer i demands.

Firms set prices to maximize profit. The first-order conditions yield a system of equations linking prices, marginal costs, and quantities (see Appendix C for the full derivation):

$$\boldsymbol{\Omega}_f(\mathbf{p}_f - \mathbf{mc}_f) = \mathbf{Q}_f/\alpha \tag{9}$$

where $\mathbf{p}_f$ and $\mathbf{mc}_f$ are the vectors of prices and marginal costs for firm f's products, $\mathbf{Q}_f$ is the vector of quantities sold, and $\boldsymbol{\Omega}_f$ is a matrix capturing how prices affect demand within the firm's portfolio.

To illustrate, consider a firm selling two goods, A and B. The first-order condition for good A is:

$$\Omega_{AA}(p_A - mc_A) + \Omega_{AB}(p_B - mc_B) = Q_A/\alpha$$

The first term captures the own-price effect that governs how raising p_A reduces demand for A. The second term is the cross-price effect. It captures the impact of raising p_A on the demand for B. This effect operates through bundles containing both goods—when buyers shift away from seller combinations that include the firm for A, they also reduce purchases of B. Multi-product firms internalize these cross-price effects when setting prices.

4.4 Entry

Let $\mathcal{H}_f$ denote firm f's current product portfolio. Introducing a new product $h \notin \mathcal{H}_f$ requires a fixed cost F_{fh} and generates a change in variable profit. This change in profits consists of a direct effect and an indirect effect. The direct effect is profit from selling the new product h itself. The indirect effect operates through bundling complementarities: adding h makes the firm more competitive for buyers who demand bundles containing both h and products already in $\mathcal{H}_f$. These buyers can now source the entire bundle from firm f, receiving the single-seller bonus γ, which raises demand for the firm's existing products. The indirect effect is unambiguously positive because markets are defined by bundles with fixed buyer populations who demand fixed quantities. Buyers choose only which firm to source from; adding product h never makes firm f less attractive for any product $A \in \mathcal{H}_f$.

There are two key forces in our setting that govern how much profits change when a firm adds a new product. First, existing clients receive the relationship bonus β_r, which raises their purchase probability for the new product. Firms with larger client bases therefore earn higher direct profits from product introductions. Second, if the firm already sells complements, introducing h creates single-seller options for buyers of bundles containing both products. The single-seller bonus γ makes the firm's bundle more attractive relative to competitors who cannot offer the same convenience, amplifying the indirect effect on existing product sales.

Prior to entry, the firm does not observe the quality ξ_{fh} or marginal cost mc_{fh} it will realize. We assume the firm draws (ξ_{fh}, mc_{fh}) from the empirical distribution G_h of incumbents selling good h. Pre-entry profit $\Pi_f(\mathcal{H}_f)$ aggregates across all bundle markets where the firm currently participates, following Equation (8). Post-entry profit $\Pi_f(\mathcal{H}_f \cup \{h\}; \xi, mc)$ depends on the realized draw and is computed at the

Bertrand-Nash equilibrium prices that prevail after entry. Expected profit gain is:

$$E[\Delta\Pi_f(h)] = E_{G_h}\left[\Pi_f(\mathcal{H}_f \cup \{h\}; \xi, mc)\right] - \Pi_f(\mathcal{H}_f) \tag{10}$$

Fixed costs are drawn from a lognormal distribution with CDF $\Phi_{FC}(\cdot; \mu, \sigma)$. Firm f introduces good h if the realized fixed cost falls below expected profit gain. The probability of entry is therefore:

$$\Pr(\text{entry}) = \Phi_{FC}(E[\Delta\Pi_f(h)]; \mu, \sigma)$$

Thus, higher expected profits increase the probability of entry. Note that we do not model strategic interactions in entry decisions. Each firm takes competitors' portfolios as fixed, which is appropriate given the low entry rates in our data: simultaneous entry by multiple firms into the same product in the same period is extremely rare.[5]

5 Estimation

We estimate the model in two stages. The first stage estimates demand parameters and recovers marginal costs from observed purchase decisions. The second stage estimates the fixed cost distribution from observed entry decisions.

Estimation requires defining valid bundles and classifying observed purchases. We define valid bundles as product combinations that appear together on a single e-way bill. We then assign each transaction to a bundle. If a single e-way bill contains a valid bundle, we classify the purchase as a same-seller bundle. If multiple transactions from different sellers within the same month form a valid bundle, they are matched into a multi-seller bundle. Otherwise, transactions are labeled as single-product purchases.[6]

[5] Two or more firms introduce a new product in the same market (product x month) in less than 0.14% of markets.

[6] The estimation sample uses eight-digit HSN product codes and 12 months of transaction data. We focus on the top 20 single-product and top 5 multi-product bundles by transaction count. We restrict the analysis to bundles of at most two products. When a transaction could match multiple bundles, we prioritize by temporal proximity and assign each transaction to at most one bundle.

5.1 Demand and Marginal Cost

This section estimates three sets of parameters: the preference parameters $(\beta_d, \beta_r, \gamma)$ on distance, relationships, and bundling; the price coefficient α; and marginal costs $\{mc_{fht}\}$.

We estimate these parameters in three steps. The first step recovers preference parameters and mean utilities jointly using a nested fixed-point procedure. For a given candidate $(\beta_d, \beta_r, \gamma)$, we solve for the mean utilities that rationalize observed market shares, then evaluate the likelihood of observed purchase decisions. The second step decomposes the recovered mean utilities into price and unobserved quality components, yielding the price coefficient α. The third step inverts firms' first-order conditions to recover marginal costs.

5.1.1 Step 1: Preference Parameters $(\beta_d, \beta_r, \gamma)$

We estimate $\boldsymbol{\theta} = (\beta_d, \beta_r, \gamma)$ using a nested fixed-point procedure in the spirit of Berry, Levinsohn, and Pakes, 1995. For each candidate $\boldsymbol{\theta}$, an inner loop recovers the mean utilities that rationalize observed market shares; an outer loop then evaluates the likelihood of observed purchase decisions given these mean utilities.

The inner loop recovers mean utilities $\{\delta_{fht}\}$ by matching predicted to observed market shares. In standard demand estimation, shares are defined within a single market. Here, however, a variety (f, h) appears in multiple bundle markets—both single-product bundles $\{h\}$ and multi-product bundles containing h. We therefore aggregate across all bundle markets containing product h to compute each variety's share.

Formally, let s^{pred}_{fht} denote the predicted market share of variety (f, h) in period

t. This share sums over all bundle markets containing h, weighting each buyer's marginal purchase probability by the quantity that buyer demands:

$$s_{fht}^{\text{pred}} = \frac{\sum_{b \ni h} \sum_{i \in I_{bt}} S_{ifh}(\boldsymbol{\theta}, \boldsymbol{\delta}) \cdot q_{ih}}{\sum_{b \ni h} \sum_{i \in I_{bt}} q_{ih}}$$

The outer sum is over all bundles b containing product h; the inner sum is over all buyers i in each bundle market. For each buyer, S_{ifh} is the marginal probability of purchasing product h from firm f (Equation 7), which depends on the choice set available in that buyer's bundle market. The denominator is the total quantity demanded for product h across all bundle markets.

