Multiproduct Exporters: Learning and Knowing.

Yelena Sheveleva, Kala Krishna
The Pennsylvania State University

November 13, 2013

Abstract

In this paper we develop and estimate a structural model of multiproduct exporters based on three empirical regularities documented using data on Chinese exporters. These regularities are as follows: (1) multi-product exporters introduce their best-selling products early; (2) more than 40% of the new products introduced by incumbent exporters are dropped due to low sales within the first year; (3) for a firm the probability of introducing a new product is positively related to the survival and success of the earlier products.

The first regularity is consistent with the idea of core competence on the demand side. The second suggests that both incumbents and new exporters face uncertainty when they introduce new products. The third is consistent with firms learning about their potential in an export market, i.e., their brand effect, as they introduce new products. We develop a model which incorporates all of these features and we estimate it structurally using data on Chinese exporters to the U.S. in the plastics industry.

First, we find that known demand shocks play an important role in whether producers enter the export market or not. Second, we find that it is important to include uncertainty about the brand effect in order to account for large attrition among new exporters. When we let firms know their brand effect precisely, only those with sufficiently high brand effects enter, and then the model cannot replicate large attrition of new products among exporters. Third, we find that while firms act consistently with learning about their brand effect, the uncertainty that firms face in conjunction with introducing new products looms large and limits the extent to which learning affects incentives of firms to add new products. Our counterfactuals show that the distribution of products among the high brand effect firms only marginally first order stochastically dominates the distribution for low brand effect firms.

Using our model we revisit the question of trade policy in the multiproduct firm setting. We simulate a decrease in the cost of introducing new products for firms. Our simulations suggest that in the presence of economies of scope and even moderate learning effects, decreasing costs of introducing subsequent products can make a significant contribution to increasing trade flows.

1 Introduction

In recent years trade economists have gained access to narrow classifications of firms' exports. This has spurred their interest in the role of multiproduct exporters. It has been documented for a number of countries that multiproduct firms are important players in international trade. While small in number they account for a large share of domestic production and international trade. This preponderance has led to the conjecture that the addition of products within firms might be a significant margin of expansion for international trade. While the efficacy of trade policies has been studied carefully in the single-product firm set up, we still lack a good understanding of how trade policies affect multiproduct firms.

The response of multiproduct firms to trade policies has been studied in two sets of frameworks: static models with unobserved firm-product heterogeneity (e.g. Bernard, Redding and Schott (2010)) and static models with core competence on the supply side (e.g. Arkolakis and Mundler (2011)). In these setups falling entry costs induce entry of only marginally productive firms or the addition of only marginally profitable products. As a result incumbent firms respond to a reduction in trade costs with introduction of products that sell in minor quantities and add little to aggregate trade flows.

In this paper we document three data regularities that static models of multiproduct exporters cannot explain. We develop a dynamic model of multiproduct firms that accounts for these empirical regularities and use it to revisit the question of trade policy.

Using information on Chinese multiproduct exporters we document the following regularities: (1) multiproduct exporters introduce their best-selling products first, everything else constant; (2) more than 40% of the new products introduced by incumbent exporters are dropped due to low sales within the first year[1]; (3) the probability that a firm introduces a new product is positively related to the survival and success of its earlier products.

The first empirical regularity is consistent with firms having prior knowledge about the success of their future product lines, while the second points to uncertainty that incumbent exporters face when they introduce new products. The third pattern is consistent with firms learning about their potential in an export market as they introduce new products.

We develop a dynamic model that can fit all three of the above data regularities. In our model, a firm that contemplates entry into the export market draws a vector of demand shocks for each of its potential product lines. This set of demand shocks is known to the firm, but is unobserved to the econometrician. We call these "known" shocks. They

[1] Javoric and Iacovone(2012) document a similar pattern for Mexican firms, except they do not distinguish between new and incumbent exporters.

capture the prior knowledge a firm has about the success of each of its potential products. The fact that firms have prior knowledge about the demand for each of their potential products makes firms introduce their best products early on in their exporting experience.

To capture the uncertainty that firms face when they introduce new product lines and to allow firms to learn about their potential in the export market, we introduce a Bayesian learning mechanism into the model. Specifically, we assume that firms are endowed with a firm-specific demand parameter, which we will call a brand effect. A firm does not directly observe its brand effect, but the brand effect influences demand for each of its products in a stochastic manner. For each product that the firm introduces, it draws a second demand shock from a known distribution with the mean given by the brand effect. These shocks are the sum of the brand effect and a random component that represents uncertainty. (In the rest of the paper we will refer to this shock as "uncertainty" shock.) Neither the firm, nor the econometrician, observe this shock until the firm launches the product and observes its sales. We call these shocks *"s shocks"* as they are the signal from which the firm infers its brand effect. The presence of this uncertainty generates products that sell poorly and are dropped shortly after launch. As the firm introduces products into the export market it learns about its underlying brand effect, which generates history dependence, i.e., firms that have been successful in the past are more likely to expand.

The model also includes time-varying cost and demand shocks that are needed to account for the intertemporal variation in the data. More importantly, we allow the cost of introducing new products to vary with the product scope of the firm. We do so to ensure that learning effects are not conflated with economies or dis-economies of scope.

The empirical questions that we ask in this paper are twofold. First, we want to understand how important each of the mechanisms that we have introduced in the model is in the data (i.e. "known" demand shocks, uncertainty and learning about the brand effect). Second, we want to understand what the implications of these mechanisms for trade policy are. In our model "known" demand shocks, uncertainty and learning about the brand effect have opposing implications for the efficacy of trade policy and the overall result depends on which of the mechanisms dominates in the data. "Known" demand shocks imply that returns to introducing new products decrease with the scope of the firm. This suggests that a decrease in the costs of introducing subsequent products will induce introduction of only less profitable products. Uncertainty about demand in the export market implies that products that firms expect to be best-selling may not turn out to be successful ex-post. Similarly, products that firms do not expect to be successful may generate unexpectedly large sales. Learning about the brand effect further suggests that there may be additional gains from reducing costs for introducing new products. For example, lower tariffs or lower market entry costs will induce some foreign firms to start exporting, i.e., introduce their first product into the export market. As these new exporters learn about their ability to serve the foreign market though their first product, some will find they are high ability firms and will continue to enter new product markets more aggressively.[2]

To answer these questions we use data on Chinese exporters to the U.S. in the plastics industry, estimate the model structurally, and use it to perform counterfactual experiments. First, we find that known demand shocks play an important role in whether producers enter the export market or not. Without "known" demand shocks significantly fewer firms enter the export market. Second, we find that including uncertainty about the brand effect is necessary to account for large attrition among new exporters. When we let firms know their brand effect precisely, only those with sufficiently high brand effects enter. In that case, the model cannot replicate disproportionately large attrition of products among new exporters. Third, we find that while firms' actions are consistent with learning about their brand effect, the uncertainty they face in conjunction with introducing new products looms large and limits the impact of learning on firms' incentives to add new products. We find that the distribution of products among the high brand effect firms only marginally first order stochastically dominates the distribution for low brand effect firms. Furthermore, even when we preclude firms from updating their beliefs about the brand effect, we are still able to (largely) replicate the distribution of firms over the number of products conditional on the brand effect. Hence, we conclude that learning affects decisions of firms to introduce new products only moderately.

Finally, we revisit the question of trade policy in the multiproduct firm setting. We consider the effect of a decline in the market entry cost for new products on aggregate exports. Specifically, we look at three scenarios: only the cost of introducing the first product decreases, the costs of introducing all but the first product decrease, and the costs of introducing all products decrease. Naturally a decrease in the cost of introduction for all products has the biggest impact on aggregate sales. What is interesting is that decreasing only the cost for the first product has less effect on aggregate sales than decreasing the cost of introduction for all but the first product. In the first case aggregate sales increase, on average, by 6% and in the second case they increase, on average, by 9% over the period of ten years. This contrasts with the results of Arkolakis and Muendler (2011) who find that a large share of the simulated increase in trade and welfare is attributable to the decline in the cost of introducing the first product. We find that more than half of the increase in aggregate sales can be attributed to decreasing costs of introducing subsequent products. Hence, in the presence of economies of scope and uncertainty, decreasing costs

[2] Albornoz et. al (2012) propose a similar argument with regard to firms expansion in new geographic markets.

of market entry for subsequent products can make a significant contribution to increasing trade flows.

1.1 Relation to the literature

Our paper is closely related to the work of Arkolakis and Muendler (2011) on the account of our first data regularity, namely, that sales of products are negatively related to their order of introduction within a firm. Using information on Brazillian exporters they document that wide-scope exporters have their sales concentrated in few top selling products.[3],[4]They interpret this regularity as evidence of decreasing product specific efficiency , i.e., core competence. Taken literally, their model predicts that prices should rise as firms introduce products further away from their core competence. Since firms would introduce products closest to their core competence first, product prices should rise for products introduced introduced later in firms' exporting careers. Table 1 shows the regression of monthly prices on the product's order of introduction within a firm, number of products attempted by a firm and the age of the product in months. Price and order of introduction are negatively related suggesting that at least in our sample of firms core competencies do not belong to the cost side. Our first data regularity is consistent with the notion of core competence set up on the demand side, or what we call "known" demand shocks.

Our model is similar in spirit to that of Timoshenko (2013) in that we also explore the dynamics of firms learning about their brand effect through exporting new product lines. The novelty of our approach is that we introduce heterogeneity in the prior beliefs of firm managers about firm-product specific success ,i.e., "known" shocks. Our strategy to separate the effects of "known" demand shocks from "uncertainty" relies on the data regularity that firms introduce their best-selling products early on. The relative magnitudes of variances of the "known" shocks and "uncertainty" shocks determine the extent to which the effects of selection (i.e., firms introduce products that they expect to be best first) are manifested in the data. If the variance of the "uncertainty" shocks is large, the importance of the selection mechanism is mitigated: even if firms introduce their best selling products first, unexpected realizations of "uncertainty" shocks that are

[3]This regularity is similar to the Bernard, Redding, and Schott (2007) stylized fact that within firm sales of multiproduct exporters follow a Pareto distribution.

[4] It is worth pointing out that this regularity could come about purely as an artifact of order statistics, rather than as a consequence of core competence. Suppose that firms were heterogeneous in productivities and drew from a distribution of demand shocks. More productive firms would take on more risk and attempt to introduce more products. Some of their products will be successful; others will not be. The maximum (minimum) sales of a firm with many products would of course be higher (lower) than the maximum (minimum) sales of a firm with fewer products. This would result in the same pattern as they document. One could tell whether this patter is just an artifact of order statistics by looking at the order of introduction of a product and its sales, as we do.

not observed until the first sale has been made, will determine which products end up as best-selling in the data.

Our paper is also indirectly related to the literature that has focused on the relationship between scope and productivity in order to explain why multiproduct firms are few in number. Nocke and Yeaple (2013) focus on the "span of control approach" (as in Lucas (1978)), and model costs of firms rising for all products as the firm's scope increases. Eckel and Neary (2008) introduce the idea of core competence: as firms introduce products further away from the core competence, the marginal cost of each new product increases. Timoshenko (2013) and Arkolakis and Muendler (2011) both embed a model of Eckel and Neary (2008) into an open economy setup. We chose not to do so because in our data we do not find evidence of a positive association between price and order of introduction as would be implied by their assumption.[5]

The paper is organized as follows. In Section 2 we describe in detail the data regularities on which we base our model. Section 3 lays out the model. Section 4 describes the estimation procedure and intuition behind identification. Section 5 presents the results and Section 6 concludes.

2 Empirical Evidence

In this section we discuss the data patterns that shape our modeling decisions. We will use information on the universe of Chinese exporters supplying the US in years 2001-2006. For each exporter we have monthly data on sales and prices for each of their exported varieties at the 6 digit level. Unless stated otherwise before we do any analysis we standardize sales (prices) for each firm-product pair relative to the mean sales (prices) across firms for a given HS 6-digit category in a given month.[6]

First, we document that a product's median monthly sales and its order of introduction within a firm are negatively related. Figure 1 depicts the log of median monthly firm-product sales plotted against their order of introduction on the horizontal axis for the cohort of firms that started exporting to the US in 2001.[7] Specifically, we take all products that were introduced first, second, etc., by all the firms in the sample over the period of four years and compute the log of median sales for each group. There is a clear decrease in the median monthly sales as the order of introduction increases. If firms had some

[5] See Table 1

[6] Before we use information on prices and sales in our analysis we standardize them as follows. For each product introduced by a firm we calculate the ratio of sales per month to average sales in that product category by all firms in that month. This scaling makes sales comparable across products: a ratio of 1.4 means the product has 40% higher sales than the average for the product.

[7] Here we are not excluding products with quotas on them. The pattern is unchanged if we do.

information about which of their potential products were going to be successful in the export market, this pattern is exactly what we would expect.

To make the case for firms knowing about the potential success of their products in the market, we need to rule out that products introduced earlier have accumulated a bigger client base and so have larger sales. To address this concern we regress the log of firm-product monthly sales on the the number of products that the firm has, the product's order of introduction, and its tenure in months. We also incorporate year and industry fixed effects. Table 2 presents the results. The coefficient on a product's age in months is positive and significant indicating that older products indeed have larger sales. Nevertheless, the effect of the order of introduction remains negative and significant. [8] The coefficient on a number of products a firm has is positive, suggesting that overall more productive firms, or firms with a higher firm-level appeal introduce more products.[9]

Figure 2 shows the log of median monthly sales vs. the order of introduction for product lines that have been exported for the same number of months (2-16 months). The figure corroborates that firms enter exporting with the products they expect to be most successful.

