Multi-product Exporters: Facts and Fiction
JEL: D24, L11, F14

Lena Sheveleva[1]

June 28, 2020

[1]The author acknowledges the useful discussions, advice and support from Milagro Rogriguez-Saborio, Kala Krishna, Peter Neary, Markus Eberhardt, James Tybout as well as her colleagues at Cardiff Business School.

Abstract

In this paper we contribute to the empirical literature on multi-product exporter by highlighting that data regularities which involve order statistics, such as best/least selling products, systematically vary with the number of products. Our insight is that instead of comparing outcomes of interest between exporters of different scope one should compare the data to patterns that would arise if the sales (and other outcomes of interest) were randomly distributed. We propose an intentionally stark statistical model of multi-product exporters where exporter outcomes are random draws from a distribution. We show that some of the patterns used to characterize multi-product exporters turn out to be consistent with a wide variety of models, including the ones that do not feature firm-specific productivity or variation in fixed costs. For example, in the data exporters with six or more products sell 8 times as much as single product exporters in their top selling products - an empirical fact that has motivated the hypothesis that large scope exporters are more productive. If large exporters drew sales from the same distribution as single product exporters they would sell more than 6 times in their best selling products. In other words, differences between how much large and small scope exporters sell in their best selling products is driven by aggregating the data across the different number of products and not y differences in firm attributes.

Finally, we extend the model to the multi-country set-up and explore whether popular measure of sales concentration within a firm, the log-ratio of sales between the best and second best selling product, is suitable to study the response of multi-product exporters to changes in trade policy and competitive environment. We show under what conditions the log-ratio of sales biases the estimated firm response.

1 Introduction

Across the world multi-product firms dominate international trade flows. Famously, Bernard et al. (2010) document multi-product firms account for 98% of the U.S. manufacturing exports value. In the light of this empirical importance, researchers are interested in understanding why multi-product exporters arise and how they respond to changes in the international trade environment. In recent years there has been some progress in answering these questions both on the empirical and theoretical fronts. For example, Helpman (1985),Ju (2003),Allanson and Montagna (2005),Nocke and Yeaple (2015),Bernard et al. (2011),Eckel and Neary (2010), Dhyne et al. (2017),Arkolakis et al. (2014),Bernard et al. (2010) are just few of the papers that explore how to model multi-product firms in international trade and how they respond to changes in international trade environment. In part, this progress has been facilitated by a number of empirical regularities documented using newly available firm-product level data.

In this paper we take a step away from behavioral models of multi-product exporters and instead focus on empirical regularities that underpinned those models. We aim to understand which ones are informative about the economic forces that drive multi-product exporters and which arise as a result of randomness and aggregation. The original motivation for this research question comes from the work of Ellison and Glaeser (1997) who show that data patterns and predictions of economic models should be compared to patterns that would arise if the outcomes of interest were randomly distributed rather than a uniform pattern or an absence of a pattern. In the context of multi-product exporters we show that some notable differences between large and small scope exporters discussed in the literature would arise even if product sales were randomly distributed across firms and products simply because we aggregate data over different number of products.

This paper is also close in spirit to the work of Armenter and Koren (2014) who highlight that many facts about the extensive margin of trade are consistent with a surprisingly large class of models because of the sparse nature of trade data. In relation to multi-product firms they show that the bins-and-balls model quantitatively reproduces the frequency of the single product exporters we observe in the data. We focus on the differences between large and small scope exporters on the intensive margin and ask which empirical regularities are consistent with a wide variety of models and which can be used to discriminate between alternative models of multi-product firms.

Take the observation that large scope exporters sell much more in their top selling products than their small scope counterparts. It has been documented across many countries and industries and helps explain why multi product exporters dominate international trade flows in terms of sales volume. In this paper instead of comparing sales of the best selling products between large and small scope exporters we compare sales that we observe conditional on scope to sales that would arise if sales were randomly distributed across firms and products. We find no evidence in the data to support that large scope exporters sell their best selling products in greater amounts than the small scope exporters beyond what we would expect if sales were randomly distributed across firms and products. The pattern that large scope exporters sell much more in their top selling products than their small scope counterparts is just an artifact of aggregating data to the firm-level across different number of products. Conditional on the number of products larger scope exporters sell no more than small scope exporters. In other words, if large scope exporters drew sales from the same distribution as single product exporters they would sell more in their top selling products than we observe them to sell in the data.

More generally to enable comparison between the large and small scope exporters one has to aggregate product level data across different number of products to the firm level. Examples include total exporter sales, average sales per product, mean sales at rank (average sales of the best/least selling products), ratio of sales between the best and the second best selling product.

Our insight is that this process of aggregation itself gives rise to differences between large and small scope exporters.

Consider the following three facts that have been documented across a large number of data-sets and countries[1].

1. Large scope exporters sell more in their best selling products than small scope exporters.

2. Large scope exporters sell less in their least-selling products than small scope exporters.

3. Average sales per product are non-monotone with scope.

These facts have been used to motivate many models of multiproduct exporters where a single firm attribute, usually a firm productivity, drives both scope (the ability to produce many products) and scale (the ability to produce at scale due to lower marginal cost). E.g. Bernard et al. (2011) ,Bernard et al. (2010), Arkolakis et al. (2014). To set ideas straight, consider a Melitz's style model where firms face random product specific demand shocks and where a single firm productivity translates into a marginal cost parameter common across all products of the firm. Lower marginal cost plays a dual role. First, it encourages high productivity firms to export more products due to higher expected profits from new products. Second, thanks to lower marginal cost large scope exporters sell their best selling products in larger amounts than small scope exporters. Furthermore, due to lower marginal costs large scope exporters can profitably export products with low demand shocks that sell in tiny amounts and so their sales in the least selling products are lower than for smaller scope exporters. Average sales per product are then indeterminate as a function of scope due to composition effects.

In this paper we show that there is no evidence in the data to support that products of the larger scope exporters sell in larger amounts than products

[1] Arkolakis et al. (2014) and Timoshenko (2015) document these patterns for Brazil, Bernard et al. (2011) and Bernard et al. (2010) for the US

of the small scope exporters beyond what we would expect if sales were randomly distributed across firms and products. In other words, the three facts above arise even in the absence of ex-ante heterogeneity: if all firms have the same marginal cost per product and product demand shocks come from the same distribution. To see the intuition, consider a two product firm and a ten product firm. Suppose that sales per product are just iid draws from the same distribution across firms and products, then the expected value of the maximum of ten products is higher than the maximum over two products. The reverse is true for the minimum. Hence it is possible that larger scope exporters have higher sales in their top-selling and lower in their least-selling products relative to the smaller scope exporters mechanically.

Formally, Facts 1 and 2 are examples of order statistics, i.e., average sales across products at a given rank, and are systematically related to the number of products for which they are calculated. This paper highlights that instead of comparing sales at rank between large and small scope exporters, one should compare the observed or predicted patterns to those that would arise if product sales were randomly distributed. Indeed, if more productive firms select into exporting more products then we should see that the large scope exporters both sell more in their top-selling products than if their sales were driven by randomness and more than their smaller scope counterparts. Similarly, if large scope exporters benefited from the economies of scope and faced lower fixed costs of exporting, they would sell less in their least-selling products both compared to the small scope exporters and less than if their sales were random across firms and products.

While in theory randomness and aggregation produce the observed patterns, whether they are sufficient to replicate them quantitatively without appealing to firm productivity and other economic forces is an empirical question. To answer it we formalize randomness in an intentionally stark statistical model of multi-product exporters. In this set up the number of exported products per firm is treated as exogenous and product revenues are realizations from some distribution F(). In our empirical implementation F() is either Pareto or log normal. We show that this statistical model can replicate the positive (negative) relationship between the mean sales in best (least) selling products and firm scope remarkably well. Log-normal provides a better fit than does Pareto which tends to over-estimate how quickly sales at a given rank increase/decrease with scope. The poor empirical performance of the Pareto stems from its failure to approximate the left tail of the sales distribution and echoes the findings of Head et al. (2014), Bee and Schiavo (2018),Fernandes et al. (2015) which show the log-normal distribution provides a better fit for the distribution of the exporter sales than the Pareto.