Given predicted shares, we recover mean utilities via the BLP contraction mapping. Let s_{fht}^{obs} denote the observed market share, computed as quantity sold divided by total market quantity. The contraction mapping iterates as follows:

$$\delta_{fht}^{(r+1)} = \delta_{fht}^{(r)} + \log s_{fht}^{\text{obs}} - \log s_{fht}^{\text{pred}}(\boldsymbol{\theta}, \boldsymbol{\delta}^{(r)})$$

until convergence. We normalize the mean utility of one variety per product-month to zero.

The outer loop searches over $\boldsymbol{\theta}$ to maximize the likelihood of observed purchase decisions. Since the inner loop recovers mean utilities as a function of $\boldsymbol{\theta}$, the likelihood is ultimately a function of $\boldsymbol{\theta}$ alone. Each buyer i in bundle market (b, t) chooses a seller combination $\mathbf{f}_{ibt}^*$; our demand model assigns probability $s_{i\mathbf{f}_{ibt}^*}$ to this choice. The log-likelihood sums over all buyers and markets:

$$\mathcal{L}(\boldsymbol{\theta}) = \sum_{b,t} \sum_{i \in I_{bt}} \log s_{i\mathbf{f}_{ibt}^*}(\boldsymbol{\theta}, \boldsymbol{\delta}(\boldsymbol{\theta}))$$

We maximize this likelihood to find the preference parameters that best explain observed purchase decisions.

5.1.2 Step 2: Price Coefficient

The mean utilities from Step 1 decompose into price and unobserved quality: $\delta_{fht} = \alpha \cdot p_{fht} + \xi_{fht}$.[7] In our baseline specification, we calibrate $\alpha = -10$.

5.1.3 Step 3: Marginal Costs

To recover marginal costs, we invert the first-order conditions of the pricing game (Equation 9). For each firm f in period t:

$$\mathbf{mc}_{ft} = \mathbf{p}_{ft} - \mathbf{\Omega}_{ft}^{-1}\mathbf{Q}_{ft}/\alpha$$

Prices and quantities are observed in the data; $\mathbf{\Omega}_{ft}$ depends on estimated parameters from Step 1, and α is the price coefficient from Step 2. This allows us to express marginal costs as a function of observables and estimated parameters.[8]

5.2 Fixed Costs

Recall that a firm introduces product h if expected gain in profits exceeds the fixed cost draw. Computing expected profit gain requires addressing two sources of uncertainty: the entrant does not know what quality and cost it will realize, and entry changes equilibrium prices.

We address both through simulation. For each potential entrant, we take 500 draws of (ξ_{fh}, mc_{fh}) from the empirical distribution of incumbents selling product

[7] We normalize prices within each HSN category by dividing by the median price across all seller-month observations. Correspondingly, quantity is recomputed as transaction value divided by normalized price, yielding a value-equivalent volume measure. This normalization ensures cross-category comparability.

[8] For products with small market shares, the inversion can produce implausible marginal costs—negative values or extreme markups. We address this by imputing marginal costs for such observations. We first compute the Lerner index $L_{fht} = (p_{fht} - mc_{fht})/p_{fht}$ for all observations. We then regress the logit-transformed Lerner index on two predictors: the within-HSN decile of firm f's sales value, and an indicator for whether h is the firm's main product. For observations with $mc \leq 0$ or $L \geq 0.95$, we impute marginal costs from the regression-predicted Lerner indices.

Table 5: *Estimation Results*

(a) *Estimated Parameters*

Parameter	Estimate
Demand Parameters	
Distance (log)	-0.689
Repeat Seller	9.123
Single-Seller Bonus	1.196
Fixed Cost Parameters	
μ	23.996
σ	2.820

(b) *Marginal Costs*

Statistic	Value
Mean MC	0.786
Median MC	0.871
% Imputed	15.529
N Observations	35,945

Notes. Left panel reports estimated demand and fixed cost parameters. Right panel reports summary statistics for recovered marginal costs.

h. For each draw, we add the entrant to the market, solve for post-entry Bertrand-Nash equilibrium prices, and compute the change in profit relative to the pre-entry equilibrium. Expected profit gain is the average across draws.[9]

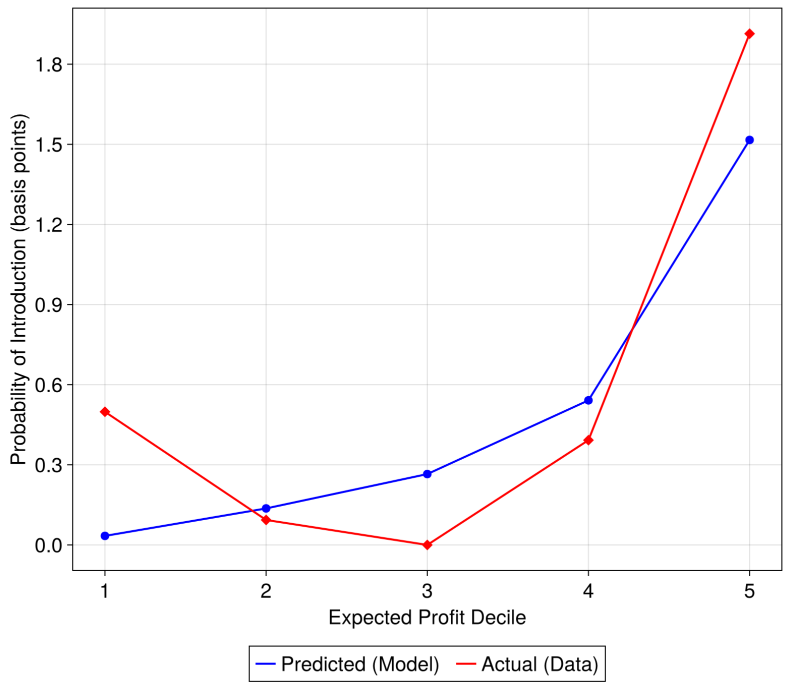
Given expected change in profit $E[\Delta\Pi_f(h)]$, we estimate the fixed cost distribution by maximum likelihood. Since fixed costs are lognormal, the probability of entry is:

$$\Pr\big(\text{entry}_{fh}\big) = \Phi\left(\frac{\log E[\Delta\Pi_f(h)] - \mu}{\sigma}\right)$$

where Φ is the standard normal CDF and (μ, σ) parameterize the lognormal distribution.

[9] Computing expected profits is costly because each potential entrant requires solving for equilibrium prices 500 times. To reduce computational burden, we employ choice-based sampling: since entry is rare, we include all observed entrants and a random sample of non-entrants, weighting by inverse sampling probabilities to correct for oversampling.

Figure 5: *Model Fit: Predicted vs. Actual Introduction Probabilities*



Notes. Predicted (model) and actual (data) product introduction probabilities by expected profit decile. Predicted probabilities computed using estimated fixed cost distribution.

5.3 Results

Table 5a reports the estimated demand parameters. Distance enters negatively, while the repeat-seller and single-seller bonuses are positive: buyers prefer nearby, familiar sellers and value bundling from a single source. Table 5b summarizes the recovered marginal costs. With prices normalized so the median equals one, the median marginal cost is 0.87. We impute marginal costs for 15.5 percent of observations following the procedure described above in footnote 8.