Recent work by Albornoz et al. (2011), Eaton et al. (2008), and Freund et al. (2008) documents that exit rates are high among new exporters, and that exporters that survive the first year experience rapid sales growth. This has been interpreted as evidence for exporters learning about their appeal in the market. Firms with low appeal drop out of exporting, while firms that remain grow rapidly. Below we show that the data suggests that even firms that have previously exported to a market still face risk in conjunction with introducing new products[10]. Tables 3 and 4 provide information on firm-product pairs that were introduced into the export market in 2001.[11] Only firms that have exported before 2001 are included in the sample so that patterns that characterize first time exporters do not influence conclusions of the exercise.

In Table 3 we show how the number of firm-product pairs introduced in 2001 evolves as the cohort ages. The second column of the table reports the total number of products that are present in a given year out of the total number of products that were introduced in 2001. For example, out of the 9,440 new product lines introduced in 2001, 4,378 of them, or about 46%, are still sold in 2002, 3,174 in 2003 and so on. The product lines present

[8] As we are restricting attention to a cohort, we need not worry about composition effects due to single product firms being young (and possibly more productive).

[9] One may expect that only the more productive firms introduce new products. Therefore, we would expect products introduced earlier to have lower median sales than products introduced later and this would should only strengthen the pattern we observe. If we sort firms by the total number of products produced and repeat the above exercise, the same pattern emerges.

[10] By new products we mean products new for a firm, not a country.

[11] Industries with quotas on them include textiles, footwear and headgear. Here other forces are at play.

in the beginning of the year are divided into three groups depending on their situation in the beginning of the following year. Column three reports the products that are still sold in the following year. Column four shows the number of products that were discontinued by firms that continued exporting. Column five reports the number of products that exit because the carrying firm quits exporting. The number in brackets is the percentage of products in the given group relative to the total number of products that were sold in a given year. For example, of the 9,440 products introduced in 2001, we see that 4,054, or about 43%, of them are discontinued by firms that continued to sell other products in 2002, while 1,008 products disappear from the export market along with the firm that introduced the product.

A striking pattern is that more than 40% of the products that were introduced in 2001 are discontinued in the same year by firms that continue to export to the same market. Attrition in subsequent years drops to about 25%. We interpret this pattern as evidence that firms face uncertainty about demand for their new products. The fact that attrition stabilizes quickly after introduction also suggests that the uncertainty about demand for a given product is resolved soon after introduction.[12]

Table 4[13]shows the average monthly sales for the cohort of firm-product pairs that were introduced into the export market in 2001, conditional on whether the firm-product pair is still exported in the following year. Specifically, it shows average sales among continuing products (products introduced in t and present in $t+1$), products dropped despite the firm remaining in the export market (products in t, which exit in $t+1$ conditional on the firm staying in $t+1$), and products dropped due to firm exit from the export market (products in t, which exit in $t+1$ along with the mother firm). The average monthly sales among products that are exported in the following period are higher than among products that are discontinued, regardless of whether the carrying firm continues to export or not. This is consistent with firms facing uncertainty about their demand shock before they observe sales of their product.

Now we consider the possibility that firms are endowed with a brand effect that is common across products, and as exporters introduce new products they learn about their firm-specific potential in the export market. Such learning would imply that firms that have introduced successful products in the past would perceive this as evidence that their brand effect is high and would introduce new products more aggressively. By analogy, firms that introduced products that were dropped shortly after introduction will perceive

[12]This pattern is consistently present in all cohorts.

[13]Average monthly sales grow as cohort ages. The fastest growth occurs in the first year after introduction across all cohorts. This pattern also suggests that learning about the demand shock is fast. Also average sales across all cohorts increase in year 2005. Large growth in year 2005 can potentially be attributed to a number of reforms undertaken in that year, including trade and bank sector liberalizations.

this as evidence of a low brand effect and they would be less likely to introduce new products. To investigate the conjecture that the past successes and failures inform the decision of the firm to expand its scope, we estimate the binary logistic probability model of the firm introducing a new product line:

$$\Pr(y_{ft} = 1|X) = G\left(\alpha + \beta_1 F\_rate_{f(t-1)} + \beta_2 n_{f(t-1)} + \beta_3 Av.sales_{f(t-1)} + \beta_4 Age_{ft} + \beta_5 Tot\_\exp_t\right) \quad (1)$$

[14]

The dependent variable y_{ft} is the indicator variable that takes value 1 if the firm indexed by f introduces at least one new product in year t. The independent regressor $F\_rate_{f(t-1)}$ denotes the failure rate for the firm f as of time t. It is computed as the ratio of the number of products the firm has introduced and abandoned by year t relative to the total number of products the firm (f) has introduced since entry into the export market. It is included to test the hypothesis that firms learn about their firm-specific brand appeal as they introduce new products. If firms learn and take into account their histories, we expect the probability of introducing a new product to be negatively related to the share of a firm's failed products.

n_{ft-1} is the number of products that the firm has introduced by year t. It is included to account for the fact that the share of products dropped may have a different effect on the probability of introducing new products depending on the scope of the firm. The number of products that a firm has introduced also may capture the fact that large scope exporters have exhausted their best selling products and now have a lower probability of introducing a new product. Finally, it is possible that the cost of entering the export market with each subsequent product decreases(increases) with scope.

$Av.sales_{ft}$ stands for the average sales per product of a firm. We include average sales per product for each firm as a measure of firm productivity to account for the fact that firms that have experienced a rise in productivity will be more likely to introduce a new product.

The age of exporter f in year t (Age_{ft}) is simply the number of years we observe the firm exporting. Age of the firm is included to account for the fact that incentives of firms to introduce new products may change with their experience in the export market. For instance, young firms may experiment with new products to learn about their brand appeal.

$Tot\_\exp_t$ is the total exports of the Chinese firms in year t. Annual aggregate sales are included to account for the changes that affect all exporters over time. To control for the firm-specific time invariant effects (productivity, industry, etc.) we use a within estimator for panel data.

[14] $G(.)$ is the pdf of the logistic distribution.

In Table 5 we present the results of the regression for firms that started exporting between 2001 and 2004 for firms operating in quota-free industries. The proxy for the history of failures has a negative sign suggesting that firms indeed take into account their history. Average sales per product and number of products per firm are marginally significant. Average sales per product has a positive sign as expected. The number of products has a negative sign consistent with the hypothesis that firms introduce their best products first. Age has negative sign suggesting the value of experimentation for the young firms. Total exports that are included to capture time effects are all near zero and insignificant.

To summarize, we document that (1) multi-product exporters introduce their best-selling products early; (2) more than 40% of the new products by incumbent exporters are dropped due to low sales within the first year; (3) for a firm the probability of introducing a new product is positively related to the survival and success of its earlier products.

In the following section we describe the model that accommodates each of these three empirical regularities.

3 Model

We develop a partial equilibrium model of multiproduct exporter behavior consistent with the data patterns we have described in the previous section. Launching new products into the export market is costly and the success of these new products is uncertain. Firms choose to introduce the kinds of products they believe will be successful in based on their individual experience.

Our model is cast in continuous time and is based on the modeling techniques of Kortum and Klette (2004), Eaton et al. (2012), and Arcidiaconno et al. (2012). With time being continuous, instead of assuming that events and decisions are made at fixed intervals of time (i.e. yearly or monthly), we assume that decisions are made at stochastic intervals of time. For instance, we have data on monthly sales of the firm, so we could say that firms have to make a sale every month. Instead, we say that firms make sales on average every x months, and let the data determine the value of x. Allowing the data determine how frequently adjustments happen has a few advantages over fixing the times when firms draw shocks and make decisions. The payoff to introducing new products changes depending on the history of the firm. A continuous time framework allows firms to revise their behavior after every event. It also simplifies the computational burden: in continuous time a firm faces a decision to introduce one more product, rather than deciding how many products to introduce over a fixed interval of time as in a discrete time model. For instance, Timoshenko (2013) allows firms to experience a change in demand and introduce a new product only once a year. If in reality the bulk of firms make decisions and draw demand shocks more frequently than once a year, then the magnitude of the estimated parameters would have to compensate for the unrealistic rigidity of the model. Choosing a small but fixed interval of time would not bias estimates, but increases the computational burden greatly as the number of times the firm's problem needs to be solved grows rapidly.

We explicitly model a small open economy in a partial equilibrium framework because we do not have enough data to confidently estimate a general equilibrium model. We also focus on the US export market only. Previous research suggests that firms face different entry costs in different markets and may even face different demand structures. Meaningfully incorporating learning across products and geographic markets would be a computationally daunting task.

Henceforth: f indexes firms, n products and their order of introduction, and t time. We start with the exposition of the cost side of the firm.

3.1 Cost

To incorporate heterogeneity arising from the production side we model the marginal cost of an n^{th} product of a firm f as:

$$c_{fnt} = \exp(-\varpi_f + u_{fnt}) w_f^{\gamma_w} k_f^{\gamma_r} \tag{2}$$

where (ϖ_f) is the firm-specific productivity shifter and (w_f) and (k_f) are the firm-specific wage rate and capital stock, respectively. The capital stock is included as a size shifter, i.e., firms with different capital stocks presumably face different rates of return on capital. The effects of the wage and capital stock on the cost of the firm are measured by γ_w and γ_r. The productivity shock(ϖ_f) is constant over time for each firm and is drawn from a normal distribution with mean m_{ϖ_f} and variance $\varkappa_{\varpi_f}$. u_{fnt} is firm-product specific time variant cost shock. The distributional assumptions about the firm-product idiosyncratic shock (u_{fnt}) will be relayed later. At this point, we will just say that it changes with intensity (Poisson rate parameter) λ_{eu}.

The monopolistic competition assumption yields the price rule for each product that the firm makes:

$$p_{fnt} = \frac{\sigma}{\sigma - 1} c_{fnt} \quad \sigma \in (1, \infty) \tag{3}$$

3.2 Demand

A representative consumer at the export destination has CES preferences. There is a mass of firms supplying multiple products to the foreign consumer. Some of these firms are Chinese in origin. These are the firms we are studying here. In particular, there are K products in the universe of exported products. A firm can produce multiple products indexed by n, but is associated with a unique variety in each product f. The number of products is finite and in the empirical implementation will correspond to the four digit international product codes (HS 4-digit). The number of varieties, or firms, that produce each product can be infinite.

With CES preferences and the assumption that the elasticity of substitution between products and varieties is the same, demand for a product n made by firm f at time t is given by:

$$q_{fnt} = (p_{fnt})^{-\sigma} \, \Phi \exp(z_{fnt}). \tag{4}$$

σ is the elasticity of substitution between products and varieties and z_{fnt} is the demand shock for firm f, product n at time t. Since in the empirical implementation we standardize data on products, Φ is the market demand shifter common to all products.

From now on the subscript n stands for the order of introduction of the product within a firm, rather than a product category.

The demand shock z_{fnt} is a composite of the firm-product specific permanent shocks μ_{fn} ("known" shock), s_{fn} (s shock), and time variant idiosyncratic firm-product specific demand shock ε_{fnt}:

$$z_{fnt} = s_{fn} + \mu_{fn} + \varepsilon_{fnt} \tag{5}$$

We assume that each firm can make any product from a fixed set of products. For each of these products a firm draws a permanent product specific demand shock μ_{fn} prior to entry into the export market from a normal distribution $N(0, \kappa^2)$. The realizations of the product specific shocks μ_{fn} are known to the firm throughout its existence. This shock captures the amount of product-specific information that a firm has about the demand it is going to face in the export market.

The firm-product permanent demand shock s_{fn} is drawn from the normal distribution with mean η_f and variance ψ^2. More precisely, $s_{fn} = \eta_f + x_{fn}$, where η_f is unobserved to the firm and x_{fn} is normally distributed with mean 0 and variance ψ. The value of ψ is a common knowledge across all exporters. The firm-specific parameter, the brand effect η_f, in turn is drawn from a normal distribution with mean zero and variance τ^2 :

$$\eta_f \sim N(0, \tau^2) \tag{6}$$

The firm does not observe the value of η_f , but knows the distribution from which it has been drawn. The beliefs of the firm about the value of η_f evolve over time with the introduction of new products into the export market.

ε_{fnt} is drawn together with u_{fnt} from a joint normal distribution $N(0, \Sigma)$. The two are potentially correlated. The two shocks change simultaneously according to the homogeneous Poisson process with rate λ_{eu}.

We introduce correlation between the time-variant cost and demand shocks to relax the consequences of the monopolistic competition assumption. The monopolistic competition assumption rules out the possibility that profit maximizing firms with higher demand shocks can charge a higher price for their products. Allowing the time variant firm-product specific cost shock u_{fnt} to be correlated with the time variant firm-product specific demand shock ε_{fnt} permits us to model the demand that the firm faces as monopolistic competition without violating the data.

3.3 Timing & Information Set

Before we proceed to describe the problem of the firm in detail we lay out a brief preview of the model and describe the timing assumptions we make. In our set up a domestic firm that contemplates entry into the export market is described by four elements: (1) its firm level productivity draw ϖ_f, (2) belief about its firm-specific brand effect η_{fn}, (3) the set of "known" permanent firm-product specific demand shocks for each of its potential products in the export market, $\overrightarrow{\mu_f} = \{\mu_{fn}\}_{n=1}^{N}$, (4) the distribution over possible realizations of the time variant shocks, u_{fnt} and ε_{fnt}. At any instant of time a firm chooses an intensity with which it introduces a new product into the export market, and wether to continue exporting each product in its current export portfolio.

The firm observes the permanent firm-product specific demand shocks, $\overrightarrow{\mu_f}$, before it starts exporting. These shocks are meant to capture the idea that potential exporters have had domestic experience in selling their products and must have learned with which products they are most likely to succeed in the export market. Even exporters that have not sold domestically or exported to other destinations would have better knowledge about the potential of the firm in a set of products it can start exporting than would an econometrician. This is the notion of the demand shock typically employed in heterogeneous demand models where the demand shock is known to the firm but is unobserved to the econometrician.