Our empirical results show that facts 1,2 and 3 can be quantitatively replicated without appealing to firm specific productivity as a single driver of scope and scale. Put another way, the firm scope is a sufficient statistic for differences between large and small scope exporters on the intensive margin. So, a model that can successfully replicate firm scope will also replicate observed differences between large and small scope exporters. In this regard, our work also speaks to the literature on the role of firm productivity in determining scope and scale of multi-product firms. Our results echo **?** who showed that scope and measured productivity are only imperfectly correlated and it's heterogeneity in fixed costs of exporting that is the primary driver of exporter scope. This has profound implications for the theory of the firm and policy implications. If there is no evidence to support thinking of firm specific productivity as a single parameter that drives both scope and scale of the exporters then the emphasis should be on other sources of firm-heterogeneity that drive exporter scope. For example, R&D that targets product innovation Dhingra (2013), or access to inputs similar to core competencies of the firms Boehm et al. (2019), market-entry cost heterogeneity Ju (2003),Eckel and Neary (2010).

Finally, we extend the statistical model to the multi-country setup to see if it can reproduce other patterns that rely on order statistics. We focus on a popular regularity that in more competitive markets multi-product exporters tend to have their sales more concentrated in the top-selling products. A measure of sales concentration that has been widely used in the multi-product literature is the ratio of sales of the best selling to the second best selling product. One reason behind its wide spread use is that theoretical models predict the ratio as a function of variables that proxy market competitiveness at a destination. For example, Mayer et al. (2014) show that when demand is linear and firms face variable mark-ups multi-product firms adjust their product mix by reallocating resources across products. They expand production of their core (low cost and high mark-up) varieties at the expense of the peripheral ones thus increasing the concentration in the core varieties. One of the ways in which they test their prediction is by regressing the log-ratio on the destination market size, market supply potential, bilateral trade costs and other proxies of market competitiveness. Using the estimated impact of market size on log-ratio they calculate that the within-firm resource reallocation contributes nontrivial 19% to aggregate productivity.[2]

Our multi-destination extension of the statistical model highlights two things. First, one should be careful when interpreting the regression coefficients as evidence of product mix adjustments. This is because unless the sales are Pareto distributed the log-ratio, which is the function of two order statistics, systematically depends on the number of products. If sales are log-normally distributed, for example, the expected value of the ratio decreases with the number of products per firm. So, if firms export fewer products to more distant or higher foreign supply potential destinations the ratio will be higher there even if sales are randomly distributed as is in our statistical

[2]Mayer et al. (2014) provide robustness checks in the appendix where they show that concentration increases is captured by Theil and Herfindalh indices as well. However, their discussion of the economic implications and the main results in the paper are based on the log-ratio measure of concentration which is vulnerable to the bias due to the unaccounted number of products.

model. Second, our quantitative results show that the regression coefficient on the market size underestimates the effect of market size on changes in product mix. On the other hand, market supply potential or bilateral distance which are associated with smaller exporter scope will have coefficients biased upwards. Intuitively, the impact of factors that are associated with firms exporting more products to a destination will be underestimated while the effect of factors associated with fewer exporters will be over-overestimate. Our results thus suggest that the within-firm resource reallocation may contribute more than 19% to aggregate productivity that Mayer et al. (2014) calibrated.

In this paper we contribute to the empirical literature on multi-product exporter by highlighting that empirical regularities which involve order statistics, such as best/least selling products, systematically vary with the number of products. Our insight is that instead of comparing sales at rank between exporters of different scope one should compare sales that are observed to those that would arise if firms drew sales from the same distribution. Some of the pattern used to characterize multi-product exporters turn out to be consistent with a wide variety of models, including the statistical model where all firms are ex-ante identical and differ ex-post only due to the number of products. For example, in the data exporters with six or more products sell 6 times as much as single product exporters in their top selling products. If large exporters drew sales from the same distribution as single product exporters they would sell more than 6 times in their best selling products. In other words, differences between how much large and small scope exporters sell in their best selling products is driven by aggregating the data across the different number of products. Finally, we show that the ratio of sales between the best and the second best selling product increases with scope of the firm for which it is calculated and that this property can bias the results of the regressions where it is used as the dependent variable.

One should, however, be careful not to interpret our results as evidence against behavioral models that feature firm productivity or resource reallocation in response to trade shocks. Beyond the well-documented patterns that we focus on here, there is work that provides direct evidence on these channels. For example, Dhingra (2013) find direct evidence that firms invest in cost-cutting technology for their core products in response to trade liberalization in Malaysia and drop peripheral products. Dhyne et al. (2017) use a novel multi-product firm production function estimation approach to estimate technical efficiencies of individual products. They find that firms are more efficient at producing core products and respond to competition by focusing more on their core products. Rather one should see our results as a study into which data moments are informative about multi-product firm behavior and which should be considered with extra care.

The paper proceeds as follows. Section 2 lays out the model and discusses the intuition behind it. Section 3 presents the calibrated results, section 4 extends the baseline model to multi country set-up and section 5 concludes.

2 Theoretical underpinnings

We now present a statistical model of multi-product exporter outcomes. We begin with a single destination version and later extend it to the multi-destination set-up.

A multi-product exporter indexed by f is a collection of products K that a firm exports in a given year. Products within a firm are indexed by $k \in \{1, ..., K\}$. The number of products a firm exports (K) is treated as exogenous and in empirical applications will be taken from the data. Product sales S_{fk} are iid draws from some continuous distribution $F()$ with a non-negative domain.

In this stark set-up, any differences between large and small scope exporters on the intensive margin are driven by aggregating random draws over the different number of exported products to the firm level. So by comparing the observed patterns to the patterns predicted by the statistical model, we will be able to identify the data patterns that are genuinely informative about the economic forces driving the intensive margin from the ones that reflect randomness and aggregation.

This stark model has four main predictions about the distribution of product sales within a firm. When products within a firm are ranked by their contribution to total sales from the best selling to the least selling the model predicts:

Prediction 1. *Bigger scope exporters sell more in their best selling products than smaller scope exporters.*

Prediction 2. *Bigger scope exporters sell less in the lowest ranked products than smaller scope exporters.*

In the model, the large scope exporters get to draw more sales shocks (one for each product) than smaller scope ones. As long as the draws are iid across firms and products, the maximum of the larger number of draws is bigger than the maximum over the smaller number of draws. The reverse is true for the minimum. We will demonstrate that both predictions are independent from the assumptions about the distribution and find strong support in the data[3].

While in our model the two facts are solely driven by randomness they are also (at least qualitatively) consistent with the models of multi-product firms where firm productivity drives both scope and scale. Faced with random demand shocks as in Bernard et al. (2011) high productivity firms have higher expected profits from a new product and so select into more products. These more productive exporters not only sell their best selling products in large amounts but also due to their high productivity are able to make profits from products that face low demand. The extent to which randomness

[3]See Arkolakis et al. (2014), Timoshenko (2015), Bernard et al. (2011) and Bernard et al. (2010)

can explain variation in the data without appealing to single attribute firm specific heterogeneity is an empirical question. We are going to address it in the next section by estimating the model from the data on Chinese exporters to the US.

Prediction 3. *Average sales per-product are constant with scope.*

Evidence on the relationship between average sales per-product and firm scope is mixed. As a rule average sales per products is not monotone with scope (see Bernard et al. (2011), Arkolakis et al. (2014)). Many models of multi-product firms avoid making predictions about average sales because they reflect composition effects (i.e. large scope exporters sell their products in both extraordinary large and small amounts). Yet, our results suggest that variation of average sales with scope cannot be replicated with randomness alone and so is a useful moment to differentiate among alternative models of multi-product exporters.

Prediction 4. *The expected value of the log ratio between sales of the best and the second best selling product is constant with scope when sales are distributed Pareto and decreases monotonically with scope when sales follow log-normal, Weibul or exponential distributions.*

In the appendix A, we present the analytical proofs for Pareto, Weibul and exponential distributions. We use numerical integration to establish the result for the log-normal distribution.

The log ratio of the best selling to the second best selling product is often used as a measure of sales concentration within a firm. It is expected to capture how firms respond to changes in competition by adjusting quantities and prices across the product range. We highlight that there is also a mechanical relationship between the number of products exported and log-ratio unless sales are Pareto distributed. This matters because within firm reallocation of resources usually coincides with adjustments on the extensive margin. Faced with tougher competition firms drop products which would have the effect of increasing the ratio even if firms did not respond optimally on the intensive margin. Hence, the effect of competition on the skewness of sales when measured using the ratio is likely to be overestimated. We explore this issue in detail when we discuss the extension of the model to the multi-destination setting in section 3.