To assess model fit, Figure 5 plots product introduction probabilities by quintile of expected profit gain. Expected profit gain is the model-implied change in firm profits from introducing a product: post-entry profit minus pre-entry profit. The figure reveals two patterns that support our model estimates. First, actual introduction probabilities increase with expected profit gain. Firms enter where the demand model predicts higher returns, and since expected profits depend on the estimated preference parameters, this upward slope confirms that distance, relationships, and bundling capture economically meaningful variation in profitability. Second, predicted probabilities closely track actual probabilities across quintiles, suggesting that the estimated fixed cost distribution rationalizes observed entry rates.

6 Counterfactual Analysis

The descriptive evidence in Section 3 showed that (1) firms rely heavily on existing clients when selling new products, and (2) product introductions increase sales of complementary products already in the firm's portfolio. The structural model formalized these patterns through two parameters: the repeat-seller bonus β_r, which captures the advantage of established relationships, and the single-seller bonus γ, which captures the value of sourcing from a single supplier.

We now use the estimated model to quantify how much these mechanisms matter for product introduction. We consider two counterfactual scenarios. The first reduces β_r by 25 percent, representing an environment where existing client relationships confer smaller advantages. The second reduces γ by 50 percent, representing an environment where single-sourcing confers smaller advantages. For each scenario, we compute product introduction probabilities and compare them to the baseline.

6.1 Reducing Repeat-Seller Bonus

In our first counterfactual exercise, we ask how product introduction would change if the preference β_r for existing supplier relationships were weaker. To do so, we reduce β_r by 25 percent. The parameter β_r captures several forces that we do not distinguish between: a pure preference for purchasing from familiar sellers, the cost of searching for new suppliers, the difficulty of evaluating unfamiliar sellers, or the risk of contracting with unknown parties. This counterfactual simulates an environment where some combination of these forces is weaker, perhaps through online platforms that reduce search costs, quality certification that eases evaluation, or repeated interactions that build trust more quickly.

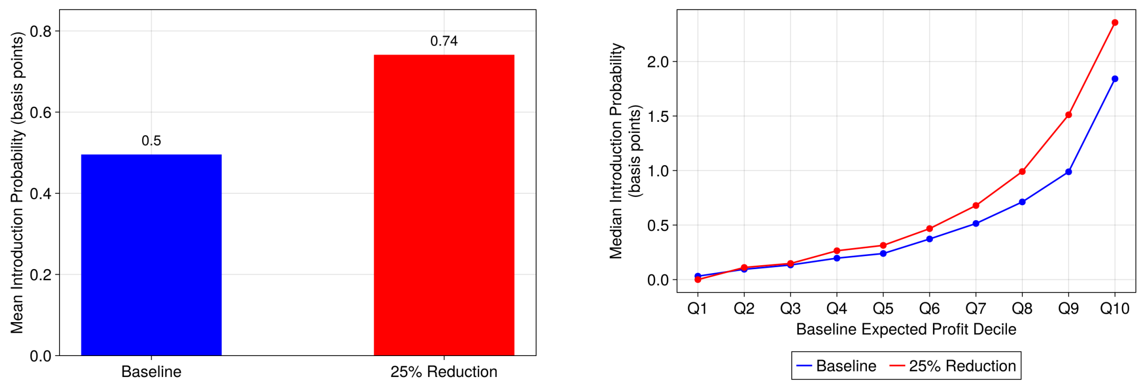
We find that reducing β_r increases product introductions. Figure 6 shows introduction probabilities under baseline and counterfactual parameters. The left panel compares mean introduction probability in aggregate. The mean rises from 0.48 to 0.72 basis points, a 50 percent increase. The right panel shows that this increase occurs across the distribution of baseline expected profit.[10]

Why does reducing β_r encourage entry? The repeat-seller preference gives firms a competitive advantage with buyers they have previously transacted with. Incumbents have accumulated relationships over time, which insulates them from competition. Potential entrants may have transacted with some of the same buyers in other product markets, but they have fewer established relationships in the market they are entering. Reducing β_r weakens this advantage, so more entrants find entry profitable.

We explore this asymmetry directly. For each firm-product-month, we compute *revenue dependence*, defined as the share of expected revenue derived from repeat

[10]For each potential introduction (firm-product-month), we compute expected profit by simulating entry. The firm draws unobserved characteristics for its new product (quality and marginal cost) and we compute the resulting profit. We average over 20 such draws to obtain expected profit. The right panel bins potential entrants into deciles by this measure and plots mean introduction probability for each decile.

Figure 6: *Impact on product entry from a 25% reduction in repeat-seller preference*



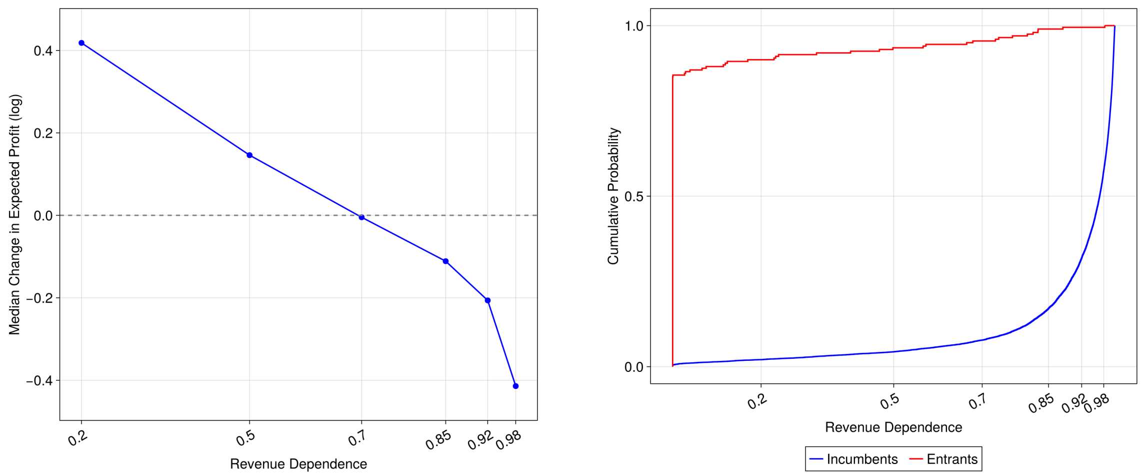
Notes. This figure shows how product introduction probabilities change when the repeat-seller bonus is reduced by 25 percent. The left panel compares aggregate introduction probability under baseline and counterfactual parameters. The right panel plots introduction probability by decile of baseline expected profit, showing that the increase occurs throughout the distribution. Weights account for choice-based sampling.

buyers.[11] This measure captures how much the firm's expected sales depend on buyers with whom it has an established relationship. The left panel in Figure 7 plots changes in profits against revenue dependence. Incumbents with revenue dependence below 0.7 gain from the counterfactual, while those above 0.7 lose out. This threshold marks where the policy switches from helping to hurting.