By construction, expected profits from a product line are directly proportional to the realization of the firm-product specific shock μ_{fn}. It is therefore optimal for the firm to start exporting with the highest μ_{fn} product. To choose how much to invest into introducing a product into the export market, the firm compares the payoff to introducing the product with the cost of choosing the intensity with which this happens. The payoff depends on the expected discounted stream of profits from the product line plus the information value from learning about the firm-specific brand effect. A firm that chooses a hazard rate of introducing a product, λ_r, will be ready do so after a period of time determined by the exponential distribution with rate parameter λ_r.

Just before starting production of a product with a given "known" demand shock μ_{fn}, the firm observes the product specific cost and demand shocks $(\varepsilon_{fnt}, u_{fnt})$. Upon observing these two shocks the firm will decide whether to proceed with the introduction of the product or not. Variation in the firm-product specific demand and cost shocks is meant to capture changes in the buyer specific relationships or physical conditions at the firm's production facilities that would result in temporary changes in cost or demand. Since conditions at the production facilities are changing over time, it is reasonable to assume that firms take into account only the distribution over the possible realizations of ε_{fnt} and u_{fnt} when they contemplate entry into new markets. Should we not make this assumption, the number of state variables that the firm has to track when it introduces a new product would increase dramatically,[15] making the problem intractable.

A firm that decides to start production of its first product after observing the time variant firm-product specific cost and demand shocks will make its first sale at the exogenous Poisson rate λ_s. After making the first shipment of the product the firm learns the permanent firm product specific shock $s_{fn} = \eta_f + x_{fn}$ that has been unknown to it until sale, and updates beliefs about its firm level effect η_f in a Bayesian manner to $(\eta_{f1}, (\tau_1)^2)$. To be concise, we omit the firm-specific index on η_f, η_{fn} s_{fn}, and x_{fn} from now on.

The updating rule is given by the following sequential update where

$$s_n = \eta + x_n$$

$$\eta_{n+1} = \begin{cases} \frac{\eta_n \psi^2 + s_n \tau_n^2}{\psi^2 + \tau_n^2} \text{ if the firm has introduced a new product and observed } s_n \\ \eta_n \text{otherwise} \end{cases}$$

$$\tau_{n+1}^2 = \begin{cases} \frac{\tau_n^2 \psi^2}{\tau_n^2 + \psi^2} \text{ if the firm has introduced a new product and observed } s_n \\ \tau_n \text{ otherwise} \end{cases}$$

So far, we have considered a number of strict timing assumptions about the sequence of shocks realizations and timing of firms' actions. These assumptions considerably simplify the estimation procedure. One such assumption that deserves particular justification is that it is enough for the firm to observe one sale of a product to learn about its permanent firm-product specific demand component. While this assumption clearly oversimplifies the process through which firms learn about the demand they face for their new product we

[15] A firm would have to keep track of the evolution of the firm-product specific shocks ϖ_{fnt} and u_{fnt} for each product that it can start exporting.

believe that it does not bias our results. Table 3 shows the attrition rates for new products are high in the first year and stabilize from the second year on. This suggests that learning about the potential of a product in the market happens quickly.

3.4 Decision to continue exporting or terminate a product line

Before we move on to consider the firm's problem of introducing a new product line into the export market we look at the decision of the firm to keep or terminate an existing product. The present discounted value of a product that the firm is currently selling can be described by the Bellman equation:

$$V(\varepsilon, u; s_n, \mu_n, \varpi) = \tag{7}$$

$$\max\left\{\frac{-F + \lambda_s[\pi(\varepsilon, u; s_n, \mu_n, \varpi) + V(\varepsilon, u; s_n, \mu_n, \varpi)] + \lambda_{\varepsilon,u} E_{\varepsilon',u'} V(\varepsilon', u'; s_n, \mu_n, \varpi))}{\rho + \lambda_s + \lambda_{\varepsilon,u}}, 0\right\}$$

The present discounted stream of profits is denoted as $V(\varepsilon, u; s_n, \mu_n, \varpi)$ where we omit firm-product specific indices on ε_{fnt} and u_{fnt} for brevity. It depends on the firm-specific productivity ϖ, the two firm-product specific permanent demand shocks: μ_n and s_n, and the firm product specific time variant cost and demand shocks: ε, u. A forward looking firm that discounts future at rate ρ and contemplates whether to keep or terminate the product anticipates that it will make a sale at rate λ_s and collect profits in the amount $\pi(\varepsilon, u; s_n, \mu_n, \varpi)$. It also takes into consideration the evolution of the product specific cost and demand shocks (ε, u), which change with exogenously given intensity λ_{eu}. The firm pays the fixed cost of exporting a product F and can terminate the product at any instant of time if the expected value of profits fails to compensate for the expenditures on fixed cost.

3.5 Introduction of new products

Now we characterize how firms introduce new products. At each instant of time the firm chooses the intensity with which it introduces a new product to the market, λ_r. It choses a value of λ_r by comparing the expected benefit from adding a product line to the flow cost of maintaining a given value of λ_r, $c_n(\lambda_r)$. The cost of choosing an intensity of starting to export a new product, $c_n(\lambda_r)$, depends on the number of products the firm has attempted to export so far. Let $v(s_n, \mu_{n+1}, \varpi)$ denote the expected present value of profits from a product before the values of the time-variant firm-product specific shocks (ε, u) are observed. We obtain it by integrating the present discounted value of a product line, $V(\varepsilon, u; s_{n+1}, \mu_{n+1}, \varpi)$, over all possible realizations of ε and u:

$$v(s_{n+1}, \mu_{n+1}, \varpi) = \int\limits_{\varepsilon \times u} V(\varepsilon, u; s_{n+1}, \mu_{n+1}, \varpi) dN(\varepsilon, u; 0, \Sigma) \tag{8}$$

The present value of introducing the $(n+1)^{th}$ product for a firm with n products depends on the updated distribution of beliefs about the brand effect (η_n, τ_n^2), productivity ϖ, and a set of product specific shocks $\overrightarrow{\mu_f}$:

$$W((\eta_n, \tau_n^2); \varpi, \overrightarrow{\mu_f}) = \tag{9}$$

$$= \max_{\lambda_r} \left\{ \frac{-c(\lambda_r, n) + \lambda_r \int\limits_{s_{n+1}} \left[v(s_{n+1}, \mu_{n+1}, \varpi) + W((\eta_{n+1}(s_{n+1}), \tau_{n+1}^2); \varpi, \overrightarrow{\mu_f}) \right] dN(s_{n+1}; \eta_n, \tau_n^2 + \psi^2)}{\rho + \lambda_r} \right\}$$

where $N(s_{n+1}; \eta_n, \tau_n^2 + \psi^2)$ represents the normal distribution with mean η_n and variance $\tau_n^2 + \psi^2$. The optimal value of λ_r depends on the expected payoff from introducing a product: the expected stream of profits plus the value of learning about its brand effect relative to the cost of choosing the intensity.

To solve the model we assume that the cost attaining a particular hazard rate,λ_r, takes the following functional form:

$$c_n(\lambda_r) = \frac{c_{2n}(1 - (1 - \lambda_r)^{1 - \frac{1}{c_1}})}{1 - \frac{1}{c_1}} \tag{10}$$

We borrow this cost function from Arkolakis (2010) as it has a number of attractive properties. First, it does not satisfy the Inada condition as λ_r goes to 0, which allows us to replicate the empirical fact that a large fraction of firms choose to introduce just one or two products. Second, it lends itself to an analytical solution. Parameter c_1 determines the curvature of the cost function, while c_{2n} is the scale parameter. c_{2n} varies with the number of products that a firm has attempted to introduce in order to allow for (dis-)economies of scope. In the empirical implementation we assume that c_{2n} may be different for products for the first seven products and remain constant for more products. [16]

Solving the first order condition yields a closed form solution for the rate of introducing new products:

[16] The reason we assume that the cost of introducing seven or more products is constant is that we have few firms that have more than seven products, which makes estimating c_{2n} for large n difficult. An alternative would have been to impose a structure on c_{2n} as a function of n. Our estimates however suggest that the costs of introducing a new product do not systematically vary with scope, introducing a functional form could lead to misleading results.

$$\lambda_r = 1- \tag{11}$$
$$\left(\frac{\int\limits_{s_{n+1}} [v(s_{n+1};\mu_{n+1},\varpi) + W(\eta_{n+1}(s_{n+1});\varpi,\overrightarrow{\mu_f}]\, dN(s_{n+1};\eta_n^{n+1},\tau_n^2+\psi^2) - W(\eta_n;\varpi,\overrightarrow{\mu_f})}{c_{2(n+1)}} \right)^{-c_1}.$$

The intensity of introducing a new product has the upper limit of one[17] and is increasing in the difference between the expected value of introducing another product line and the value of maintain the current scope.

Our empirical model consists of key structural equations: the demand equation(4), price rule (3) and marginal cost function(2) , product introduction equation(9), and market participation decision for each product(7). In the following section we will proceed to describe the estimation routine.

4 Estimation and Identification

4.1 Data

In our structural estimation exercise we focus on firms that operated in the plastics industry. We chose the plastics industry because it was free from quotas and tariff restrictions during the sample period. We limit ourselves to information about exporters in 2001-2004, because our model is not equipped to handle the implications of a number of reforms that took effect in 2005, e.g., banking sector reform. In Appendix 3 we verify that the patterns we have documented for the universe of Chinese exporters persist in this subsample of firms that we use for the structural estimation.

We prepare the sample of firms for the estimation as follows. Using the firm survey dataset we obtain information on the wage rate, capital stock, firm registration date and start date of exporting for each firm. For firms that export directly[18], we add to the firm level information data on monthly sales and prices of products at the 4-digit level from the universe of customs transactions. After excluding firms that exited and reentered the sample we are left with 5,860 potential exporters of different ages, of which 645 enter into exporting in 2001. We track these firms from the moment of their entry to the end of 2004.

[17] The upper limit of one is never binding

[18] i.e., those firms that have a match in the customs data

4.2 Estimation Routine

We estimate all of the parameters in the described model except for the elasticity of substitution between the products, which set to values to values established in the literature. We use $\sigma = 8$ based on estimates in Das et. al. (2007), and the instantaneous discount factor, $\rho = 0.02$, which corresponds to the annual discount factor of 0.98, as in Arcidiaconno et. al. (2012). In order to identify the remaining 22 parameters we will use the indirect inference approach. Using actual data we create a vector of moments to match, $\overrightarrow{\beta_d(\theta_{true})}$. Given an initial guess for the set of parameter estimates, θ_0, we generate continuous time data using the model and then aggregate it to the same level as the observed data. Using the aggregated data we construct a counterpart of the targeted moments vector, $\overrightarrow{\beta_s(\theta_0)}$. Then using a global search genetic algorithm we numerically solve for the value of $\hat{\theta}$ such that it minimizes the objective function given by:

$$(\overrightarrow{\beta_s(\hat{\theta})} - \overrightarrow{\beta_d(\theta_{true})})' W (\overrightarrow{\beta_s(\hat{\theta})} - \overrightarrow{\beta_d(\theta_{true})}) \tag{12}$$

where W is a weighting matrix. We use a diagonal weighting matrix, whose elements are given by $Var^{-1}(\overrightarrow{\beta_d(\theta_{true})})$. We compute each element of $Var^{-1}(\overrightarrow{\beta_d(\theta_{true})})$ using the bootstrap.

For each firm in the firm-level survey we observe its first year of operation, annual wages, capital stocks, and exporting status. For exporters we have data on sales and prices for each of the exported products at the 4-digit level. In the simulation routine for each firm in the firm level survey we draw a productivity shock (ϖ_f) from the distribution of firm productivities $N(m_{\varpi_f}, \varkappa^2_{\varpi_f})$, and firm-specific brand effect, η_f, from the distribution of the brand effects $N(0, \tau^2)$, and a vector of time-invariant firm-product shocks for each of the 25 products that the firm can potentially make $\{\mu_{fn}\}_{n=1}^{N=25}$. Given the draws of primitive shocks, we solve the firm's problem to obtain the policy function: the intensity of introducing a new product. Given the intensity of introducing a product, λ_r, we draw the time when the firm starts production from the exponential distribution with the intensity λ_r. Next, we draw the time of the first sale from the exponential distribution with intensity λ_s, common to all firms and products. After the first sale the firm observes the demand shock signal $s_n = \eta_f + x_{fn}$ and updates its beliefs about the distribution of the demand shocks it is facing for its subsequent product, and chooses a level of intensity of introducing a new product.

For each product the firm introduces, we draw a vector of transitory cost and demand shocks, as well as the times of their change. Using the policy function of the firm, we determine which products are going to be sold and which are going to be dropped. The products with positive present values remain and we simulate sales of these products.

Eventually for each firm we will have an array of simulated prices, revenues and times of sale. We use this information to construct analogues of the annual and monthly datasets that we observe in the data.

4.3 Identification

Discussion of identification is informal. The mechanisms presented in the model are relatively complex and the primitive parameters are jointly responsible for generating the distribution of sales, prices, and products. To achieve identification we choose moments that are sensitive to changes in some parameters but not to changes in other parameters. The values of the data and their simulated counterparts are shown in tables 6 and 7 in the order they are discussed here.

First of all, we face the problem of separately identifying variance of the brand effects from the variance of the productivity shocks. τ^2 denotes the variance of the distribution of the brand effects (η_f). m_{ϖ_f} and $\varkappa_{\varpi_f}$ are the mean and variance of the normal distribution from which firm specific productivity shocks, ϖ_f, are drawn. In the model, differences in productivities across firms influence the distribution of prices, while firms' innate brand effects influence sales conditional on prices. To pin down the parameters that govern the distribution of firm-specific productivities we target the {0.25,0.5,0.75,0.98} percentiles of the distribution of the firm level price indices (average price across all products of the firm and across time).