3 Empirical Test of the Model

In this section, we test the predictions of the statistical model in the data. When we compare sales at rank in the data to the sales at rank that arise from the statistical model we show that large scope exporters in fact sell less in their top-selling products than if they drew their sales from the same distribution as single product exporters (Prediction 1). Similarly, once we take into account aggregation large scope exporters sell no less than small scope exporters in their least-selling products (Predictions 2). The model is less able to replicate the variation in average sales per product (Prediction 3). In terms of the log ratio of sales between the best selling to the second best selling product the model replicates the rate at which ratio decreases with scope but underestimates average ratio at a given scope (Prediction 4).

3.1 Estimation Details

In testing the predictions of the statistical model we use data on firm-product sales available at the HS-6 level for Chinese exporters to the US in 2003. In Table 1 we verify the multi-product exporter facts documented for other countries in the Chinese data. It splits the sample based on the number of products each firm exported and reports average total exporter sales, average sales of the best and least selling products for each group. While small in number large scope exporters dominate the export market in terms of their sales. Single product exporters, for example, account for 41 percent of exporters while exporters with 5 or more products account for only 20 percent of all exporting firms. Yet, exporters with 5 or more products sell 10 times as much as single product exporters. Exporters with 5 or more products sell almost 8 times as much as single product exporters in their best selling products. They also sell drastically more in their least selling than small scope exporters.

Table 1 – Summary Statistics for the Chinese Exporters to the US.

Scope	1	2	3 - 4	5+
Exporter Share with Scope	41	20	19	20
Mean Total Firm Sales	907	1,826	3,141	10,185
Mean Top Ranked Product	907	1,679	2,689	7,184
Mean Bottom Ranked Product	907	147	45	8

The table splits the sample of exporters from China to the US by the number of products that they exported in 2003 and reports share of exporters with 1,2,3-4,5 and more products, average total firm sales, mean sales of the best and least selling products.

To test the predictions of the statistical model we first estimate the distribution of product sales $F()$ and then use the estimated distribution to simulate the moments about which the statistical model makes predictions. We then compare the simulated moments to the data and infer which patterns can be explained with randomness and aggregation alone, and which ones are genuinely informative about the economic forces that produced the data.

To take into account the systematic variation in products (at the HS group level), we adjust annual sales data used in Table 1 relative to the average annual sales across all firms selling the same product category in the US.

We use the log-normal and Pareto distributions to approximate the empirical distribution of product sales and estimate the parameters of each distribution using either Maximum Likelihood (ML) or the Simulated Method of Moments (SMM) approach. With the latter we target sales at rank statistics. The ML approach is a more stringent test of the model since we do not directly target the patterns we want to explain in the estimation. If the distribution $F()$ accurately describes the data on sales both approaches should yield similar results.

3.1.1 Log-normal Distribution

If sales S_{fk} follow the log-normal distribution with the location and shape parameters μ and variance τ, the natural logarithm of sales $ln(S_{fk})$ follows the normal distribution with mean μ and variance τ. The top panel of Table 2 reports the ML and SMM estimates of parameters from fitting the normal distribution to the data on the natural logarithm of sales. The bootstrapped standard errors are in parenthesis and indicate that estimates are statistically significant. Both approaches yield very similar results and both closely track the actual distribution of the log sales in Figure 1. The ML approach yields only slightly thicker tails due to higher estimated variance. This is what one would expect if the theoretical distribution approximates the empirical one reasonably well.

Predictions of the statistical model that we have set out to test involve order statistics, i.e., best, second best, least selling products. The predictions of the statistical model will only be able to quantitatively match the data if the model can replicate the correlation between order statistics and exporter scope that we observe in the data. Figure 2(a) compares the mean sales at rank conditional on firm scope in the data and simulated from the model at the ML and SMM estimates. Each panel looks at average sales at rank for firms with the same scope. Take panel three, for example. Here for each firm that exported three products in 2003 we rank firm's exported products by their contribution to the firm's total sales. So the product with the highest contribution is assigned rank one. We report average sales for products at the same rank among three product firms calculated from the actual data as well as data simulated from log-normal distribution estimated using the ML and SMM. In each case the order statistics track the actual data remarkably well.

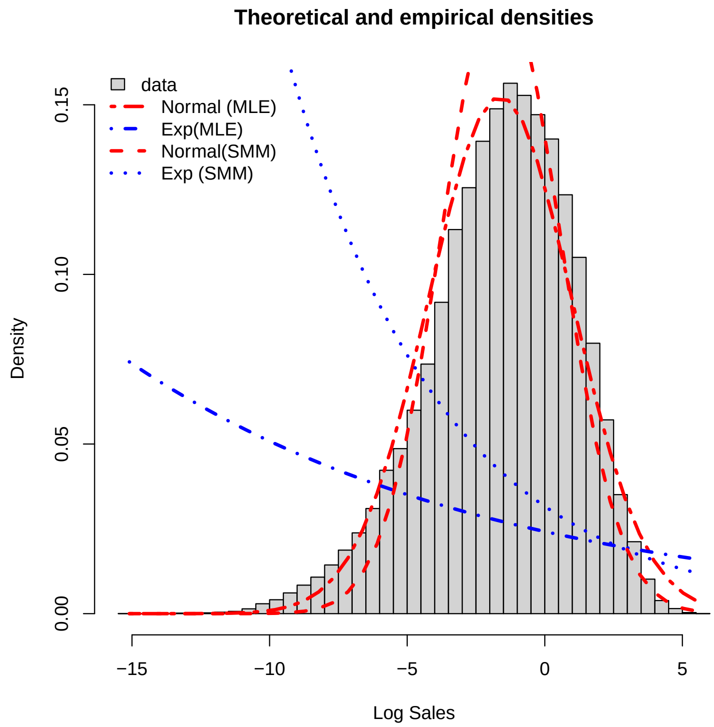


Figure 1 – The figure shows the estimated density of log-sales under the hypotheses that sales are Pareto or log-normally distributed using the maximum likelihood and the simulated method of moments approaches against the empirical distribution of the log-sales.

Table 2 – Distribution parameter estimates.

	MLE	SMM
	Log-normal	
μ	-1.6 ***	-1.6 ***
	(0.0155)	(0.00018)
σ	2.62 ***	2.12 ***
	(0.008)	(0.00024)
Log-likelihood	-104,570.4	
	Pareto	
ν	0.074 ***	0.13***
	(0.00043)	(0.0007)
α	0.000059	0.000003 ***
		(0.0000006)
Log-likelihood	-684,81.37	
	Log-normal Single-product exporters sub-sample	
μ	-1.73 ***	
	(0.02)	
σ	2.62 ***	
	(0.014)	
Log-likelihood	-41,704	