The right panel shows why this matters for entry. It compares the distributions of revenue dependence for incumbents and potential entrants, computed under baseline parameters for the same product-month combinations. Two patterns emerge. First, the CDF for entrants lies to the left of the CDF for incumbents, confirming that entrants have systematically lower revenue dependence. Second, most incumbents fall below the 0.7 threshold. Together, these facts imply that reducing β_r primarily benefits entrants, who compete at a disadvantage when the repeat-seller preference

[11] We compute revenue dependence separately for each firm-product-month. A buyer is classified as a repeat buyer of firm f in month t if their firm purchased any product from f prior to month t. For each buyer in the market, we use the estimated demand model to predict their probability of choosing firm f, weight by expected quantity, and sum. Revenue dependence is the ratio of this sum for repeat buyers to the sum across all buyers.

Figure 7: *Revenue dependence and change in profits from lower repeat-seller preference*



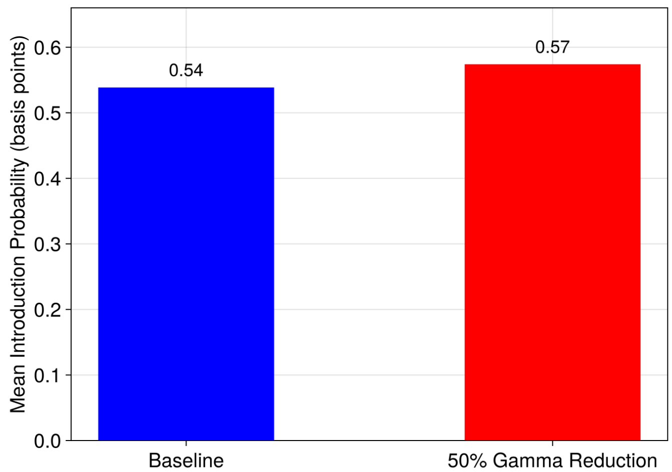
Notes. This figure illustrates heterogeneity in the effects of reducing the repeat-seller bonus. Revenue dependence measures the share of a firm's expected revenue from buyers with whom it has a prior relationship. The left panel shows that incumbents with high revenue dependence lose from the counterfactual, while those with low revenue dependence gain. The right panel compares the distribution of revenue dependence for incumbents and potential entrants in the same markets, showing that entrants are systematically less reliant on repeat buyers.

is high.

6.2 Reducing Single-Seller Bonus

The single-seller bonus γ rewards multi-product firms with broad portfolios because a high γ implies that buyers prefer sourcing multiple products from the same seller, and these multi-product firms can offer that convenience. This preference for single-sourcing may be driven by transaction costs, simplified logistics, or the overhead of managing multiple supplier relationships. We do not separately identify these channels; γ captures their combined effect. In our second counterfactual exercise, we reduce γ by 50 percent to simulate an environment where the ability to single-source confers smaller benefits for consumers. This could arise from logistics platforms that

Figure 8: *Impact on product entry from a 50% reduction in single-seller bonus*



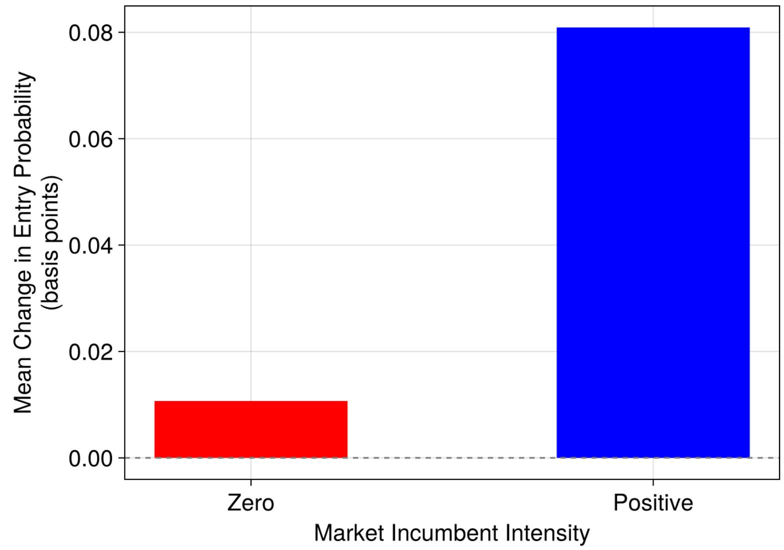
Notes. This figure shows aggregate introduction probability under baseline parameters and after reducing the single-seller sourcing bonus by 50 percent. Weights account for choice-based sampling.

make multi-supplier coordination easier, or from procurement systems that reduce the cost of managing multiple vendors.

When γ falls, the benefit of single-sourcing weakens, and purchase decisions depend more on the individual merits of each product. Incumbents who benefitted from offering complete bundles see that advantage erode. This creates entry opportunities. Figure 8 shows that the mean product introduction probability rises from 0.42 to 0.44 basis points under the counterfactual.

This aggregate effect is modest compared to the repeat-seller counterfactual, but the average masks substantial heterogeneity across markets. To show this, we compute a measure of single-sourcing intensity: the fraction of a firm's expected revenue from multi-product bundles where the firm supplies all products. Markets differ in how much their incumbents rely on this single-seller advantage. Figure 9 plots the change in entry probability against market incumbent intensity, defined as the revenue-weighted average single-sourcing intensity among incumbents in each

Figure 9: *Entry probability gains by incumbent single-seller intensity*



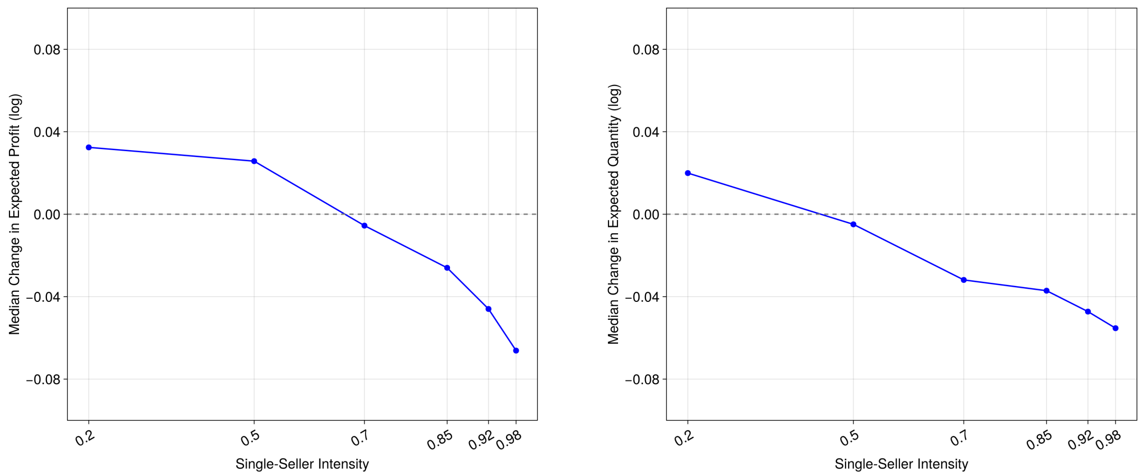
Notes. This figure shows how entry probability gains vary with incumbent reliance on single-sourcing. Market incumbent intensity is the revenue-weighted average single-seller intensity among incumbents in each product-month market, where single-seller intensity measures the share of revenue from multi-product bundles supplied entirely by one firm. Entry gains are larger in markets where incumbents rely more heavily on single-sourcing.

product-month market.[12] Entry probability gains are larger in markets where incumbents rely more heavily on single-seller bundling. In these markets, reducing γ erodes more of the advantage enjoyed by incumbents and creates greater opportunities for entrants.