In order to identify the variance of the distribution of brand effects, τ^2, we need to disentangle the price effect on firm sales from the brand effect. To do so we use information on prices and sales along with the model to derive an empirical measure of brand effects. In the model, the composite demand shock is given by $z_{fnt} = \frac{s_{fnt}}{p_{fnt}^{1-\sigma}\Phi}$ and Φ is constant across products, firms, and time. The ratio of sales to prices to the power of 1-σ ($\frac{s_{fnt}}{p_{fnt}^{1-\sigma}} \propto z_{fnt}$) is informative about firms' demand shocks net of productivity effects. Now, we construct a statistic that is informative about firm-specific brand effects. We will refer to it as η_f-*proxy*. First, we compute a proxy of a product specific demand shock z_{fn}-*proxy*, by averaging across monthly values of $\frac{s_{fnt}}{p_{fnt}^{1-\sigma}}$. Then, we compute η_f-*proxy* as an average across z_{fn}- *proxies*. For example, if a given firm f has introduced n_f products by the end of our sample $\eta_f\text{-}proxy = \frac{\sum_{n=1}^{n_f} z_{fn}-proxy}{n_f}$. We then match the {0.25,0.5,0.75,0.98} percentiles of the distribution of the brand effect proxies (η_f-*proxy*).

Another pair of parameters whose identification is tricky are the variances of the distribution of the unobserved demand shock x_{fn} and "known" demand shock μ_{fn}. The former, x_{fn}, is drawn from $N(0, \psi^2)$ and the latter (μ_{fn}) from $N(0, \kappa^2)$. Both of these shocks determine quantities of products sold in the market, as well as the probability of introducing a new product. A crucial distinction between the two is that μ_{fn} is known by the firm throughout its existence. As for the x_{fn}, the firm only observes the composite $s_{fn} = x_{fn} + \eta_f$ after the product has been produced and exported. Intuitively the distribution of x_{fn} pins down the share of products that are dropped by the firm after introduction due to low sales. Uncertainty about the realization of x_{fn}, and consequently s_{fn}, is the only source of uncertainty in the model and explains why firms introduce products that are quickly dropped due to low sales, as we observe in the data.

The presence of prior knowledge about the success of new products, or the "known" demand shocks μ_{fn}, implies that on average firms will select into their best-selling products early in their exporting career. The relative magnitudes of ψ and κ, the variances of x_{fn} and μ_{fn} respectively, determine the extent to which the effects of selection are manifested in the data. If ψ is large the importance of the selection mechanism is mitigated. Even if firms introduce their best selling products first, unexpected realizations of x_{fn}, and hence s_{fn}, will determine which products end up as best selling in the data.

To identify κ, we use information on the relationship between the product's demand shocks and it's order of introduction within a firm. The value of z_{fn}-*proxy* is informative about the underlying $\mu_{fn}+x_{fn}$. We regress the obtained value of z_{fn}-*proxy* of a permanent firm-product specific demand shock on its order of introduction within a firm and the number of products the firm has attempted over the sample period. We include the number of products attempted to account for the fact that more productive firms would export more products.

Parameters that determine how costly introduction of new products is (c_1 and $c_{21}, c_{22}, c_{23}, c_{24}, c_{25}, c_{26}, c_{27}$) influence the distribution of firms over the number of products. c_1 determines how much firms adjust the intensity of introducing a new product in response to changes in the expected profits from introducing new products. Hence c_1 not only determines how much firms adjust their intensities in response to learning more about their brand effect, but also how much intensities differ across firms with different productivities and different sets of "known" demand shocks. For lower values of c_1 we expect to have less dispersion in the number of products per firm than when c_1 is large. To this end, we match the number of firms that have from one to ten products.

In our model both the learning mechanism and the economies(dis-economies) of scope imply that firms' incentives to expand are influenced by the products they have introduced in the past. The implications of the two mechanisms differ as follows. Learning makes firms that have introduced successful products in the past more likely to expand, and firms that have introduced products that sold poorly less likely to expand. If costs of introducing new products are increasing (decreasing) with scope it affects all firms that have introduced a certain number of products, regardless of how well these products sold.

To help identify the scope parameters separately from the effect of learning we match the coefficients of a Poisson regression where the dependent variable is the number of products introduced in year y and the independent variables describe the state of the firm in the previous year $y-1$. The first independent variable is the proxy of a firm's beliefs about its brand effect after introducing n_f products by the end of year y. We denote it as η_{fny}-*proxy*. We first need to compute a statistic that is informative about the firm-specific demand shock. For each product the firm has introduced by the end of year $y-1$ we will compute an average of $\frac{s_{fnm}}{p_{fnm}^{1-\sigma}}$ over the number of months a product has been sold since introduction until the end of the year $y-1$. Finally, we approximate the belief of the firm about its brand effect, η_{fny}-*proxy*, as the average value of $\frac{s_{fnm}}{p_{fnm}^{1-\sigma}}$ over the number of products the firm has introduced by the end of the year $y-1$. We also include the number of products the firm has introduced and the number of products a firm has dropped by the end of the previous year.

The η_{fny}-*proxy* and the number of dropped products help to capture how much a firm responds to changes in their beliefs about its underlying brand effect. They help pin down the values of c_1, ψ, and τ. The number of products a firm has attempted so far is informative about the effect of scope parameters, i.e., $c_{21}, c_{22}, c_{23}, c_{24}, c_{25}, c_{26}, c_{27}$. If costs of introducing new products are decreasing with the number of products, we expect large scope firms to introduce products more intensively, despite the fact that incentives to introduce new products decrease as firms run out of their "best-selling" products. Similarly, if costs of introducing new products are increasing with the number of products we expect that the intensity with which firms introduce new products will decrease with scope beyond what is implied by the value of κ, which governs the extent to which firms appear to introduce their best products first in the data.

The age of firms at the time when they start exporting to the US helps us identify the mean and variance of the firm-specific productivity distribution and cost of introducing the first product into the export market. Our model implies that firms that took longer to start exporting since registering are less productive relative to firms that start earlier.

Even though we have information on the registration date of all firms, we treat firms that were registered before 1999 as if they did so in 1999 to avoid dealing with long term dynamics. Firms that have been producing domestically before 1999 (as early as 1916) may have experienced a change in productivity, ownership, etc. over such a long period of time. Our model rules that out and would misinterpret firms that took a long time to start exporting as unproductive.

The fixed cost of exporting a product is pinned down by the mean sales among the products that are dropped. The intensity of making a sale λ_s is identified by matching the average number of shipments per year.

In order to identify the variance-covariance matrix of the distribution of the time variant cost and demand shocks, Σ, and the hazard rate of a change in these two shocks,λ_{eu}, we target moments of monthly price and sales distributions. We match the four quantiles {0.25,0.5,0.75,0.98} of monthly prices and backed out demand shock distributions. We also target correlation between prices and sales.

The correlation between prices and wages, and the correlation between prices and firms' capital stocks pin down the effect of wage rates and capital stocks on the marginal cost of the firm.

5 Results

In this section, we present the estimates of the parameters, as well as the counterfactual experiments.

5.1 Estimates

Table 8 shows the estimates of the demand side parameters. The first parameter in the table, τ, is the standard deviation of the firm-specific brand effects. The second parameter, ψ, determines the degree of uncertainty that firms face in the export market when they introduce new products. These two parameters govern how long it takes a firm to learn about its brand effect. The larger ψ and τ are, the harder it is for a firm to extract information about its brand effect from the signal it receives when it makes a sale. The value of ψ determines the residual uncertainty that persists even when the brand effect is fully known by the exporter. Figure 3 shows the evolution of the variance of the exporter's beliefs about its underlying brand effect as it introduces new product lines. The variance of the perceived brand effect shocks falls by 20% after the introduction of the first product. After the introduction of the second product, the variance falls by another 13%. The incremental effects of subsequent new signals decrease monotonically, so that most of the learning happens with the introduction of the first few products.

The third row of the table shows the estimate of κ, the standard deviation of the "known" shocks. Even though it is small in magnitude our counterfactuals suggest that it plays an important role in determining entry of firms into exporting.

The value of λ_s we estimate translates into a product being shipped to the export market at an average rate of 5 times per year, or approximately every 2.5 months. This is consistent with the data sample average. All of the learning and demand estimates are significant.

Table 9 shows the estimates of the supply side parameters. The first two parameters are the mean and standard deviation of the firm-level productivity distribution. Comparing the standard deviation of the productivity shocks and the standard deviation of the firm-specific brand effects suggests that the bulk of firm heterogeneity lies on the demand side. The wage coefficient β_w is positive and significant as would be expected. The coefficient on capital stock, a measure of firm's size, has a positive sign, and is insignificant. This is consistent with the absence of size effects.

The value of λ_{eu}, the hazard rate of transient costs and demand shocks changing, translates into such changes occurring on average twice a year. $\Sigma_{11} = .283$ and $\Sigma_{22} = .973$ point to large variation in prices and quantities over time for each product. The variance of the transient cost shocks is particularly large. This suggests that firm-specific productivity is not sufficient to account for variation in costs. The correlation between the cost and demand is small but still positive.

The first row of Table 10 gives the estimate of the instantaneous fixed cost of exporting a product. It translates into the fixed cost of exporting a product being about 39% of the average profits from a product line across all firms and products. The scope parameters c_{21}, c_{22}, c_{23}, c_{24},c_{25}, c_{26} and c_{27} are roughly decreasing with the number of products attempted, consistent with limited economies of scope.

5.2 Uncertainty, Experimentation, Selection into Products, and their Consequences.

Next we use our estimates to assess the importance of "uncertainty" shocks, learning about the brand effect, and the "known" demand shocks in generating the histories of firms observed in the data.

We will consider several scenarios. In the full-information scenario we generate the data under the assumption that firms know their brand effect exactly. We continue to have firms draw their brand effects from the population distribution, $N(0, \tau^2)$, but assume these are known to the firm. To consider the "no learning" scenario we preclude firms from updating their beliefs about their brand effect draw. To understand the role that the "known" demand shocks play in determining the behavior of firms we simulate the model setting the variance of the "known" shocks to zero.[19]

In each scenario we simulate a pool of producers that contemplate entry into exporting. We allow entry of new firms for three quadrimestres (or one year) and then restrict entry of new firms into exporting and track this cohort of firms for another nine quadrimestres

[19]We might have just made the known demand shock unknown, i.e., added it to the unknown demand shock instead. We chose not to do this as this would raise the noise in the model and change the extent to which firms could learn about their brand effect.

(three years).

5.2.1 Baseline case

Our model predicts that firms learn about their underlying brand effect through introducing new products. Firms that have a high brand effect or firms that come to believe they have a high brand effect, will add products more intensively. Conversely, firms that come to believe they have a low brand effect stop expanding. This has implications for the distribution of the number of products per firm for a cohort of exporters. We expect that as the cohort matures, the distribution of the number of products per exporter for firms with a high brand effect will first order stochastically dominate the distribution for those with lower brand effects. This is more so when the variation in the "uncertainty" shocks is low and learning happens quickly. Figure 4 illustrates the case when we set the standard deviation of the "uncertainty" shocks at $\psi = 0.2$ instead of the estimated $\psi = 0.9$. It shows the distribution of firms over the number of products for a cohort of exporters that started exporting this year conditional on their brand effect. We consider two groups of exporters: those with a brand effect higher than 0.1 and those with brand effect below -0.1.

The first panel of the figure 4 shows the high brand effect exporters and low brand effect exporters in their first year with $\psi = 0.2$. There are 41 high brand firms and 49 low brand ones entering the export market during one year. In the first year the two groups of exporters exhibit similar distributions of firms over the number of products. This is natural since prior to entry exporters don't know about their brand effect. As the cohort ages, high brand effect firms add a larger number of products, so that their distribution moves to the right and grows taller. This pattern becomes clearer as the cohort ages in the third and fourth panels of the Figure 4.

Now consider figure 5. This figure is analogous to Figure 4 except that we have used the estimated variance of the "uncertainty" shocks of .9. In this scenario, the uncertainty that exporters face is much larger than in the previous scenario. This means learning is more difficult now and firms need to introduce more products to learn about their brand effect. Thus, the difference in the number of products that high and low brand effect firms introduce over time is smaller.

Comparing figures 4 and 5 we can also see that when uncertainty is reduced fewer firms enter, and fewer new products are introduced. This is expected because forecasted profits increase in ψ: when uncertainty is large even firms with a low productivity and a low brand effect stand a chance to launch a profitable product with a high unobserved firm-specific demand shock.

5.2.2 No Learning case

When we eliminate the learning mechanism we continue to have firms draw brand effects from the population distribution, but preclude firms from updating their beliefs. Firms perceive that the unobserved demand shock s_n is drawn from the normal distribution with mean zero and variance given by $\tau^2 + \psi^2$ at all times. Here we use the estimated value of $\psi = 0.89$.

Without learning we don't expect to see differences in firms' incentives to introduce new products. Figure 6 shows the distribution of firms over the number of products for a cohort of exporters conditional on the brand effect. As in the previous counterfactual we consider firms that draw a brand effect less than -0.1 (113 entrants) and those with a brand effect greater than 0.1 (107 entrants). Note that the distributions for low and high brand effect firms still differ, but this is just because of selection. High brand effect firms drop products less often even though they never learn about their brand effect.

Comparing the distributions of firms over the number of products conditional on the brand effect in the baseline scenario and the analogous distribution of firms in the no learning scenario in Figures 5 and 6 respectively, one can see that the distributions in the two cases are similar. This reiterates the observation that just a moderate amount of learning is sufficient to rationalize the product introducing behavior of exporters in the data.

5.2.3 Full information

Now we consider the full information counterfactual where firms observe their brand effect draw, η_f, but still face uncertainty about the firm-product specific demand shocks.