Bootstrapped standard errors in the parenthesis.
* p<0.10, ** p<0.05, *** p<0.01

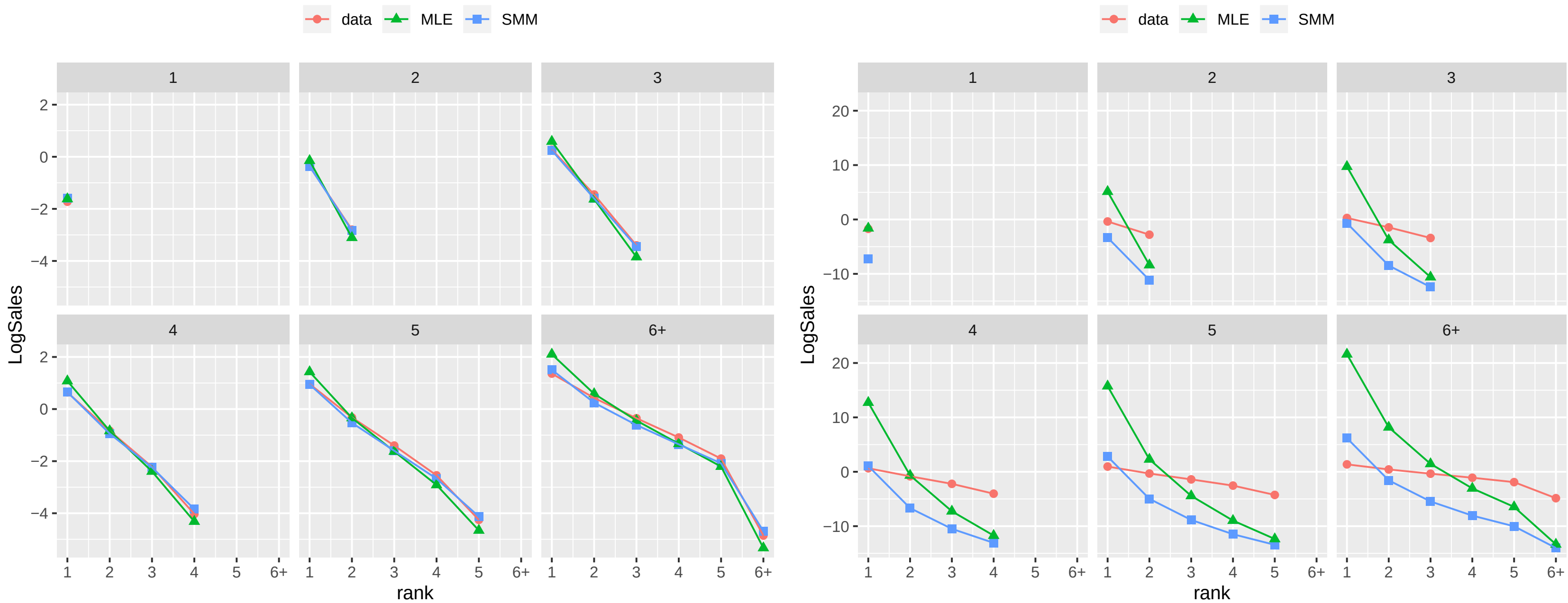
3.1.2 Pareto Distribution

If sales S_{fk} are Pareto distributed with a location parameter α and shape parameter ν then $ln(S_{fk}) - ln(\alpha)$ is distributed exponentially with rate parameter ν. To estimate the parameters of the Pareto distribution with the maximum likelihood we first calibrate the location parameter α to the smallest sale value in the sample, $\hat{\alpha} = min(S_{fk})$, and then estimate the rate parameter ν by fitting the exponential distribution to data on $ln(S_{fk}) - ln(\hat{\alpha})$. With the SMM approach we jointly estimate α along with ν by targeting sales at rank conditional on the number of exported products. The second panel of Table 2 reports the estimates based on the SMM and MLE approached. The estimates are quite different which indicates that the distribution is likely misspecified. Figure 1 compares the estimated exponential distribution with the the actual distribution of log-sales and the fitted normal density. Regardless of how parameters are estimated Pareto performs poorly in replicating the sales distribution. This is in line with Head et al. (2014) who find that log-normal fits the entire distribution of firm export sales much better than does Pareto. Similarly, in Figure 2(b) the order statistics simulated from the statistical model whether estimated using the ML or SMM fit the data poorly. Both tend to over-estimate how quickly sales at rank decrease with rank at a given scope.

3.1.3 Quantifying the Fit

To quantify the fit of the statistical model we use a modified version of R^2 which measures the ability of the model to explain the sales at rank pattern. Let r^D_{kK} be the expected value of sales of a product with rank k for firms with K products calculated from the data, let r^S_{kK} be the simulated analogue, and let $\bar{r}^D$ be the simple mean of sales. G is given in equation 1

$$G = 1 - \frac{\sum_{k,K} \left(r^D_{kK} - r^S_{kK}\right)^2}{\sum_{k,K} \left(r^D_{kK} - \bar{r}^D\right)^2} \tag{1}$$



(a) Sales distributed log-normally.

(b) Sales distributed Pareto.

Figure 2 – In each panel the x-axis reports the rank and the y-axis reports average sales across products at a given rank for exporters with the same scope. Product rank is assigned for each product based on it's contribution to total sales of the exporter at the destination. A product with the highest share of sales has rank 1. Scope is indicated at the top of each panel. The last panel in 2(a) and 2(b) respectively shows sales at rank for firms with 6 or more products products, the rank "6+" on the x-axis corresponds to the least selling product,i.e., the y-axis shows average sales in the least selling product for exporters with 6 or more products.

and it measures the share of variation in the order statistics explained by the statistical model relative to the share of variation explained by the simple mean. The values of G for the log-normal distribution and Pareto respectively are shown in Table 3. The statistical model when sales are log-normally distributed accounts for 99.3 percent of variation in order statistics with ML estimates and 99.9 with SMM. Pareto distribution in contrast performs worse than a simple average of sales with the SMM estimate and accounts for 83.2 percent with ML estimates. The results show that if sales are distributed log-normally then the statistical model replicates the variation in sales at rank much better than if sales are Pareto distributed.

Table 3 – The value of G for simulated method of moments and maximum likelihood estimators.

	MLE	SMM
Pareto	0.832	-1.919
Log-normal	0.993	0.999

In the following subsection we show that statistical model with sales-distributed log-normally quantitatively replicates the observed data patterns. If sales are distributed Pareto the statistical model predicts even bigger differences between the large and small scope exporters than in the data.

3.2 Test of the Statistical Model

In this section, we compare the predictions of the statistical model to the patterns in the data. Rather than comparing the outcomes of interest between large and small scope exporters we focus on comparing predictions of the statistical model to the patterns in the data. Thus we are able to identify when differences between the large and small scope exporters arise due to aggregation and when these differences call for an economic explanation.

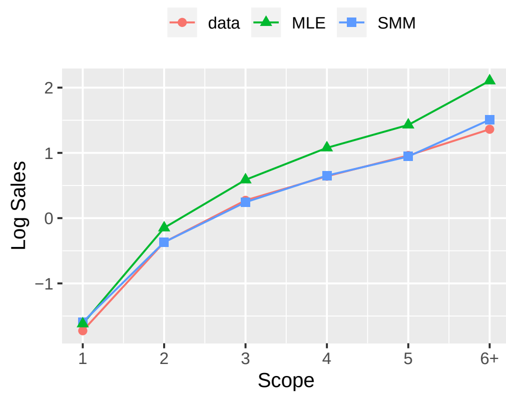
In Figures 3(a) and 3(b) we address prediction 1. Figure 3(a) compares the mean sales of the best selling products as a function of exporter scope in the data to the simulated counterpart from the log-normal distribution. Comparing the sales between large and small exporters in the data, exporters with 6^+ products on average sell more than 6 times as much as single product exporters in their top-selling products. If instead of comparing sales of firms with different number of products we compare averages sales of the top-ranked products in the data to those predicted by the statistical model in Figure 3(a) we see that large scope exporters sell no more or less in their best selling product than we would expect if sales were iid across products and firms. To see this consider the MLE and SMM series correspond to the moments simulated based on the SMM and ML estimates in the top panel of Table 2. The pattern simulated from the SMM estimates tracks the data closely[4] and shows that differences between large scope and small scope exporters arise even if firm-product sales are iid random draws. The pattern simulated from the ML estimates predicts that large scope exporters predicts should sell more in their top selling products that we observe in the data. Differences between the predictions of the statistical model based on the MLE and SMM are driven by differences in variance. SMM and ML estimates have identical location parameter estimates but the shape parameter σ is bigger in the ML case and implies a greater variance of sales. This illustrates that higher variance of sales translates in bigger differences between large and small exporters.

In Figure 3(b) we simulate the best selling products from the statistical model where $F()$ is set to be Pareto with parameters estimated in the second panel of Table 2. While the moments simulated from the Pareto distribution poorly track the data it's worth pointing that they predict much bigger differences between large and small scope exporters than we observe. The patterns simulated from the log-normal and Pareto both illustrate that one should be careful not to interpret the pattern that large scope exporters sell

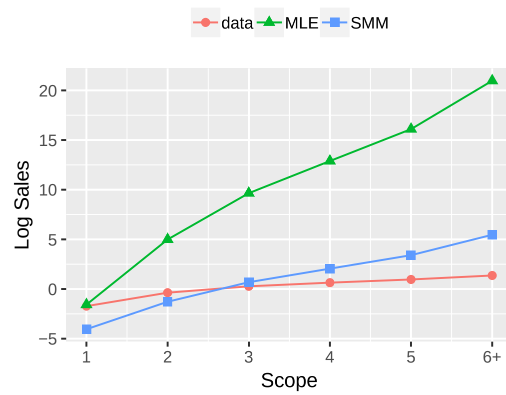
[4]This is not surprising since we have explicitly targeted order statistics conditional on firm scope

more in their top-selling products as evidence of large scope exporters being more productive.