Figure 10 examines how the counterfactual affects incumbents. Those who rely most heavily on single-seller bundling lose the most when the bundling advantage weakens. The left panel confirms this pattern: incumbents with high single-seller intensity experience larger profit declines, while those with low intensity are largely unaffected. The right panel shows a similar pattern for expected quantity.

[12] For each firm, we compute expected revenue from all bundles containing two or more products, using the estimated demand model. Single-seller intensity is the share of this revenue from bundles where the firm supplies all products. Market incumbent intensity is the revenue-weighted average across incumbents in each product-month market.

Figure 10: *Single-seller intensity and change in profits from lower bundling bonus*



Notes. This figure shows how reducing the single-seller bonus affects incumbents. Single-seller intensity measures the share of a firm's revenue from multi-product bundles where the firm supplies all products. The left panel shows that incumbents with high single-seller intensity experience larger profit declines. The right panel shows a similar pattern for expected quantity.

6.3 Discussion

The above two counterfactuals quantify some of the demand-side mechanisms that shape product scope. Buyers prefer purchasing from sellers they have previously transacted with and value the convenience of sourcing multiple products from a single supplier. These preferences favor firms with established buyer relationships and broad product portfolios, making it difficult for firms without these assets to introduce new products. Reducing repeat-seller preferences or single-sourcing preferences weakens the position of established incumbents and increases entry rates for potential new products. However, the impact of these factors differs substantially in our setting. A 25% reduction in the repeat-seller bonus and a 50% reduction in the single-seller bonus produce qualitatively similar effects on entrants. Thus, the preference for prior sellers appears to be the more consequential mechanism.

7 Conclusion

This paper studies how demand-side factors shape firms' decisions to expand their product portfolios. Using administrative transaction records from India's e-way bill system, we document three empirical patterns: firms rely on existing clients for over half of new product sales; introducing a product increases sales of complementary products already in the firm's portfolio; and market access to existing clients, particularly for complements, predicts which products firms introduce. We develop a structural model that formalizes these patterns through two mechanisms: a repeat-seller preference that captures the advantage of established relationships, and a single-seller bonus that captures buyers' preference for sourcing multiple products from the same firm. The estimated model fits observed introduction probabilities well and allows us to quantify the role of each mechanism.

Counterfactual exercises reveal that both mechanisms affect product introduction, though to different degrees. Reducing the repeat-seller preference by 25 percent increases mean introduction probability by 50 percent. This is because incumbents have accumulated buyer relationships over time; potential entrants have not. Thus, when the repeat-seller preference falls, this incumbency advantage weakens, and product introduction probability rises. Reducing the single-seller bundling bonus by 50 percent produces a smaller aggregate effect, though gains concentrate in markets where incumbents rely heavily on bundling. These results show that buyer preferences over supplier relationships and multi-product sourcing create barriers to product introduction that favor established firms with broad portfolios and deep customer bases.

These findings also speak to firm growth in developing economies. They suggest that the constraint on product variety may not just be a lack of productive capacity or entrepreneurial talent, but rather the difficulty of breaking into established buyer-seller networks. In settings where formal contracting is costly and information about suppliers is limited, buyer-seller relationships play a large role in shaping which firms can expand. Our results show that firms without established customer networks face barriers when introducing new products. Understanding these demand-side forces complements supply-side explanations and may help explain why product scope remains concentrated among a small number of firms.

References

Bernard, A. B., Blanchard, E. J., Van Beveren, I., & Vandenbussche, H. (2019). Carry-along trade. *The Review of Economic Studies*, *86*(2), 526–563.

Bernard, A. B., Redding, S. J., & Schott, P. K. (2010). Multiple-product firms and product switching. *American Economic Review*, *100*(1), 70–97.

Berry, S., Levinsohn, J., & Pakes, A. (1995). Automobile Prices in Market Equilibrium. *Econometrica : journal of the Econometric Society*, *63*(4), 841.

Boehm, J., Dhingra, S., & Morrow, J. (2022). The comparative advantage of firms. *Journal of Political Economy*, *130*(12), 3025–3100.

Callaway, B., & Sant'Anna, P. H. (2021). Difference-in-Differences with multiple time periods. *Journal of Econometrics*, *225*(2), 200–230.

Eaton, J., Eslava, M., Jinkins, D., Krizan, C. J., & Tybout, J. R. (2025). A search and learning model of export dynamics. *Journal of International Economics*, *157*(100).

Eckel, C., & Neary, J. P. (2010). Multi-product firms and flexible manufacturing in the global economy. *Review of Economic Studies*, *77*(1), 188–217.

Flagge, M., & Chaurey, R. (2014). *Firm-product linkages and the evolution of product scope.*

Foster, L., Haltiwanger, J., & Syverson, C. (2016). The slow growth of new plants: Learning about demand? *Economica*, *83*(329), 91–129.

Gentzkow, M. (2007). Valuing New Goods in a Model with Complementarity: Online Newspapers. *The American Economic Review*, *97*(3), 713–744.

Goldberg, P. K., Khandelwal, A. K., Pavcnik, N., & Topalova, P. (2010). Multiproduct firms and product turnover in the developing world: Evidence from india. *Review of Economics and Statistics*, *92*(4), 1042–1049.

Goldberg, P. K., & Reed, T. (2023). Presidential address: Demand-side constraints in development: The role of market size, trade, and (in)equality. *Econometrica*, *91*(6), 1915–1950.

Gourio, F., & Rudanko, L. (2014). Customer capital. *The Review of Economic Studies*, *81*(3), 1102–1136.

Leone, F., Macchiavello, R., & Reed, T. (2025). The high and falling price of cement in africa. *American Economic Journal: Applied Economics*, *17*(2), 1–40.

Mayer, T., Melitz, M. J., & Ottaviano, G. I. P. (2014). Market size, competition, and the product mix of exporters. *American Economic Review*, *104*(2), 495–536.

Monarch, R. (2022). It's not you, it's me: Prices, quality, and switching in u.s.-china trade relationships. *The Review of Economics and Statistics*, *104*(5), 909–928.

Press Information Bureau. (2021, April 1). *7.12 crores e-way bills generated in march 2021, highest ever since launch of e-way bill system*. Government of India, Ministry of Finance.

APPENDIX

Customer-Driven Product Scope

Contents

A Data Construction Details

This appendix describes the e-way bill system, the raw data, and our sample construction.

A.1 E-Way Bill System Details

Each e-way bill consists of two parts. Part A records transaction details: the GSTIN and PIN code of buyer and seller, the invoice number, date, and value, and the HSN code identifying the product. Part B captures transportation details: the vehicle number, mode of transport, and distance traveled. A separate e-way bill is required for each invoice, so a single truck carrying goods covered by multiple invoices generates multiple e-way bills.