Figure 7 shows the distribution of firms over the number of products for a cohort of exporters by brand effect. We consider firms with a brand effect less than -0.1 and greater than 0.1. The pattern that comes across is the disproportionately large entry of high brand exporters. 167 exporters with high brand and 48 with low brand effect enter exporting. The low brand effect exporters that enter are the highly productive exporters that introduce new products intensively, and so we see a few large scope low brand effect firms in panels 3-4 that show matured exporters.

Comparing Figure 7 to its baseline scenario analogue Figure 5 suggests that uncertainty about the brand effect is needed to account for the share of low brand effect firms entering exporting and generating attrition among new exporters.

5.2.4 No "known" demand shocks

When we set the variance of the "known" demand shocks to zero, firms no longer select their best products to introduce first and their incentive to introduce new products will not decrease over time. This also means that now firms are identically uncertain about the overall demand for their new products. One would expect that fewer firms will introduce their first product, or enter exporting.

Figure 9 shows the distribution of firms over the number of products for a cohort of exporters under the baseline and no "known" demand shocks assumption. As expected the number of products in the no "known" demand shocks scenario falls relative to the baseline. From the first panel in Figure 9 we can see that fewer firms enter in the no "known" demand shocks than in the baseline scenario.

5.2.5 Aggregate magnitudes

Now that we have a better understanding how each of the mechanisms is reflected in the data, we are interested in their implications for aggregate trade flows. Here we consider the implications of the learning mechanism, uncertainty about the brand effect and the "known" demand shocks on aggregate sales.

Figure 10 shows aggregate sales in each of the four counterfactual scenarios we have considered so far. Implications of the full information and no "known" shocks scenarios are clear. In the first case, high brand effect firms will introduce new products more intensely and drive aggregate sales up. In the no "known" shocks case fewer firms enter exporting and as a result aggregate sales will be lower than in the baseline case. Aggregate sales in the no learning and baseline cases are very similar.

To understand why, consider Figure 11 ,which compares the distribution of the number of products resulting in the no learning and the baseline cases. One can see that the number of new exporters is in fact higher in the no learning case than in the learning case. This means that expected present value of entry is in fact higher when firms are not learning. This happens because expected profits are increasing in both τ and ψ.[20] Overall, the two distributions appear similar in each of the four years. The number of small scope firms is larger in the no learning scenario because firms that fail in the first few products do not cease to attempt to introduce new products as they do when learning takes place. There are a few large scope exporters in the baseline scenario. The fact that aggregate sales in the two scenarios are similar suggests that sales of the large scope exporters in the learning scenario contribute to aggregate sales just as much as a large

[20]Since τ is not decreasing over time, as it does when firms are learning, firms at any scope have equal incentive to add new products to their exporting portfolio.

number of single product exporters in the no learning case.

To summarize, our counterfactual experiments have identified the channels through which the mechanisms incorporated in our model (uncertainty, "known" demand shocks, and learning about brand effect) operate. First, "known" shocks influence firms' decisions of whether to start exporting or not. In other words, firms cope with uncertainty that they face in the export market by introducing products that they expect to be successful first, i.e., products that have sold well domestically. This resonates with the empirical regularity that firms usually start exporting with products that they have sold domestically.[21] Second, we find that uncertainty is large. Even though firms update their beliefs in a Bayesian manner, the learning is so noisy that their behavior is similar to firms that do not update their beliefs.

In the following section we reduce costs of introducing new products and evaluate the quantitative implications of the policy.

5.3 Costs of introducing new products.

In our model the cost of introducing the first product is included in the cost of starting to export for a firm as it does so with the first product. Here we consider the effect of a decline in the cost of introducing new products into the export market on aggregate exports. Specifically, we consider three scenarios. First, we consider a 25% drop in the cost of introducing the first product (25% drop in c_{21}). Second, we decrease the cost of introducing all of products for a firm (25% drop in c_{21},c_{22},c_{23},c_{24},c_{25},c_{26},c_{27}). Finally, we consider a 25% decrease in the cost of introducing all but the first product (25% drop in c_{22},c_{23},c_{24},c_{25},c_{26},c_{27}).

We simulate a cohort of potential exporters that start production at the beginning of time (i.e., 2001) and gradually enter exporting. We follow their activities in the export market for thirty quadrimesters[22]. Figure 12 shows the result. Naturally a decrease in costs for all products has the biggest impact on sales. What is interesting is that decreasing only the cost for the first product has less of an effect on aggregate sales compared to decreasing the cost of introduction for subsequent products, but not the first product. In the first case aggregate sales increase by an average of 6% and in the second case they increase by an average of 9% . Decreasing the costs of introducing all but the first product disproportionately affects the more productive exporters, and those exporters who had high s shocks draws and believe that they have a high brand effect.

[21] Javoric and Iacovone document this pattern for Mexican firms.

[22] In this counterfactual we consider firms entering throughout the simulation time.

The finding that decreasing the cost of introducing subsequent products has at least as much effect on aggregate sales as does decreasing the cost of introducing the first product relates to the two experiments with entry costs performed in Arkolakis and Muendler (2011). In the first scenario they lower the cost of introducing only for the first product. In the second, they reduce the cost of introduction into the export market for all products. They find that the results from the two experiments are similar to each other, and conclude that the simulated increase in welfare is attributable to a decline in the firm's entry cost for the first product. Their explanation is that product efficiency decreases fast along with costs of introducing new products. As a result only wide scope exporters find it profitable to introduce new products, but these products sell in minor amounts and matter little for bilateral trade.

It is fair to say that, since our model is a partial equilibrium one and is estimated for just a single industry we could not derive implications for multi-country trade flows to make it directly comparable to the results of Arkolakis and Muendler (2012). Nevertheless, our counterfactual suggests that in the presence of uncertainty and even moderate learning effects, decreasing the costs of introducing subsequent products can make a significant contribution to increasing trade flows.

6 Conclusion

In this paper we quantify the importance of the three mechanisms in determining observed firm outcomes. The three mechanisms are: uncertainty firms face about demand when introducing new products (1), firms learning about their brand effects (2), and firms' prior knowledge about their potential in each of their products prior to product introduction (3). To do so we develop a dynamic model of multiproduct exporters with heterogeneity both on the demand and supply side. We estimate it using information on firms in the Chinese plastics industry and detailed information on their exports to the US.

We find that the incorporating "known" demand shocks and "uncertainty" into the model is empirically important in order to account for firms' product introducing behavior. We find that "known" shocks play a significant role in determining new exporter behavior and that "uncertainty" in the export market is high, making learning noisy.

Next, we revisit the question of trade policy in multi-product setting. We simulate a decrease in the cost of introducing new products for firms. Our simulations suggest that in the presence of economies of scope and even moderate learning effects, decreasing costs of introducing subsequent products can make a significant contribution to increasing trade flows.

References

[1] Ariel Pakes, Richard Erickson. "Empirical Implications of Alternative Models of Firm Dynamics." *Journal of Economic Theory*, 79,1-49,

[2] Andrew B. Bernard, Stephen Redding, and Peter K. Schott. "Multiple-product firms and product switching." *American Economic Review*, Volume 100, Issue 1, pages 70-97.

[3] Andrew Bernard, Stephen Redding and Peter K. Schott. "Multi-Product Firms and Trade Liberalization." *The Quarterly Journal of Economics*, Volume 126, Issue 3, pages 1271-1318. 2011.

[4] Bee Yan Aw, Mark J. Roberts, Daniel Yi Xu. "R&D Investment, Exporting, and Productivity Dynamics." *The American Economic Review*, Volume 101, Number 4, June 2011, pages 1312-1344.

[5] Boyan Javanovic. "Selection and Evolution of Industry." *Econometrica*,Voumel 50,Number 3, May 1982.

[6] C. A. Hidalgo, B. Klinger, A.L. Barabasi, R. Hausmann. "The Product Space Conditions the Development of Nations." *Science* 317, 481-487, 2007

[7] Costas Arkolakis, Theodore Papageorgiou. "Firm Selection, Growth and Learning", Mimeo,Yale University, and NBER Pennsylvania State University,December 31, 2009

[8] Costas Arkolakis. "Market Penetration Costs and the New Consumers Margin in International Trade." *Journal of Political Economy*, University of Chicago Press, vol. 118(6), 2011, pages 1151-1199.

[9] Costas Arkolakis, Marc-Andreas Muendler."The Extensive Margin of Exporting Products: A Firm-level Analysis."

[10] Facundo Albornoz, Hector Calvo-Pardo, Gregory Corcos, Emanuel Ornelas. "Sequential exporting: how firms break into foreign markets." CentrePiece - The Magazine for Economic Performance 364, Centre for Economic Performance, LSE. Downloaded at http://ideas.repec.org/e/pal164.html.,2012.

[11] Gregory Crawford, Matthew Shum. "Uncertainty and learning in pharmaceutical demand." *Econometrica*, Volume 73, Issue 4, pages 1137-1173. July, 2005.

[12] Joseph LeCates. "Perception Spillovers Between Bundled Goods: A Model of Physician Learning from Chemotherapy Choice", Cornell University.

[13] Jonathan Eaton, Marcela Eslava, Maurice Kugler, James Tybout. " The Margins of Entry into Export Markets: Evidence from Colombia", in Elhanan Helpman, Dalia Marin, and Thiery Verdier, eds., The Organization of Firms in a Global Economy, Cambridge, MA: Harvard University Press, 2008.

[14] Jonathan Eaton, Marcela Eslava, C. J. Krizan, Maurice Kugler, James Tybout, David Jinkins (2011). "A Search and Learning Model of Export Dynamics." Downloaded at strategy.sauder.ubc.ca/schiff/seminars/EEKKT10b.pdf.

[15] Lucia Foster, John C. Haltiwanger, Chad Syverson. "The Slow Growth of New Plants: Learning about Demand?" NBER Working Paper No.17853 Issued in February 2012

[16] Leonardo Iacovone, Beata S. Javorcik. "Multi-Product Exporters: Product Churning,Uncertainty and Export Discoveries." *The Economic Journal*, Volume 120, Issue 544, pages 481-499, May 2010

[17] Peter Arcidiaconno, Pat Bayer, Jason Blevins, and Paul Ellickson. "Estimation of Dynamic Discrete Choice Models in Continuous Time with an Application to Retail Competition." Updated May 2, 2013.

[18] Sanghamitra Das, Mark J. Roberts, James R. Tybout. "Market Entry Costs, Producer Heterogeneity, and Export Dynamics." *Econometrica*,Volume 75, Issue 3, pages 837-873, May 2007

[19] Volker Nocke and Stephen Yeaple, 2013. "Globalization and Multiproduct Firms." Discussion Paper Series of SFB/TR 15 Governance and the Efficiency of Economic Systems 408, Free University of Berlin, Humboldt University of Berlin, University of Bonn, University of Mannheim, University of Munich.

[20] Zhanar Akhmetova and Cristina Mitaritonna. "A Model of Firm Experimentation under Demand Uncertainty with an Application to Multi-Destination Exporters"

7 Appendix 1. Tables.

Table 1: Product monthly prices and its order of introduction. Plastics industry.

	Dependent variable: log of monthly product prices
Age in months	0.0008
	(0.0007)
Order of introduction	-0.0281***
	(0.0013)
# of products	-0.0215
	(0.0358)
N	34549
r^2	0.3667

Standard errors in parentheses.
$^{*}p < 0.05$, $^{**}p < 0.01$, $^{***}p < 0.001$
Includes year fixed effects.

Table 2: Product monthly sales and its order of introduction. All firms excluding textiles.

Dependent variable: Monthly product sales

	REG1	REG2
# of products	0.0064***	0.0038***
	(0.0004)	(0.0004)
Order of introduction	-0.0943***	-0.0194***
	(0.0014)	(0.0019)
Age in months		0.0199***
		(0.0003)
N	738,239	738,239
r^2	0.0070	0.0118

Standard errors in parentheses.
$^{*}p < 0.05$, $^{**}p < 0.01$, $^{***}p < 0.001$
Includes year and industry fixed effects.

Table 3: Evolution of the number of products introduced in 2001. Sample includes all firms excluding textiles, footgear and headgear that started exporting before 2001. HS 4-digit category.

t	Number of products introduced in 2001 and sold in t.	Sold in t and $t+1$ (% of total).	Sold in t only, carrying firm is present in $t+1$ (% of total).	Sold in t only, carrying firm exits in $t+1$ (% of total).
2001	9,440	4,378 (46%)	4,054 (43%)	1,008 (11%)
2002	4,378	3,174 (72%)	909 (21%)	295 (7%)
2003	3,174	2,426 (76%)	547 (17%)	201 (6%)
2004	2,426	1,861 (77%)	444 (18%)	121 (5%)
2005	1,861	1,345 (72%)	401 (22%)	115 (6%)

Table 4: Average monthly sales per product. Cohort of products introduced in 2001.

t	Mean sales for products introduced in 2001 and sold in t.	Mean sales for products in t, which exit in $t+1$ conditional on firm staying.	Mean sales for products in t, which exit in $t+1$ along with the firm.
2001	0.56	0.24	0.24
2002	0.82	0.38	0.51
2003	0.95	0.43	0.40
2004	1.05	0.31	0.55
2005	1.72	0.41	0.32

Table 5: Logit, FE

Dep.var. takes value 1 if a firm has introduce at least one new product in a year, 0 otherwise.	
Share of products dropped after one year in total number of products ($F\_rate_{f(t-1)}$)	-0.9518***
Number of products ($n_{f(t-1)}$)	-0.2822**
Av. sales per product ($Av.sales_{f(t-1)}$)	0.0416**
Age	-0.1205***
Total exports	0
Obs.	4,198

Standard errors in parentheses.
$^{*}p < 0.05$, $^{**}p < 0.01$, $^{***}p < 0.001$

Table 6: Data vs. Simulated Moments.
($Q_{.p}(.)$ stands for p^{th} percentile of variable (.))