Figure 3 – Average sales in the top-selling product for exporters of different scope.



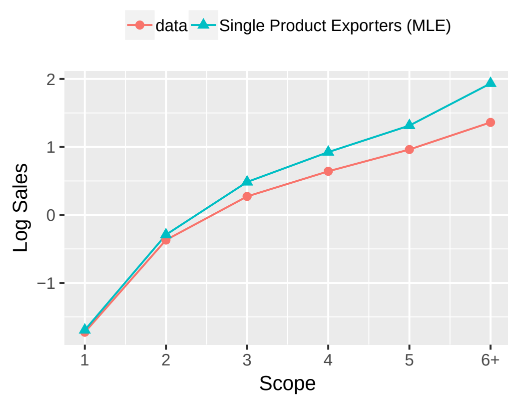
(a) Sales distributed log-normally. MLE series corresponds to the log-normal distribution with a mean of -1.6 and variance 2.12, and SMM series corresponds to the log-normal distribution with a mean of -1.6 and variance 2.62.



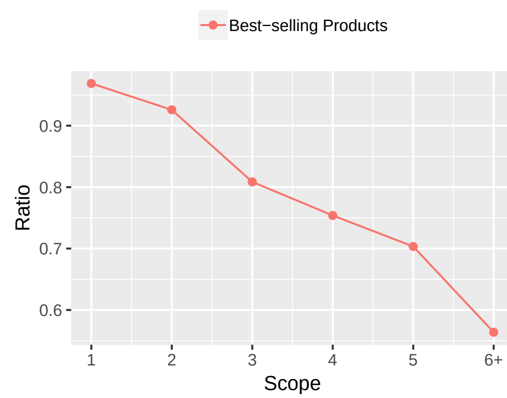
(b) Sales distributed Pareto. SMM series corresponds to Pareto with location parameter 0.000059 and shape parameter 0.074 estimated by the SMM approach. MLE series corresponds to Pareto with location parameter 0.000003 and shape parameter 0.13 estimated by the ML approach.

To highlight this point further we ask how much more large scope exporters would sell in their best selling products if their sales came from the same distribution as single scope exporters. To this end we estimate the location and scale parameters of the log-normal distribution from the data on single-product exporters alone. The location parameter is estimated to be -1.73 and scale parameter to be 2.61. The details are in the bottom panel of Table 2. In Figure 4(a) we compare the simulated mean sales of the best selling products conditional on exporter scope to the data. Mean sales of the single product exporters in the data and the simulation are virtually identical by construction. The statistical model predicts that if larger scope exporters drew sales from the same distribution as single product exporters they would sell more in their top selling products than we observe in the data. In Figure 4(b) we plot the ratio of the actual mean sales of the top-selling products at a given exporter scope to the simulated analogues. For single product exporters this ratio is 1 but as scope increases the ratio decreases to 0.6 for exporters with 6 and more products. This means that if large scope exporters drew their sales from the same distribution as the single scope exporters their sales in the top-selling products would have been 40 percent higher on average than we observe in the data. In other words, once we consider that calculating average sales of the best selling product conditional on exporter scope involves taking a maximum over a different number of products, large scope exporters appear to sell less in their best selling products compared to the smaller scope exporters.

Figure 4 – In the simulation, sales are distributed log-normally and the location and shape parameters estimated from the data on single product exporters only.



(a) Average sales in the top-selling product for exporters of different scope.



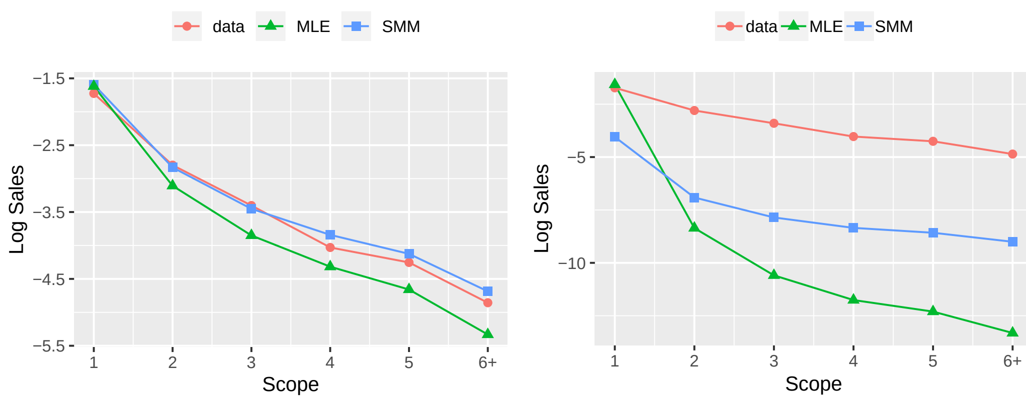
(b) Ratio of the mean best selling product in the data to the mean best selling product simulated from the statistical model conditional on exporter scope.

Figures 5(a) and 5(b) zoom into the least selling products by exporter scope to address Prediction 2. Figure 5(a) compares the prediction of the statistical model to the data if sales are distributed log-normally and Figure

5(b) if sales are distributed Pareto. The pattern simulated from the log-normal distribution based on the SMM estimates narrowly tracks the data and predicts that large scope exporters sell only slightly more in their least-selling products than we observe in the data. The ML estimates, on the other hand, imply that large-scope exporters sell-less in their least selling products than in the data. Although Pareto falls short of quantitatively replicating the data quite dramatically, it too predicts that average sales of the least-selling product decrease with scope. This suggests the observation that large scope exporters sell less in their least selling products arises because we aggregate sales across different number of products and should not be viewed as evidence of firm specific productivity or fixed costs of exporting decreasing with scope.

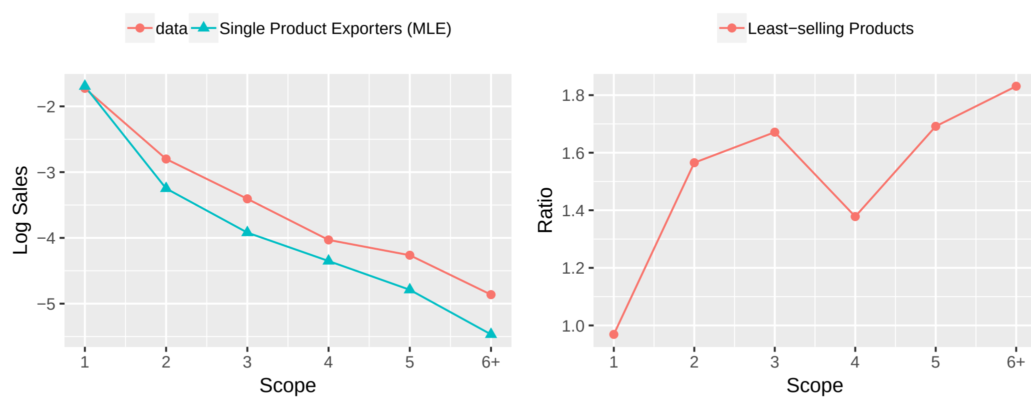
In Figures 6(a) and 6(b) we compare what we observe firms to sell in their least-selling products conditional on exporter scope to what they would sell if the sales were drawn from the log-normal distribution fitted to the data on single product exporters (see the bottom panel of Table 2). While large scope exporters sell more in their least selling products than small scope exporters Figure 6(a) shows that the differences would have been even bigger if large scope exporters drew sales from the single product distribution. Figure 6(b) highlights just how much more. In the data exporters with two products sell 60 percent more and exporters with 6 or more products sell more than 80 percent than we would expect if the firms drew sales from the single product sales distribution.

Figure 5 – Average sales of the least selling products by exporter scope



(a) Sales distributed log-normally. **(b)** Sales distributed Pareto.

Figure 6 – In the simulation, sales are distributed log-normally and the location and shape parameters estimated from the data on single product exporters only.



(a) Average sales in the top-selling product for exporters of different scope.