The system includes enforcement mechanisms at highway checkpoints. GST officers can intercept any vehicle and verify that its cargo matches the e-way bill documentation. Failure to produce a valid e-way bill results in penalties of Rs. 10,000 or the tax evaded, whichever is higher, and the goods and vehicle may be detained. In the first three years alone (April 2018 to March 2021), over 1.8 billion e-way bills were generated (Press Information Bureau, 2021). This reflects widespread adoption among formal sector firms.

A.2 Raw Data Characteristics

The raw data span 24 monthly files containing approximately 439 million transaction records for the period April 2018 through March 2020. Each record corresponds to a single e-way bill and includes:

- GSTIN and PIN code of consignor (seller)
- GSTIN and PIN code of consignee (buyer)
- Invoice number, date, and value
- HSN code and product description
- Quantity and unit of measurement
- Transport mode and vehicle number
- Distance traveled (kilometers)

A.3 HSN Product Classification

The Harmonized System of Nomenclature (HSN) is an international standard developed by the World Customs Organization. HSN codes identify traded products at varying levels of granularity: the first two digits indicate the chapter (broad category), the next two the heading, and subsequent digits provide finer detail.

Indian GST requires businesses with annual turnover above Rs. 5 crore (600K USD) to report at least four HSN digits. The most common categories include cement (HSN 2523), pharmaceuticals (HSN 3004), auto parts (HSN 8708), and ceramic tiles (HSN 6907).

A.4 Sample Construction

We apply the following filters to the raw data:

1. **Restrict time period:** Exclude transactions dated on or after April 1, 2020, to avoid COVID-19 disruptions.

2. **Remove intra-firm transactions:** Remove shipments between establishments of the same firm (where buyer and seller share a tax identifier).
3. **Retain valid HSN codes:** Drop records with missing or invalid HSN codes. Codes shorter than four digits after standardization account for 5.4% of transactions but only 0.4% of total value.
4. **Clean sales and quantities:** Exclude records with zero or negative quantities or values.
5. Winsorize transaction values at the 99.999$^{\text{th}}$ percentile.

After these filters, the cleaned data contain 53.8 million transactions involving 242,474 unique selling firms and 649,503 unique buying firms.

Additional restrictions for main analysis. We impose two further restrictions to focus on established firms with regular activity:

1. **Firm activity:** Retain only firms active for at least 13 months over the sample period. This filter reduces the seller count from 242,474 to 39,964 (16.5% retained).
2. **Product activity:** Retain only seller-product pairs with more than 50 transactions over the sample period. This filter removes sporadic or occasional product lines.

Table A.1 summarizes sample attrition at each step.

A.5 Reduced-Form Analysis Sample

For the reduced-form event study analysis in Section 3, we aggregate products to the 4-digit HSN level.

Table A.1: *Sample Attrition*

Step	Transactions	Sellers	% Retained
Raw data (Apr 2018–Mar 2020)	439,000,000	—	—
After all cleaning filters	53,800,000	242,474	12.3%
After firm activity filter (>12 months)	46,100,000	39,964	85.7%
After product activity filter (>50 txns)	42,600,000	39,344	92.4%

Notes. This table shows the number of transactions and unique sellers at each stage of sample construction. The cleaning filters include: date restriction (pre-COVID), intra-firm removal, HSN validation, value/quantity cleaning, and outlier winsorization. Percentages in the final column show retention relative to the previous step.

Table A.2 presents summary statistics for this sample. The resulting panel contains 1.70 million seller-product-month observations across 39,964 firms and 1,029 product categories. The median firm-product-month observation records Rs. 4.60 lakh in sales across 7 transactions to 4 distinct buyers. The mean is substantially higher (Rs. 49 lakh per firm-product-month), reflecting a right-skewed distribution driven by differences in firm size.

Under the product introduction definition described in Section 2, this sample yields 2,368 product introductions across 1,300 firms, representing 2.7% of all firm-product pairs. The typical introduction occurs 9 months after the firm's entry into the data. Most firms that introduce products add just one or two new categories (median 1, mean 1.8), though a few expand into dozens.

Table A.2: *Sample Summary Statistics (4-Digit HSN)*

Variable	Value
Total Observations	42,602,732
Unique Seller Firms	39,344
Unique Buyer Firms	120,633
Unique Products (HSN4)	1,116
Unique E-way Bills	24,963,653
Total Transaction Value (Billion USD)	97.670
Median Transaction Value (Hundred USD)	8.160
Median Transactions per Seller	175
Median Products per Seller	1

Notes. This table presents descriptive statistics for the reduced-form analysis sample, aggregated to the 4-digit HSN level. The sample includes only firms that were active for at least 13 months and excludes within-firm shipments. Transaction values are in Indian Rupees (1 USD ≈ 85 INR).

B Stylized Model

This appendix develops a stylized model of product introduction that generates testable predictions for our reduced-form analysis. The model incorporates two key features of demand. First, existing client relationships provide additional utility from repeat purchases. Second, buyers receive additional utility from purchasing products together, which leads to complementarity between products. We use this framework to derive predictions about what drives product introduction and what happens after introduction.

B.1 Setup

Consider a market with two products, A and B. A firm f currently sells product A and is deciding whether to introduce product B. Buyers, indexed by i, can purchase products individually or as a bundle. After introduction, buyers choose among four options: purchase only A, purchase only B, purchase both as a bundle, or purchase neither. Each buyer demands q_i units conditional on purchase. Prices are exogenous.

Buyer i's utility from purchasing product A, product B, or the bundle AB is:

$$\begin{aligned}
U_A^i &= \mu_f + \mu_A + \beta_r \cdot \mathbf{1}[\text{repeat}_{if}] + \beta_d \cdot \log d_{if} - \alpha \cdot p_A + \nu_A^i \\
U_B^i &= \mu_f + \mu_B + \beta_r \cdot \mathbf{1}[\text{repeat}_{if}] + \beta_d \cdot \log d_{if} - \alpha \cdot p_B + \nu_B^i \\
U_{AB}^i &= U_A^i + U_B^i + \Gamma \\
U_0^i &= 0
\end{aligned}$$

Utility from purchasing product A or B depends on firm and product characteristics μ_f, μ_A, μ_B, whether the buyer has an existing relationship with the firm, distance, and prices. The terms ν_A^i and ν_B^i capture idiosyncratic buyer preferences for each product.

The third equation defines utility when both products are purchased together as a bundle. Buyers receive utility from each product, $U_A^i + U_B^i$, plus an additional utility Γ. When $\Gamma > 0$, products are complements: buyers value the bundle more than the sum of its parts. This formulation follows Gentzkow (2007). The last equation normalizes the outside option to zero.

B.2 Choice Regions

Each buyer chooses the option with highest utility. Figure B.1 illustrates the choice regions in (U_A, U_B) space. When $\Gamma = 0$, buyer i purchases both products if and only if $U_A^i > 0$ and $U_B^i > 0$. When $\Gamma > 0$, the bundle region expands. A buyer may now purchase the bundle even when one product has negative standalone utility, provided the combined utility $U_A^i + U_B^i + \Gamma$ exceeds zero.