Moment	Data	Simulated
Firms' productivity distribution $N(m_{\varpi_f}, \varkappa^2_{\varpi_f})$.		
Firm level price, quantiles		
$Q_{.25}(p_f)$	-1.200	-1.016
$Q_{.5}(p_f)$	-0.786	-0.455
$Q_{.75}(p_f)$	-0.248	-0.191
$Q_{.98}(p_f)$	2.567	0.4
Firms' brand effects distribution $N(0, \tau^2)$.		
Firm brand effect ($\eta_f - proxy$)		
$Q_{.25}(\hat{\eta_f})$	-9.319	-4.721
$Q_{.5}(\hat{\eta_f})$	-4.0249	-4.033
$Q_{.75}(\hat{\eta_f})$	0.377	-3.3026
$Q_{.98}(\hat{\eta_f})$	33.956	0.139
Distribution of "uncertainty"(x_{fn}) shocks $N(0, \psi^2)$.		
Share of products dropped	0.213	0.213
Distribution of "known"(μ_{fn}) shocks $N(0, \kappa^2)$.		
$\overline{z_{fn} - proxy} = \alpha + \beta_1 n_f + \beta_2 max(n_f)$		
α (se)	-4.856(0.269)	-4.464(0.147)
β_1 (se)	-0.127(0.044)	-0.093 (0.044)
β_2 (se)	0.0497(0.0291)	0.107(0.0341)
Cost of introducing new products ($c_1, \{c_{2n}\}_{n=1}^7$).		
Share of firms that have introduced N products		
N=1,2	0.325, 0.187	0.283,0.233
N=3,4	0.142, 0.076	0.183, 0.067
N=5,6	0.0526, 0.0371	0.066, 0.0333
N=7,8	0.023, 0.0263	0.050,0.033
N=9,10	0.0263, 0.020	0.050,0.002
Cost of introducing new products ($c_1, \{c_{2n}\}_{n=1}^7$). Learning effects, i.e. ψ.		
#new products in year y=$\alpha + \beta_1 \eta_{f}\hat{_{ny-1}} + \beta_2$# products dropped (y-1) $+\beta_3 n_{fy-1} + \beta_4 year$		
α	112.183 (92.159)	126.75 (72.691)
β_1	0.048 (0.01)	0.028 (0.052)
β_2	-0.002 (0.0231)	-0.044(0.097)
β_3	0.095(0.006)	0.063(0.012)
β_4	-0.056(0.046)	-0.063(0.036)
Firms' productivity distribution $N(m_{\varpi_f}, \varkappa^2_{\varpi_f})$ & c_1.		
Share of firms born in 1998 started exporting in		
2001	0.124	0.08
1999	0.065	0.005
2000	0.0534	0.0231
Fixed cost of exporting a product F.		
Mean sales among dropped products	-2.620	-2.144
Frequency of shipments, λ_s.		
Average number of shipments per year	5.664	6.82

Table 7: Data vs. Simulated Moments (Continued)

Moment	Data	Simulated
Distribution of transitory cost (u_{fnt}) and demand shocks (ε_{fnt}). Σ. λ_{eu}.		
Firm-product price, quantiles		
$Q_{.25}(p_{fny})$	-1.244	-0.981
$Q_{.5}(p_{fny})$	-0.610	-0.457
$Q_{.75}(p_{fny})$	-0.131	-0.215
$Q_{.98}(p_{fny})$	5.470	0.482
Firm-product sales, quantiles		
$Q_{.25}(s_{fny})$	-2.134	-2.247
$Q_{.5}(s_{fny})$	-0.615	-0.718
$Q_{.75}(s_{fny})$	0.6803	1.035
$Q_{.98}(s_{fny})$	4.815	5.589
Covariance between prices and capital stocks	0.1122	0.0014
Unit cost of production.		
Covariance between prices and sales	0.105	-0.919
Covariance between prices and wages	0.0323	-0.0211

Table 8: Learning and Demand Parameters.

	Parameter	Point estimate(st.error)
τ	St.dev. of brand effects	0.4486 (0.068)
ψ	St.dev. of unobserved demand shocks	0.8854 (0.011)
κ	St.dev. of known demand shocks	0.1316 (0.012)
Φ	Demand shifter	0.0107 (0.034)
λ_s	Hazard rate of making a sale	0.1938 (0.277)

Table 9: Cost of production.

	Parameter	Point estimate(st.error)
$m_{\varpi f}$	Mean of productivity shocks distribution	-0.2266(0.01)
$\varkappa_{\varpi_f}$	St.dev. of productivity shocks distribution	0.1706 (0.0018)
β_k	Elasticity of marginal cost wrt. to capital stock	0.0015 (0.0014)
β_w	Elasticity of marginal cost wrt. to wage	0.0255 (0.0076)
λ_{eu}	Hazard rate of a change of transient cost and demand shocks	0.0365 (0.036)
Σ_{11}	St.dev of the transient demand shocks	0.2831 (0.0006)
Σ_{12}	Covariance of the transient cost and demand shocks	0.0073 (0.0036)
Σ_{22}	St.dev. of the transient cost shocks	0.973 (0.0386)

Table 10: Estimates of parameters governing introduction of new products: Economies of Scope.

	Parameter	Point estimate(st.error)
F	Fixed cost of exporting a product	0.0375 (0.003)
	Cost of introducing a new product	
c_1	Curvature	0.0114 (0.01)
c_{21}	Cost shifter for the 1^{st} product	0.5447 (0.0011)
c_{22}	Cost shifter for the 2^{nd} product	0.4307 (0.0019)
c_{23}	Cost shifter for the 3^{rd} product	0.2495 (0.027)
c_{24}	Cost shifter for the 4^{th} product	0.3630 (0.0065)
c_{25}	Cost shifter for the 5^{th} product	0.4940 (0.0396)
c_{26}	Cost shifter for the 6^{th}	0.1121 (0.0032)
c_{27}	Cost shifter for the 7^{th} and higher products	0.065 (0.0462)

8 Appendix 2. Figures.

Figure 1. Median monthly product sales vs. order of introduction.

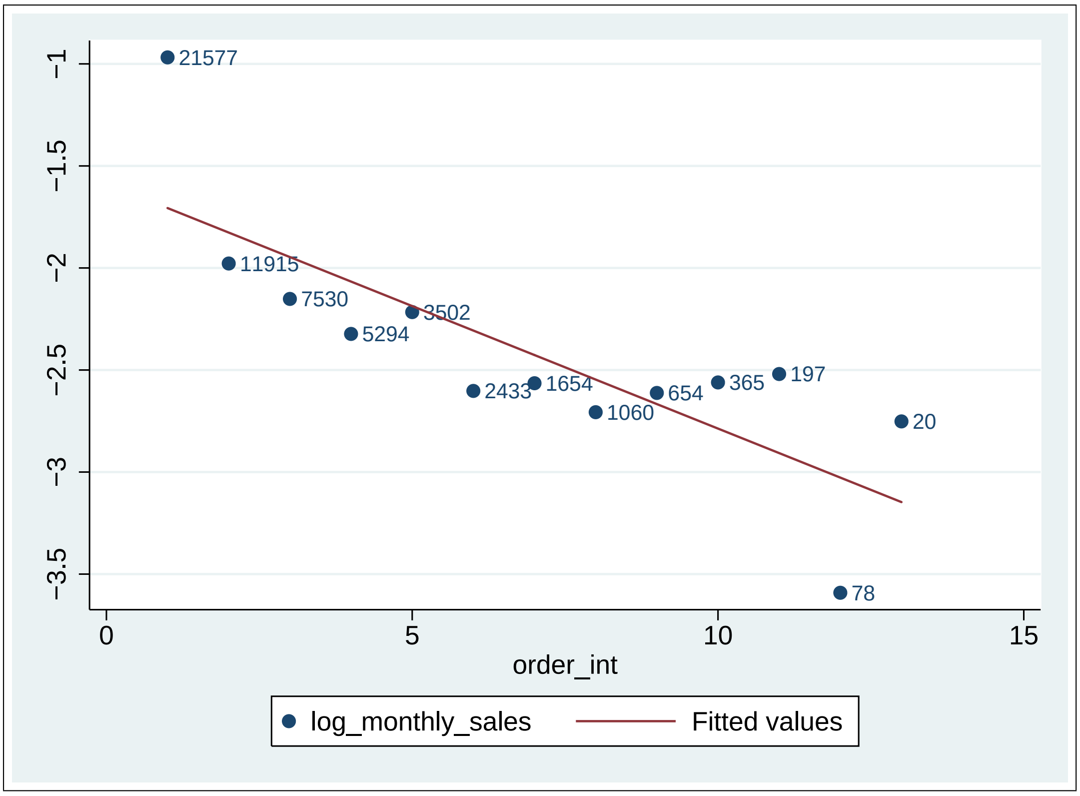


Figure 2. Median monthly product sales vs. order of introduction conditional on the number of months the product has been exported.

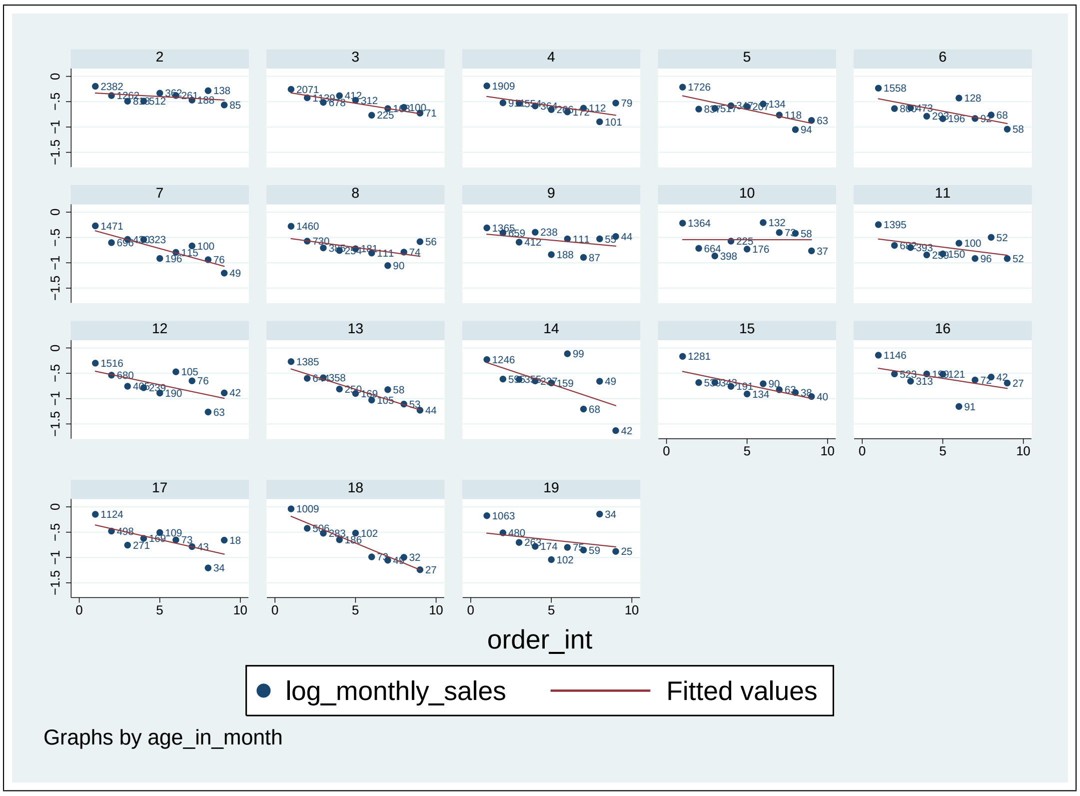


Figure 3. Evolution of the variance of the firm's beliefs about its brand effect as a function of the number of products introduced.

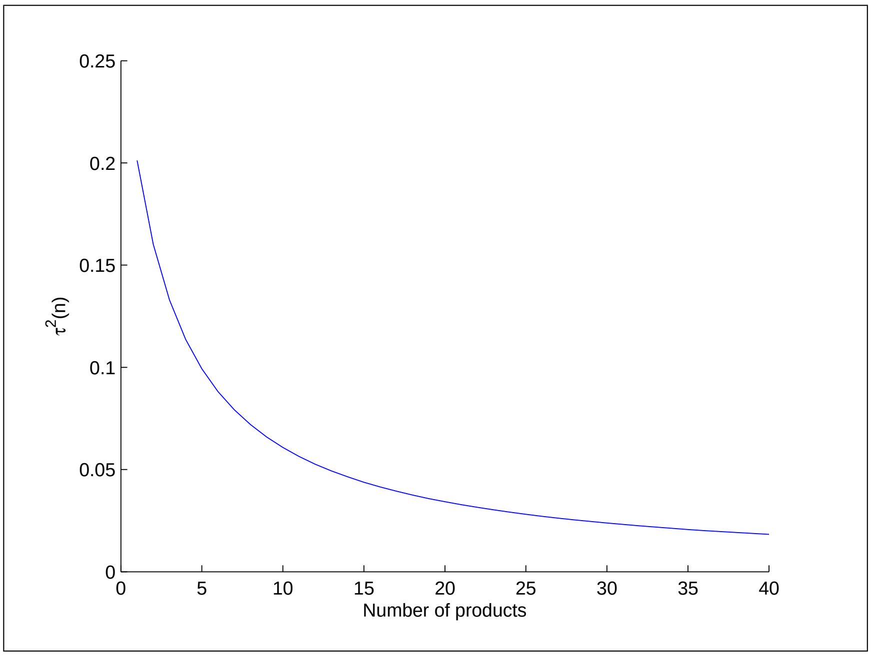


Figure 4. Distribution of firms over the number of products conditional on the brand effect. Baseline ($\psi = 0.2$). Cohort of firms that entered exporting in the same year, i.e., in year 1.