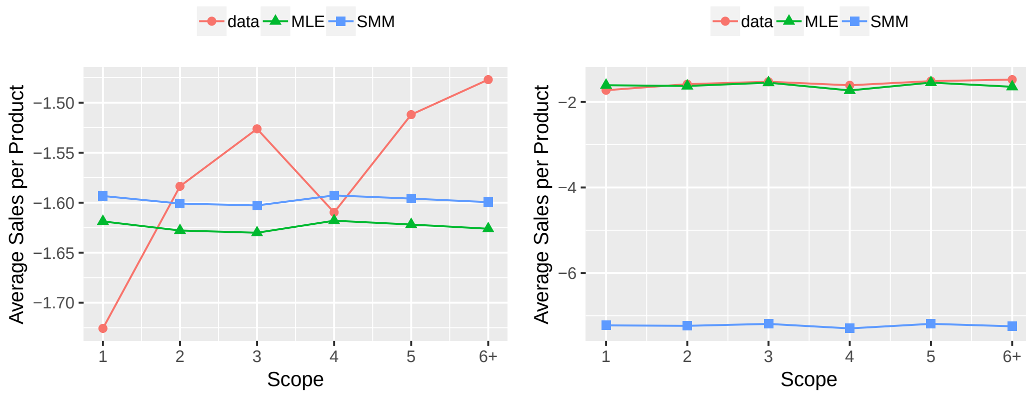
(b) Ratio of the mean best selling product in the data to the mean best selling product simulated from the statistical model. t

So far, we have shown that the relationship between the mean best/least selling products and exporter scope arises due to aggregation even in the absence of ex-ante firm heterogeneity. This ability of the statistical model to replicate the sales at rank pattern indicates that comparing sales of the best and least selling products between large and small scope exporters is misleading. In fact, comparing data to the benchmark of the statistical model highlights that large scope exporters sell no more than single product exporters in their top selling products but sell more in their least-selling products. This is in contrast to the conclusion one would reach by directly comparing mean sales at the highest/lowest rank for firms of different scope.

Our next result shows that randomness is not sufficient to capture the extent of variation in average sales per product that we observe in the data with either log-normal or Pareto distributed sales. In Figures 7(a) or 7(b) average sales per product tend to increase with scope in the data and are virtually independent of scope in the simulation. This is consistent with Prediction 3 and suggests that average sales is a useful moment to characterize multi-product exporters. Furthermore, average sales in Figures 7(a) or 7(b) suggest that large scale exporters on average sell more than smaller scope exporters, although the relationship between scope and average sales per product is not monotone.

Log-normal distribution that fits the overall distribution of sales also produces reasonable magnitude for average sales per product even though it does not replicate the relationship between scope and average sales per product in the data. With Pareto distribution average sales per product predicted by the statistical model depend on how the parameters of the distribution were estimated. Average sales per product based on SMM estimates provide an unreasonably low estimate of average sales per product, even though they produced the better fit for mean sales at rank.

Figure 7 – Average sales of per product by exporter scope.



(a) Sales distributed Log-Normally. **(b)** Sales distributed Pareto.

According to Prediction 4, the log ratio of sales of the best selling to the second best selling product is constant with scope when sales are distributed Pareto and decreases monotonically with scope when sales follow log-normal, Weibul or exponential distributions. In Figure 8 we show the mean of the log ratio decreases with exporter scope both in the data and the simulation assuming sales are log-normally distributed. The model misses the levels but replicates the rate of decline in the ratio with scope. Note, that we do not directly target the ratio moments in the data.

The relationship between the ratio of sales and the number of products is rarely of interest in its own right. Instead, researchers use the ratio as a measure of skewness of the exporter's product mix and focus on how competition impacts it. However, to the extent that changes in the competitive environment influence, the number of products firms export some of the variation in the ratio will be driven by the changes on the extensive margin rather than exporter response on the intensive margin. In the next section, we explore when log-ratio is a useful measure of a firm's concentrating their sales in their best selling products and when it mechanically reflects the variation in the number of products.

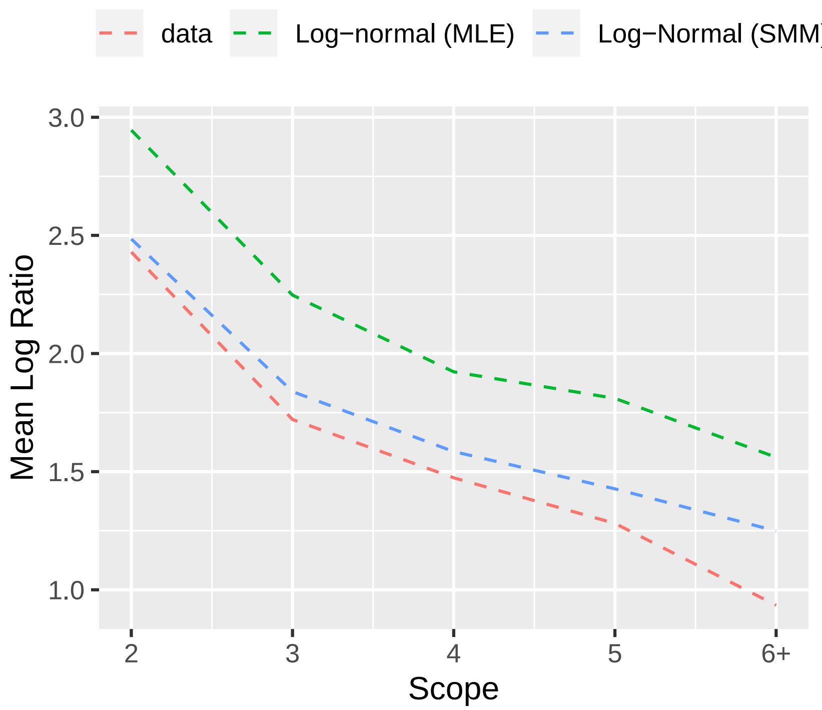


Figure 8 – Mean ratio of the best selling to the second best selling product.

4 Multi-destination Extension

In the previous section we have shown that when sales are log-normally distributed the log ratio of sales of the best selling to the second best selling product is systematically related to the number of products a firm sells in the market: smaller scope exporters have a higher ratio. The ratio, in turn, has been used to measure how multi-product firms respond to changes in the intensity of competition. The best known example of this is Mayer et al. (2014). In their seminal paper on multi-product exporters, they show that tougher competition in an export destination shifts down the entire distribution of markups across products and induces multi-product exporters to concentrate on their core products[5]. With product sales distributed Pareto an increase in sales concentration is equivalent to an increase in the log-ration. While Mayer et al. (2014) consider alternative measures of concentration, their preferred formulation of the model's testable prediction relies on the log-ratio.

Specifically, Mayer et al. (2014) regress the log-ratio of sales between the best and the second best selling product on destination market size, it's supply potential[6] and other variables that proxy toughness of competition in a market[7]. As long as firm-product sales are distributed Pareto using log-ratio to capture the reallocation of resources across the fixed product range is a perfectly valid approach. However, this is unlikely to be the case in the data, and the log-ratio captures variation in the number of products a firm exports to a market above and beyond the reallocation of resources across products within the firm. This is a problem because exporter scope at a destination itself varies with market toughness. Bernard et al. (2010) and Feenstra and Ma (2007) among others document that firms export a wider product range to larger markets and trim their scope in high supply potential destinations.

If firms export more products to large market size destinations then the statistical model predicts negative correlation between market size and the log-ratio. Pro-competitive effects of a larger market that work through firms reallocating resources to lower cost (higher mark-up) products in contrast imply a positive correlation. It's even trickier with the supply potential. Firms

[5] Core products correspond to the lowest marginal cost products

[6] The supply potential is used to proxy for the geography of a destination that does not rely on country-level data for that destination. It is typically constructed as the aggregate predicted exports to a destination based on a bilateral trade gravity equation (in logs) with both exporter and importer fixed effects as well as the standard bilateral measures of trade barriers/enhancers. Following Mayer et al. (2014) the foreign supply potential is a related measure of a destination's foreign supply potential that does not use the importer's fixed effect when predicting aggregate exports to that destination. By construction, foreign supply potential is thus uncorrelated with the importer's fixed-effect.
The supply potential data by Head and Mayer (2011) is available online at http://www.cepii.fr/anglaisgraph/bdd/marketpotentials.htm.

[7] The regressions include country-specific random effects on firm-demeaned data.

export fewer products to high supply potential destinations, so the statistical model predicts a positive correlation between the log-ratio and market supply potential. Pro-competitive effects captured by the high-supply potential also imply a positive correlation with the log-ratio. So, when one regresses the log-ratio on the measures of market competitiveness the coefficients reflect a combination of product mix adjustments and product scope adjustments. To the extent that one wants to measure the effects of competition on the reallocation of resources across a given product range one will underestimate the effect of market size on changes in product mix and over-estimate the effect of market supply potential.