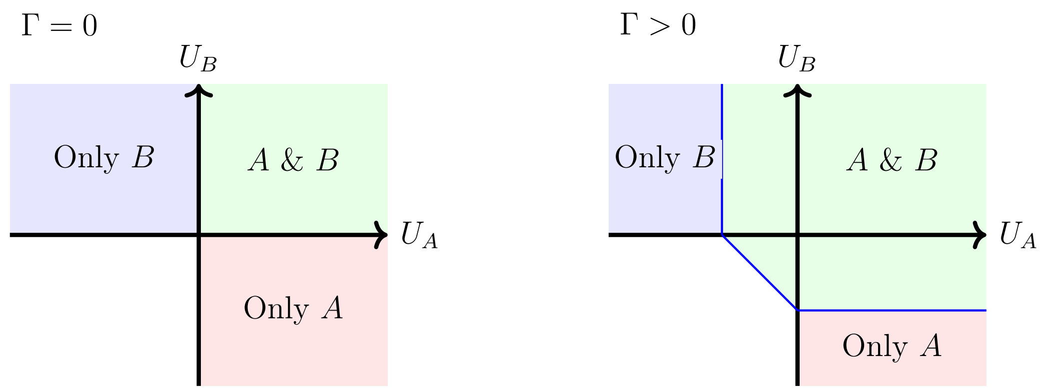
This expansion has implications for product introduction. Buyers with $U_A^i \in (-\Gamma, 0)$ would not purchase A alone, but may purchase it as part of the bundle. When $\Gamma > 0$, introducing B expands demand for A by converting these marginal non-buyers into bundle purchasers. This makes introducing B more profitable than when $\Gamma = 0$, and firms are more likely to introduce products that complement their existing offerings.

B.3 Predictions

We now derive testable predictions from this framework.

Prediction 1: Product introduction is driven by market access, with different roles for existing clients and new buyers. Consider first the case where $\Gamma = 0$. A firm is more likely to introduce a product when located near buyers with high demand for

Figure B.1: *Choice Regions Under Different Complementarity*



(a) *No complementarity* **(b)** *Positive complementarity*

Notes. Choice regions in utility space. When $\Gamma = 0$, buyers purchase both products if and only if $U_A > 0$ and $U_B > 0$. When $\Gamma > 0$, the bundle region expands: buyers may purchase both products even when one has negative standalone utility, provided the combined utility exceeds the threshold.

that product. Expected sales of the new product B are:

$$Q_B = \int q_i \cdot \mathbf{1}\{U_B^i > 0\}\, dF(i)$$

With simplifying distributional assumptions, this expression yields:

$$\log(\text{Sales of } B) \approx \lambda_f + \lambda_B + \log\left(\theta^e \sum_{i \in \mathcal{E}} \frac{q_i}{d_{if}^{|\beta_d|}} + \theta^n \sum_{i \in \mathcal{N}} \frac{q_i}{d_{if}^{|\beta_d|}}\right)$$

where $\mathcal{E}$ denotes existing clients and $\mathcal{N}$ denotes new buyers. The summation terms capture distance-weighted demand potential, which we call *market access.* Firms with high-demand buyers nearby—high q_i and low d_{if}—have higher market access and are more likely to introduce. Existing clients and new buyers contribute differently to expected sales. If existing relationships provide additional utility, we expect $\theta^e > \theta^n$: a given level of demand from existing clients translates into higher sales than the same demand from new buyers.

Prediction 2: Product introduction increases sales of existing products, and com-

plementarity drives introduction. When $\Gamma > 0$, the increase in sales from introducing product B has three components:

$$\Delta Q = \int q_i \cdot \Big[\underbrace{\mathbf{1}\{U_B^i > 0\}}_{\text{sales of } B} + \underbrace{\mathbf{1}\{0 > U_A^i > -\Gamma\} \cdot \mathbf{1}\{U_A^i + U_B^i > -\Gamma\}}_{\text{higher bundled sales of } A} + \underbrace{\mathbf{1}\{0 > U_B^i > -\Gamma\} \cdot \mathbf{1}\{U_A^i + U_B^i > -\Gamma\}}_{\text{higher bundled sales of } B} \Big] \, dF(i)$$

The first term captures direct sales of B, as in Prediction 1. The second term captures buyers with $U_A^i \in (-\Gamma, 0)$ who would not purchase A alone but now purchase it as part of the bundle. The third term is symmetric for B. Introducing a new product therefore increases sales of existing products, and firms are more likely to introduce products that complement their existing offerings.

Prediction 3: The effect on prices of existing products is ambiguous. We explore this prediction through numerical simulation. We find a non-monotonic relationship between Γ and the price of existing products: prices can either increase or decrease after product introduction, depending on the degree of complementarity. This ambiguity arises from two opposing forces. First, lowering prices becomes more attractive. Before introduction, a price cut on A only increased sales of A. After introduction, it increases sales of both A and B via the bundle. Second, raising prices also becomes more attractive. Buyers who were previously marginal now receive additional utility from the bundle, making them inframarginal. The firm can raise prices without losing them, allowing it to extract surplus.

C Derivation of First-Order Conditions

This appendix derives the first-order conditions for pricing and expresses them in matrix form. The resulting system allows us to recover marginal costs from observed prices, quantities, and estimated demand parameters. Following the main text, we suppress time subscripts throughout. Recall that markets are defined by a bundle $b = \{h_1, \ldots, h_K\}$, and buyers choose among seller combinations $\mathbf{f} = (f_1, \ldots, f_K)$ with choice probabilities $s_{i\mathbf{f}} = \exp(V_{i\mathbf{f}}) / \sum_{\mathbf{f}'} \exp(V_{i\mathbf{f}'})$. Mean utility is $\delta_{fh} = \alpha \cdot p_{fh} + \xi_{fh}$, where $\alpha < 0$ is the price coefficient.

C.1 Choice Probability Derivatives

We first derive how prices affect choice probabilities. The key result expresses the derivative $\partial s_{i\mathbf{f}} / \partial p_{fh}$ in terms of the marginal probability S_{ifh} that buyer i purchases product h from firm f. This decomposition separates own-price effects from cross-price substitution.

Let $D = \sum_{\mathbf{f}'} \exp(V_{i\mathbf{f}'})$ denote the denominator, so that $s_{i\mathbf{f}} = \exp(V_{i\mathbf{f}})/D$. The derivative is:

$$\begin{aligned}
\frac{\partial s_{i\mathbf{f}}}{\partial p_{fh}} &= \frac{D \cdot \frac{\partial \exp(V_{i\mathbf{f}})}{\partial p_{fh}} - \exp(V_{i\mathbf{f}}) \cdot \frac{\partial D}{\partial p_{fh}}}{D^2} \\
&= \frac{\exp(V_{i\mathbf{f}})}{D} \cdot \frac{\partial V_{i\mathbf{f}}}{\partial p_{fh}} - \frac{\exp(V_{i\mathbf{f}})}{D} \sum_{\mathbf{f}'} \frac{\exp(V_{i\mathbf{f}'})}{D} \cdot \frac{\partial V_{i\mathbf{f}'}}{\partial p_{fh}} \\
&= s_{i\mathbf{f}} \left(\frac{\partial V_{i\mathbf{f}}}{\partial p_{fh}} - \sum_{\mathbf{f}'} s_{i\mathbf{f}'} \cdot \frac{\partial V_{i\mathbf{f}'}}{\partial p_{fh}} \right)
\end{aligned}$$