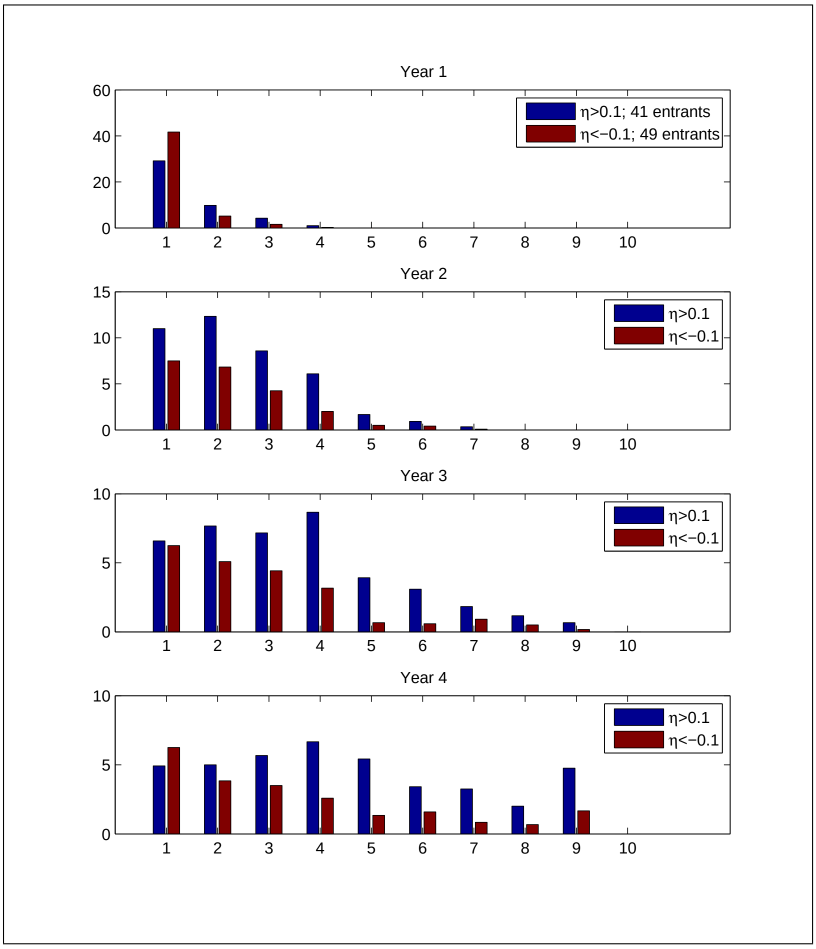


Figure 5. Distribution of firms over the number of products conditional on the brand effect. Baseline case ($\psi = 0.89$). Cohort of firms that entered exporting in the same year, i.e., in year 1.

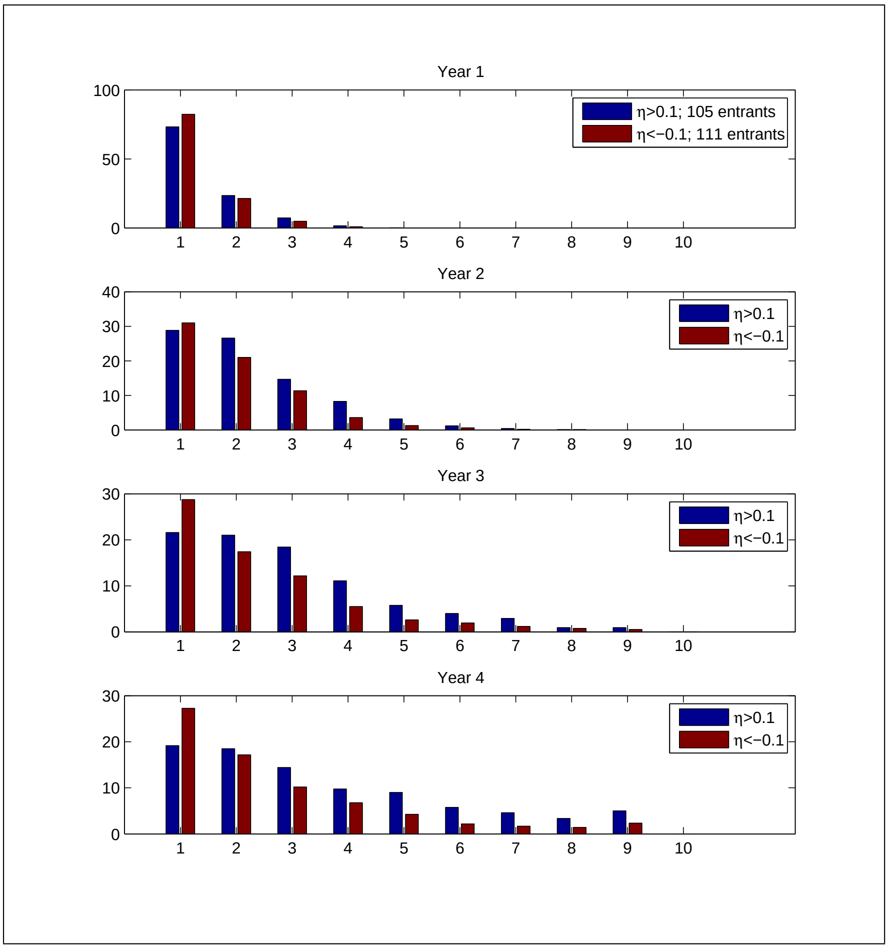


Figure 6. Distribution of firms over the number of products conditional on the brand effect. No learning ($\psi = 0.89$). Cohort of firms that entered exporting in the same year, i.e., in year 1.

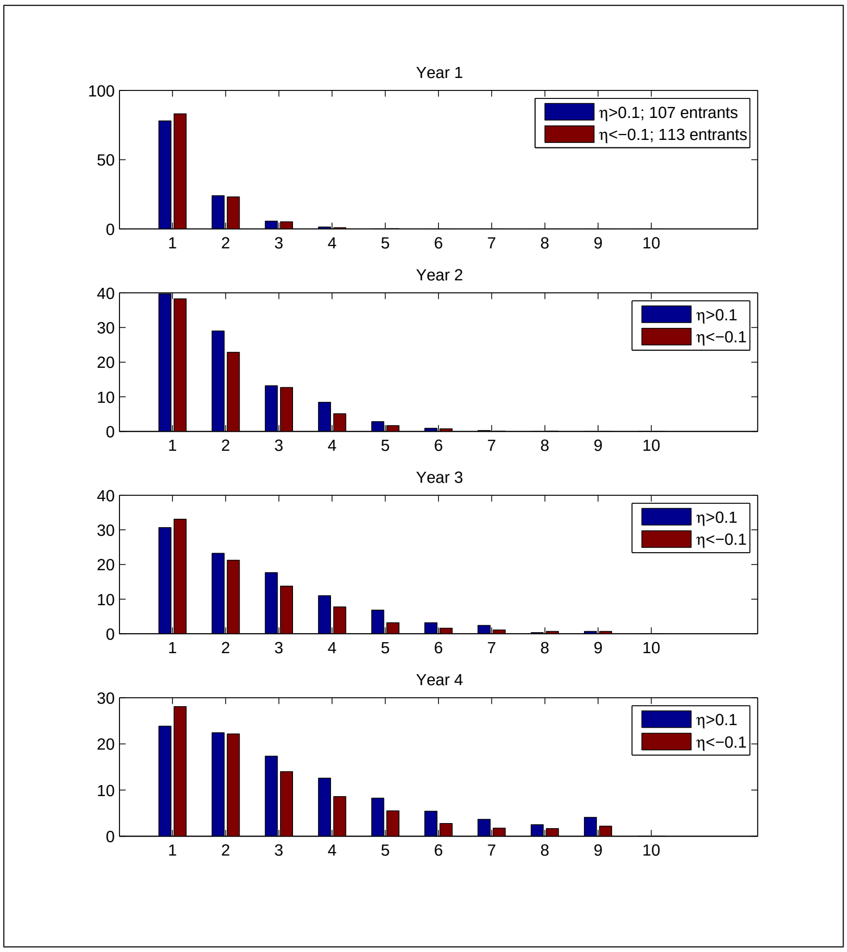


Figure 7. Distribution of firms over the number of products conditional on the brand effect. Full information ($\psi = 0.89$). Cohort of firms that entered exporting in the same year, i.e., in year 1.

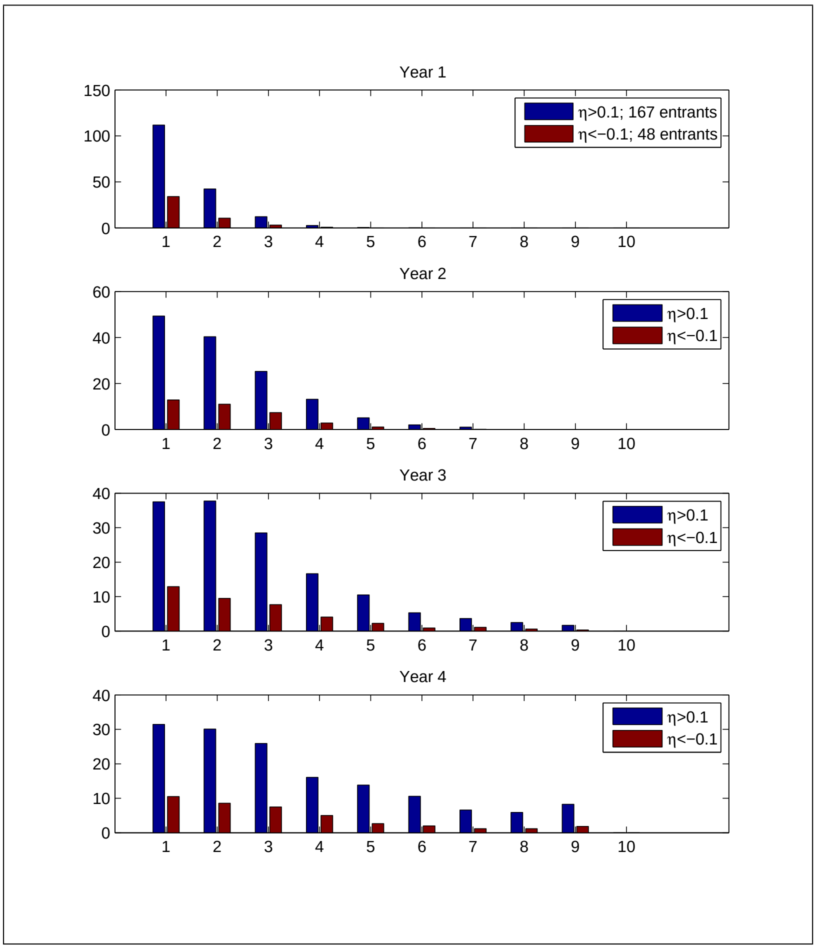


Figure 8. Share of products dropped relative to the total number of products in the four scenarios.

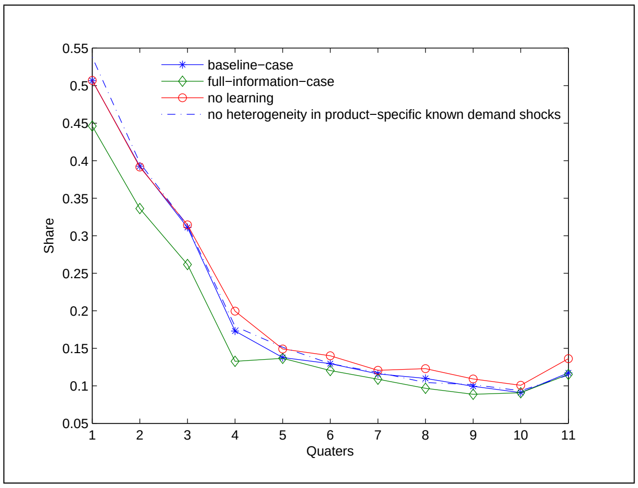


Figure 9. No "known" demand shocks vs. baseline case. Distribution of firms over the number of products. Cohort of firms that entered exporting in the same year, i.e., in year 1.

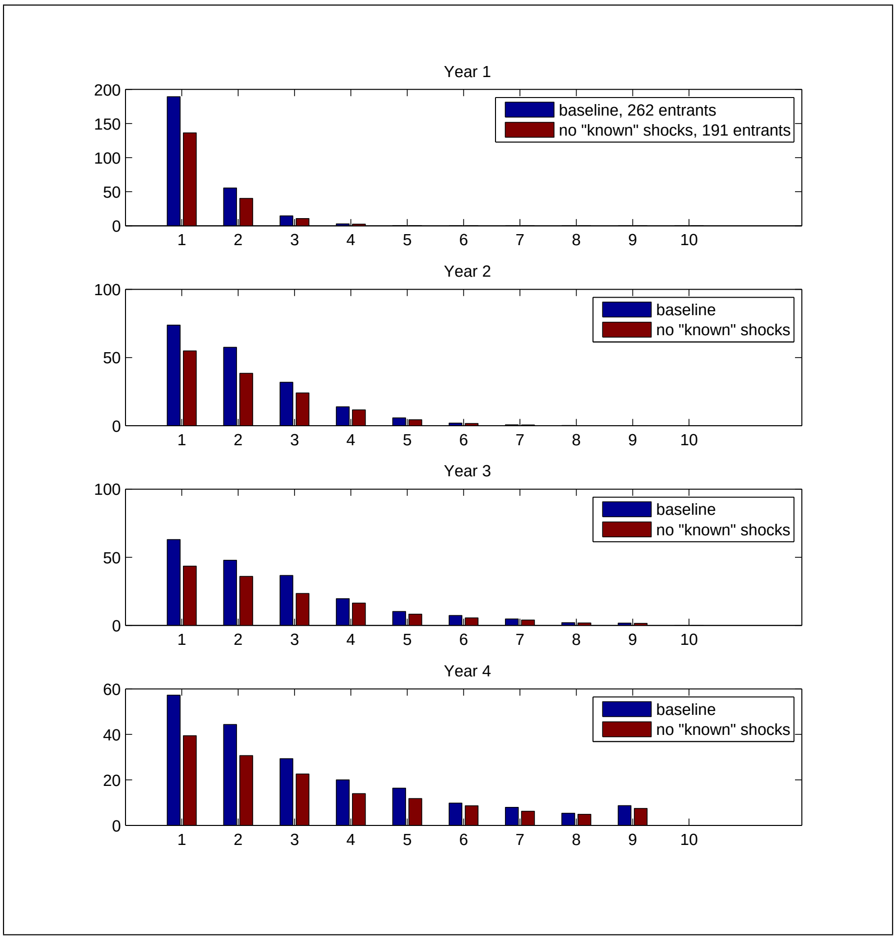


Figure 10. Aggregate sales of the cohort.