Whether variation in the number of products exported alone is sufficient to generate statistically significant relationships between the log-ratio and the variables that proxy intensity of competition in the market is an empirical question. To answer it we extend the statistical model to the multi-country set-up. We then simulate the statistical model to obtain a data set where firm outcomes are driven by randomness alone. We then replicate the regressions in Mayer et al. (2014) to evaluate the impact of the market competitiveness measures on the log-ratio in the actual and simulated data sets. Any statistically significant results using the simulated data set will indicate that the results are driven by variation in the number of products rather than changes on the intensive margin.

We extend the statistical model to the multi-product set-up as follows. As with a single destination case a multi-product exporter indexed by f is a collection of products K_d that a firm exports in a given year to a given destination d. Products that a firm exports to a destination are indexed by $k \in \{1, ..., K_d\}$. The number of products a firm exports to a destination d (K_d) is treated as exogenous. Product sales S_{fkd} are iid draws from the same log-normal distribution.

To estimate the parameters of the sales distribution $F()$ we use maximum likelihood (ML) approach where we treat sales of the Chinese exporters to

Table 4 – Country Level Summary Statistics

	Mean	Std. Dev.	25%	50%	75%	95%
Mean Country Log Ratio	1.78	0.63	1.54	1.72	1.91	2.46
Mean Country Scope	2.80	0.45	2.60	2.73	2.95	3.72
Number of Exporters	584	1473	17	82	512	2627
Log GDP	23.4	2.27	21.9	23.1	25.1	27.3
Log Supply Potential	14.9	0.99	14.3	14.6	15.4	16.9
Log GDP per Capita	7.85	1.64	6.45	7.78	9.10	10.5
Log Distance	9.02	0.53	8.83	9.06	9.39	9.63
Observations	176					

The sample contains all countries to which at least one Chinese multi-product exporter sells at least two products in 2003 and includes only manufacturing producers. The first three rows describe the variation in the outcome variables across the different destinations to which Chinese exporters sell. The *Mean Country Log Ratio* is the mean log-ratio of sales of the best selling to the second best selling product across all firms exporting to a destination. The *Mean Country Scope* is the average number of products that firms export to a destination and *Number of Exporters* is the number of firms exporting to a destination. The Table also contains information on the variables that we will use to proxy market toughness of a destination: *Log GDP*,*Log Supply Potential*, *Log GDP per Capita* and *Log Distance* are the determinants of the market toughness that were used in Mayer et al. (2014).

176 destinations as if they come from a single distribution.

Table 4 provides some summary statistics for the country level variables we use. The sample includes all destinations to which at least one Chinese exporter sells at least two products in 2003 and only includes manufacturing firms involved in production of its exports. The number of multi-product exporters selling to a destination appears to be highly skewed: the median number of exporters is only 82 while the 95^{th} percentile is 2,627. Average exporter scope on the other hand is much more balanced across destinations and varies between 2 and 4. The Table also contains information on the determinants of the market toughness (*Log GDP*, *Log Supply Potential*, *Log GDP per Capita*, and *Log Distance*) that we will use in replicating the analysis in Mayer et al. (2014) in the regression analysis.

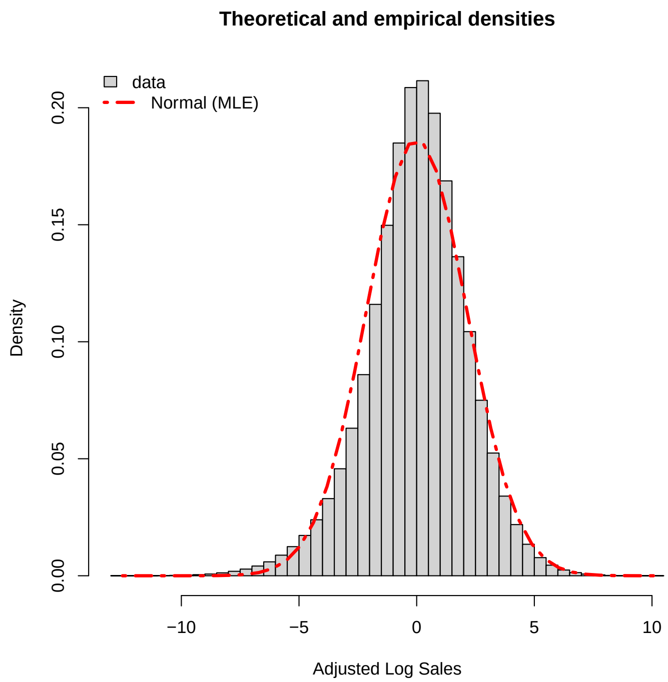


Figure 9 – The density of firm-product-destination log sales and the estimated density $N(0, 2.14)$.

In it's simplest form the multi-country extension of the model has two parameters: the location and shape parameter that govern the distribution of sales. In the data, we have to contend with the fact that sales systematically vary with destinations and product categories (6-digit HS groups). To take this into account we purge the sales of the destination-product fixed effects. We do this by first estimating equation 2 where the dependent variable is the log of sales and γ_{kd} captures the destination-product fixed effect.

$$lnS_{fkd} = \gamma_{kd} + \epsilon_{fkd} \tag{2}$$

We then use the residual ϵ_{fkd} to estimate parameters of sales distribution $F()$. Using the ML approach and the assumption that ϵ_{fkd} is normally distributed yields the variance of 2.14 with standard error of 0.002 and mean equal to 0 by construction. Figure 9 compares the estimated theoretical distribution and the actual distribution. While the log-normal distribution performs poorly in the left tail it otherwise is able to approximate the empirical distribution of log sales well.[8]

We simulate data from the estimated log-normal distribution with mean 0 and standard deviation 2.14 for each exporter at a destination and use it to study the relationship between the simulated log-ratio and country level variables that measure the level of competition at the destination. See Table 4 for the list of country level variables. With data simulated from the statistical model, sales per product are completely random and any relationship between the log-ratio and the variables that proxy market toughness is driven entirely by aggregating data across different number of products.

First, to establish the relationship between the ratio and the various measures of market toughness in the Chinese data we replicate the regressions from Table 2 in Mayer et al. (2014) in columns (1) and (5) in Table 5. In column (1), we look only at the effect of market size and foreign supply potential, in column (5) we also include the explicit geographic barriers. Only the *log of GDP* variable is highly significant across regressions. In magnitude the coefficient on the *Log GDP* variable is similar to what Mayer et al. (2014) find for the French exporters. To the extent that GDP captures intensity of competition we observe that the effect on the log-ratio is positive - firms concentrate their resources in their best selling products. The foreign supply potential that is significant in the French data is not significant in any of the formulations in the Chinese data.

Next, we reproduce the regressions with the simulated data in columns (3) and (7). In the simulated data the product sales are just iid draws from the log-normal distribution. Hence any effect of market competitiveness measures that we find is driven by aggregating data on sales across different

[8]Using raw data on log of sales does not change the estimate of the variance. The results can be found in the appendix

number of products to the firm level. If a lot of small-scope exporters sell at a destination the ratio will be mechanically higher than if the destination was dominated by large scope exporters. Both in columns (3) and (7) the coefficient on GDP is negative and significant reflecting that market size encourages firms to export more products which mechanically lowers the log-ratio. The foreign supply potential variable is positive in both regressions. Interestingly it becomes significant in the simulated data when we control for bilateral trade barriers between China and it's trading partners in column (7). This is because including both geographic barriers and supply potential helps differentiate between good geography and distance as many high supply potential markets are located far away from China[9]. Distance and contiguity in column (7) are also statistically significant. The log of distance enters with a negative sign while contiguity has a positive effect. These signs are consistent with firms exporting a narrow range of products to faraway destinations and a wider range to countries nearby. To check that the variation in the simulated data is indeed driven by the variation in scope we use the number of products as the dependent variable in columns (4) and (8) respectively. The signs on the independent variables are consistent with variation in scope driving the results in the simulated regression.