From the utility specification, price p_{fh} only affects seller combinations where firm f supplies product h:

$$\frac{\partial V_{i\mathbf{f}}}{\partial p_{fh}} = \mathbf{1}[f_h = f] \cdot \alpha \cdot w_{ih}$$

where w_{ih} is the quantity weight. Substituting yields:

$$\begin{aligned}\frac{\partial s_{i\mathbf{f}}}{\partial p_{fh}} &= \alpha \cdot w_{ih} \cdot s_{i\mathbf{f}} \cdot \left(\mathbf{1}[f_h = f] - \sum_{\mathbf{f}'} s_{i\mathbf{f}'} \cdot \mathbf{1}[f'_h = f] \right) \\ &= \alpha \cdot w_{ih} \cdot s_{i\mathbf{f}} \cdot (\mathbf{1}[f_h = f] - S_{ifh})\end{aligned}$$

where $S_{ifh} = \sum_{\mathbf{f}'} s_{i\mathbf{f}'} \cdot \mathbf{1}[f'_h = f]$ is the marginal probability that buyer i purchases product h from firm f.

The term $(\mathbf{1}[f_h = f] - S_{ifh})$ determines the sign of the share response. Consider a price increase for variety (f, h). Seller combinations that include firm f for product h become less attractive. For these combinations, the indicator equals one, so the term is positive. Since $\alpha < 0$, their shares decrease. The lost demand must go somewhere: seller combinations that do *not* include firm f gain share. For these combinations, the indicator is zero, so the term is negative. Since $\alpha < 0$, their shares increase.

C.2 First-Order Condition

We now derive the firm's optimality condition for pricing. Firm f's profit and its derivative with respect to p_{fh} are:

$$\Pi_f = \sum_{b \in \mathcal{B}} \sum_{i \in I_b} \sum_{\mathbf{f}' \in \mathcal{C}_b} s_{i\mathbf{f}'} \left(\sum_{k=1}^{K} \mathbf{1}[f'_k = f] \cdot (p_{fh_k} - mc_{fh_k}) \cdot q_{i,h_k} \right)$$

$$\frac{\partial \Pi_f}{\partial p_{fh}} = \sum_{b \ni h} \sum_{i} \left[\underbrace{\sum_{\mathbf{f}'} s_{i\mathbf{f}'} \cdot \mathbf{1}[f'_h = f] \cdot q_{ih}}_{\text{Direct Revenue Effect}} + \underbrace{\sum_{\mathbf{f}'} \frac{\partial s_{i\mathbf{f}'}}{\partial p_{fh}} \left(\sum_{k \in b} \mathbf{1}[f'_k = f](p_{fk} - mc_{fk}) q_{ik} \right)}_{\text{Demand Diversion Effect}} \right]$$

The direct revenue effect captures the additional revenue from selling at a higher price, as each unit sold contributes more to profit. The demand diversion effect captures the cost of losing sales as buyers substitute away. Raising the price of product

h reduces demand for seller combinations that include the firm for h, lowering sales volume. For multi-product firms, there is an additional effect. Buyers who substitute away also stop purchasing the firm's other products in the same bundle. The inner sum over $k \in b$ captures the margin on each product the firm sells. This is multi-product pricing internalization, where the firm accounts for the impact of pricing h on its other products.

The direct revenue effect equals $Q_{fh} = \sum_{b \ni h} \sum_i S_{ifh} \cdot q_{ih}$, the total quantity of product h sold by firm f. Substituting the share derivative from above and setting $\partial \Pi_f / \partial p_{fh} = 0$ yields the first-order condition:

$$Q_{fh} + \sum_{b \ni h} \sum_i \alpha w_{ih} \sum_{\mathbf{f}'} s_{i\mathbf{f}'} (\mathbf{1}[f'_h = f] - S_{ifh}) \left(\sum_{k \in b} \mathbf{1}[f'_k = f](p_{fk} - mc_{fk}) q_{ik} \right) = 0$$

C.3 Deriving the Ω-Matrix System

The goal is to express the FOC in matrix form, where a single matrix Ω captures how each product's margin contributes to the optimality condition for every other product in the firm's portfolio. We proceed by reordering the summations to isolate margin terms.

Multiplying by $-1/\alpha$ and moving Q_{fh} to the right-hand side gives

$$\sum_{b\ni h}\sum_{i} w_{ih}\sum_{\mathbf{f}'} s_{i\mathbf{f}'}(S_{ifh}-\mathbf{1}[f'_h=f])\left(\sum_{k\in b}\mathbf{1}[f'_k=f](p_{fk}-mc_{fk})q_{ik}\right)=\frac{Q_{fh}}{\alpha}$$
$$\sum_{b\ni h}\sum_{i}\sum_{k\in b}(p_{fk}-mc_{fk})\left(w_{ih}\cdot q_{ik}\sum_{\mathbf{f}'} s_{i\mathbf{f}'}(S_{ifh}-\mathbf{1}[f'_h=f])\mathbf{1}[f'_k=f]\right)=\frac{Q_{fh}}{\alpha}$$
$$\sum_{b\ni h}\sum_{k\in b}(p_{fk}-mc_{fk})\left(\sum_{i} w_{ih}\cdot q_{ik}\sum_{\mathbf{f}'} s_{i\mathbf{f}'}(S_{ifh}-\mathbf{1}[f'_h=f])\mathbf{1}[f'_k=f]\right)=\frac{Q_{fh}}{\alpha}$$
$$\sum_{k\in\mathcal{H}_f}(p_{fk}-mc_{fk})\underbrace{\left(\sum_{b\supseteq\{h,k\}}\sum_{i\in I_b} w_{ih}\cdot q_{ik}\sum_{\mathbf{f}'} s_{i\mathbf{f}'}(S_{ifh}-\mathbf{1}[f'_h=f])\mathbf{1}[f'_k=f]\right)}_{\Omega_{hk}}=\frac{Q_{fh}}{\alpha}$$

The first line rearranges the FOC. The second swaps $\mathbf{f}'$ and k to bring the margin outside. The third swaps i and k. The fourth swaps bundles and products, iterating over all k in $\mathcal{H}_f$ and bundles containing both h and k. The underbrace defines the matrix element Ω_{hk}, and the FOC becomes

$$\sum_{k\in\mathcal{H}_f}\Omega_{hk}(p_{fk}-mc_{fk})=Q_{fh}/\alpha$$

This holds for every product h in firm f's portfolio. In matrix notation,

$$\mathbf{\Omega}(\mathbf{p}-\mathbf{mc})=\mathbf{Q}/\alpha$$

which matches Equation 9 in the main text.

To illustrate, consider a firm selling products 1 and 2:

$$\Omega_{11}(p_1-mc_1)+\Omega_{12}(p_2-mc_2)=Q_1/\alpha$$
$$\Omega_{21}(p_1-mc_1)+\Omega_{22}(p_2-mc_2)=Q_2/\alpha$$

Consider the first equation. The term Ω_{11} captures the own-price effect—how raising p_1 affects demand for product 1. The term Ω_{12} captures the cross-price effect—how raising p_1 affects demand for product 2 through bundle substitution.