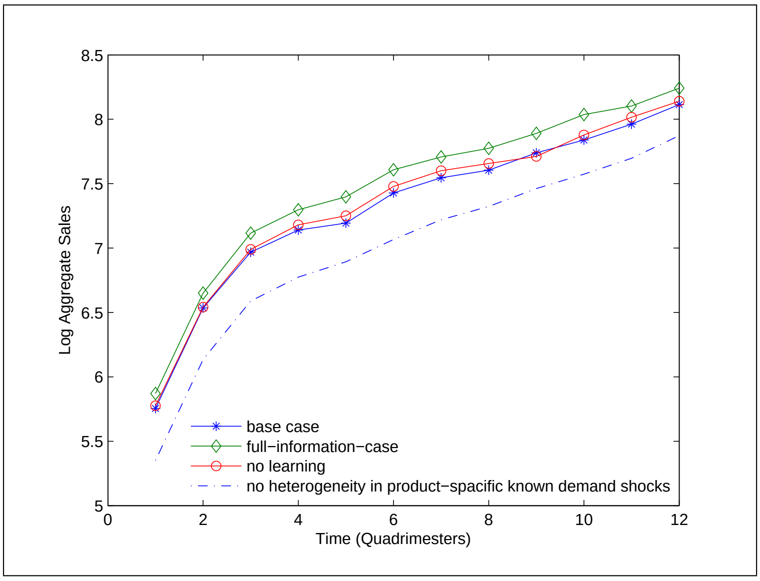


Figure 11. No learning vs. baseline scenarios ($\psi = 0.89$). Distribution of firms over the number of products. Cohort of firms that entered exporting in the same year.

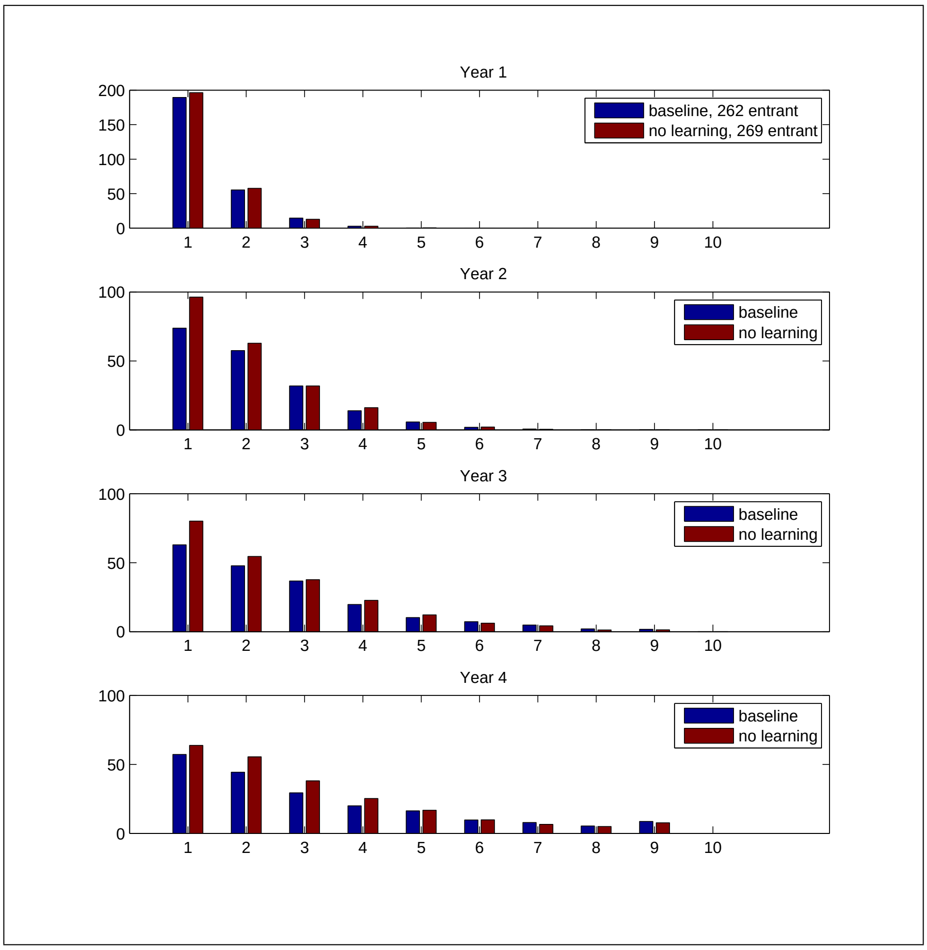
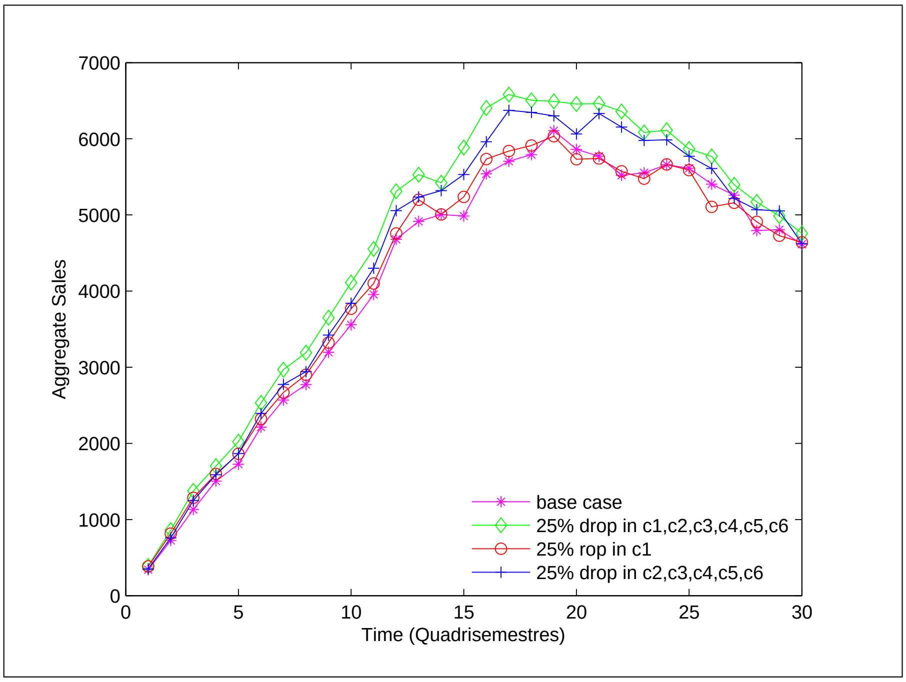


Figure 12. Quadrisemestre aggregate sales. Decreasing the cost of introducing new products.



9 Appendix 3. Empirical Regularities in the Plastics Industry.

In this appendix we verify the data patterns that we have demonstrated in Section 2 for the subsample of exporters operating in the plastics industry between 2001 and 2004. Here we use the HS 4-digit definition of a product.

First, we consider the relationship between the product's order of introduction and its sales. We regress the log of firm-product monthly sales on the the number of products that the firm has, the product's order of introduction, and its tenure in months. We also incorporate year fixed effects. Table 11 presents the results. As in Table 2, the effect of a product's order of introduction on it's sales is negative and significant. Similarly, the coefficient on a product's age is positive and significant. The coefficient on a number of products a firm retains its positive sign.

Table 11: Product monthly sales and its order of introduction. Plastics industry.

Dependent variable: Monthly product sales

	REG1	REG2
# of products	.058***	0.047***
	(.003)	(0.003)
Order of introduction	-.223***	-0.161***
	(.007)	(0.009)
Age in months		0.047***
		(0.004)
N	34521	34521
r^2	0.026	0.03

Standard errors in parentheses.

$^{*}p < 0.05$, $^{**}p < 0.01$, $^{***}p < 0.001$

Includes year fixed effects.

Next we verify for the plastics industry that a large share of new products are discontinued shortly after their introduction due to low sales. Table 12 is the analogue of Table 3. It shows how the number of products in a cohort of products introduced in 2002 by firms that started exporting in 2001 evolves over the years for the subset of firms in the plastics industry. The second column shows that 40% of products introduced in 2002 are dropped by firms that continue to export in 2003. In the following year attrition drops to 30%. Table 13 shows average monthly product sales as the cohort ages. Sales among products dropped by firms that continue exporting are lower than for products that are still exported in the following year.

Table 12: Evolution of the number of products introduced in 2002 by firms that started exporting in 2001. Sample includes only firms that operated in the plastics industry in 2001-2004. HS 4-digit category.

t	Number of products introduced in 2002 and sold in t.	Sold in t and $t+1$ (% of the total).	Sold in t only, the carrying firm is present in $t+1$ (% of the total).	Sold in t only, carrying firm exits in t (% of the total).
2002	450	245(54%)	181(40%)	24(5%)
2003	245	147(60%)	74(30%)	24(9.8%)

Table 13: Average monthly sales per product among products introduced in 2002 by firms that have started exporting in 2001. Sample includes only firms that operated in the plastics industry in 2001-2004. HS 4-digit category.

t	Mean sales for products introduced in 2002 and sold in t	Mean sales for products in t, which exit in $t+1$ conditional on firm staying.	Mean sales for products in t, which exit in $t+1$ along with the firm
2002	0.56	0.27	0.46
2003	0.87	0.21	0.44

In Table 14 we verify the pattern again by looking at the evolution of the number of products at the half year intervals. Here we consider the cohort of products introduced in the first half of 2002 by firms that started exporting in the first half of 2001 in the plastics industry. A similar pattern emerges.

Table 14: Evolution of the number of products introduced in the first half of 2002 by firms that started exporting in the first half of 2001. The sample includes only firms that operated in the plastics industry in 2001-2004. HS-4 digit category.

t	Number of products introduced in 2002, 1-6 and sold in t.	Sold in t and $t+1$ (% from total).	Sold in t only, carrying firm is present in $t+1$ (% from total).	Sold in t only, carrying firm exits in t (% from total).
2002,1-6	182	104(57%)	75(41%)	3(2%)
2002,7-12	104	73(70%)	24(23%)	7(7%)
2003,1-6	73	56(77%)	16(22%)	1(1%)
2003,7-12	56	48(86%)	7(13%)	1(2%)
2004,1-6	48	40(83%)	5(10%)	3(6%)

Finally, we consider wether firms that have introduced successful products in the past are more likely to expand further. To this end we estimate the binary logistic probability model of the firm introducing a new product line as we did in Section 2. Here we omit the total exports variable since it was not significant in the full sample of the universe of Chinese exporters. Table 15 presents the results. The coefficient on the proxy for the history of failures has a negative sign confirming that firms indeed take into account their history. The standard error for the coefficient is higher than what we saw when we considered the universe of Chinese exporters in Table 5. This is not surprising since in the plastics industry we only have 645 observations and the panel is shorter.

Interestingly the coefficient on the number of products a firm has attempted is larger and more significant compared to the Table 5. The coefficient on the age of a firm has a positive sign. The variable age in this regression cannot be directly compared to the age variable in the analogous regression in Section 2 because there we have considered multiple cohorts of firms, while here we have only firms that started exporting in 2001.

Table 15: Logit, FE. Plastics industry.

Dep.var. takes value 1 if a firm has introduce at least one new product in a year, 0 otherwise.	
Share of products dropped after one year in total number of products ($F\_rate_{f(t-1)}$)	-1.146^{**}
Number of products ($n_{f(t-1)}$)	-2.116^{***}
Av. sales per product ($Av.sales_{f(t-1)}$)	$-.007$
Age	1.061^{***}
Obs.	645

Standard errors in parentheses.
$^{*}p < 0.05$, $^{**}p < 0.01$, $^{***}p < 0.001$

Overall, we conclude that the patterns we have documented for the universe of Chinese exporters are present in the subsample of exporters operating in the plastics industry.

10 Appendix 4. Standard errors.

We compute standard errors numerically according to the formula:

$$Var(\hat{\theta}) = \left(\frac{\partial m(\hat{\theta})}{\partial \theta}' W \frac{\partial m(\hat{\theta})}{\partial \theta} \right)^{-1} \frac{\partial m(\hat{\theta})}{\partial \theta}' W [m(\hat{\theta}) m(\hat{\theta})'] W \frac{\partial m(\hat{\theta})}{\partial \theta} \left(\frac{\partial m(\hat{\theta})}{\partial \theta}' W \frac{\partial m(\hat{\theta})}{\partial \theta} \right)^{-1} \tag{13}$$

where $m(\hat{\theta}) = \overrightarrow{\beta_s(\hat{\theta})} - \overrightarrow{\beta_d(\theta_{true})}$, and W=$Var^{-1}(\overrightarrow{\beta_d(\theta_{true})})$