Finally, we introduce the Scope variable in the regression of the actual log-sales ratios on measures of market competitiveness in columns (2) and (6). In both cases the coefficient on the scope variable is negative and statistically significant. What is more interesting is that it changes the magnitude of the market size coefficient. Comparing the regressions in column (1) and (2) the coefficient on the GDP variable increased from 0.035 to 0.05 which is a 30 percent increase with the inclusion of the Scope variable. Similarly

[9]In French data that Mayer et al. (2014) use distance from France is highly correlated with good geography and hence a high supply potential for that destination. The correlation between the log-distance and log-supply potential is 78 percent. This is not the case in Chinese data, the correlation between the foreign supply potential and distance in logs is negative and relatively small (-0.2191)

the coefficient on the GDP variable increases from 0.034 in column (5) to 0.05 in column (6). This is consistent with log-ratio reflecting variation in scope beyond variation in concentration. Controlling for scope suggests that the original results under-estimate the adjustment of the product mix by as much as 30 percent. The coefficient on the supply potential variable is still statistically insignificant but changes the sign from positive to negative again implying that in the original formulation the coefficient under-estimates the negative effect of the supply potential on product mix. Note however, that the scope and product mix which we approximate with the log-ratio are likely to be simultaneously determined so using scope as an independent variable is not econometrically correct.

Table 5 – Country Level Summary Statistics

	Log Ratio			Scope	Log Ratio			Scope
	Data		Simulation		Data		Simulation	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
GDP	0.0347***	0.0505***	-0.0443***	0.189***	0.0338***	0.0509***	-0.0479***	0.209***
	(10.40)	(14.54)	(-7.42)	(6.05)	(8.24)	(15.15)	(-7.80)	(9.34)
Supply Potential	0.00124	-0.00195	0.00912	-0.0490	0.00612	-0.00488	0.0225*	-0.118***
	(0.11)	(-0.21)	(0.86)	(-1.53)	(0.47)	(-0.45)	(2.12)	(-4.19)
Scope		-0.0749***				-0.0752***		
		(-6.97)				(-6.96)		
Distance					0.0331**	-0.00206	0.0617***	-0.430***
					(2.96)	(-0.17)	(4.10)	(-4.43)
Contiguity					0.000720	0.0329*	-0.0901*	0.338
					(0.04)	(2.04)	(-2.45)	(1.37)
GATT					0.00971	0.0377	-0.0554	0.272*
					(0.26)	(1.01)	(-1.46)	(2.06)
Observations	102758	102758	102758	102758	102758	102758	102758	102758
Within R^2	0.000770	0.0148	0.00105	0.0116	0.000883	0.0148	0.00156	0.0252
R^2	0.00193	0.0160	0.00214	0.0198	0.00213	0.0161	0.00281	0.0411

Notes: All columns use Wooldridge's (2006) procedure with country-specific random effects on firm and HS2 demeaned data, with a robust covariance matrix estimation. t-statistics in parentheses.

In Table 5 the simulated data results highlight that variation in the number of products can drive variation in the log-ratio even when firms don't optimize on the intensive margin. As such the log-ratio may be less than an ideal measure of changes in sales concentration in response to changes in competition. We have shown that the coefficient on the marker size is biased downward while the coefficient on the foreign supply potential is biased upward. This has implications for the measures of economic significance of the product mix adjustment mechanism. The GDP coefficient in the baseline regression can be interpreted as the average elasticity of the log-ratio with respect to the destination's GDP. Mayer et al. (2014) use this elasticity to estimate that the within firm resource reallocation on the intensive margin contributes 19 percent to aggregate productivity change. Our results suggest the true effect of GDP on skewness measured by the log ratio of sales is larger than the baseline regression estimates and would imply that Mayer et al. (2014) underestimate the contribution of the product mix channel to aggregate productivity growth.

More generally, the results of this section illustrate that the log-ratio, as well as other measures that are based on order statistics should be treated with care in the multi-product firm analysis as they systematically vary with the number of products per firm.

5 Conclusion

In this paper, we propose a stark statistical model of multi-product exporter to help separate which well-documented facts about multi-product exporters are genuinely informative about the economic forces behind the observed outcomes and which ones arise as an artefact of aggregating data across a different number of products. Our results show that patterns that rely on order statistics should be treated with caution. For example, we find that the empirical regularity about multi-product exporters that large scope exporters sell more in their best selling products that have been used to motivate models featuring firm-specific productivity can be replicated even if sales are just random draws from a distribution. Similarly, we can quantitatively replicate that large scope exporters sell less in their least selling products compared to their smaller scope counterparts without resorting to firm productivity or product-specific fixed costs.

Our results also show that the log-normal distribution provides a much better fit than Pareto distribution, so often employed in theoretical models. While log-normal distribution allows reproducing order statistics across firms of different scopes remarkably well. Pareto distribution over-predicts sales of the least selling products especially for large scope firms. This may have potential implications for estimating fixed costs of exporting for multi-product exporters when product shocks are drawn from the Pareto distribution.

The Pareto distribution also predicts that the log-ratio of the sales in the best selling to the second best selling product is constant with scope. With Weibull, exponential and log-normal distributions the ratio is systematically related to firm scope. Large scope exporters have a lower ratio than the small scope exporters. We show that this has implications for regression analysis that relies on exploring the effect of export market competitiveness on product mix measured by the log-ratio of the best selling to the second best selling product sales.

Our key insight is that when we look at differences between large and small scope exporters that have been traditionally invoked to motivate a single-attribute model of multi-product exporters we find that those differences are consistent with randomness and aggregation. This has implications for the ways in which we model multi-product exporters. If there is no evidence that large scope exporters sell any more in their best-selling products than small scope exporters it calls for models where scope and scale are not necessarily driven by a single parameter.

References

Albornoz, F., H. F. C. Pardo, G. Corcos, and E. Ornelas (2012): "Sequential exporting," *Journal of International Economics*, 88, 17 – 31.

Allanson, P. and C. Montagna (2005): "Multiproduct firms and market structure: An explorative application to the product life cycle," *International Journal of Industrial Organization*, 23, 587–597.

Arkolakis, C., S. Ganapati, and M.-A. Muendler (2014): "The Extensive Margin of Exporting Products: A Firm-level Analysis Online Supplement," .

Armenter, R. and M. Koren (2014): "A Balls-and-Bins Model of Trade," *American Economic Review*, 104, 2127–51.

Bee, M. and S. Schiavo (2018): "Powerless: gains from trade when firm productivity is not Pareto distributed," *Review of World Economics*, 154, 15–45.

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

——— (2011): "Multiproduct firms and trade liberalization," *The Quarterly Journal of Economics*, 126, 1271–1318.

Boehm, J., S. Dhingra, and J. Morrow (2019): "The Comparative Advantage of Firms," Tech. rep., CEPR Discussion Paper No. DP13699 Available at SSRN: https://ssrn.com/abstract=3379615.

De Loecker, J. (2011): "Product differentiation, multiproduct firms, and estimating the impact of trade liberalization on productivity," *Econometrica*, 79, 1407–1451.

Dhingra, S. (2013): "Trading Away Wide Brands for Cheap Brands," *The American Economic Review*, 103, 2554–2584.

Dhyne, E., A. Petrin, V. Smeets, and F. Warzynski (2017): "Multi product firms, import competition, and the evolution of firm-product technical efficiencies," Tech. rep., National Bureau of Economic Research.

Eckel, C., L. Iacovone, B. Javorcik, and J. P. Neary (2015): "Multi-product firms at home and away: Cost- versus quality-based competence," *Journal of International Economics*, 95, 216 – 232.

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

Ellison, G. and E. L. Glaeser (1997): "Geographic concentration in US manufacturing industries: a dartboard approach," *Journal of political economy*, 105, 889–927.

Feenstra, R. and H. Ma (2007): "Optimal choice of product scope for multiproduct firms under monopolistic competition," Tech. rep., National Bureau of Economic Research.

Fernandes, A. M., P. J. Klenow, S. Meleshchuk, M. D. Pierola, and A. Rodriguez-Clare (2015): "The intensive margin in trade: moving beyond pareto," *Manuscript, Stanford University.*

Head, K. and T. Mayer (2011): "Gravity, market potential and economic development," *Journal of Economic Geography*, 11, 281–294.

Head, K., T. Mayer, and M. Thoenig (2014): "Welfare and trade without Pareto," *American Economic Review*, 104, 310–16.

Helpman, E. (1985): "Multinational Corporations and Trade Structure," *The Review of Economic Studies*, 52, 443–457.

Iacovone, L. and B. S. Javorcik (2010): "Multi-Product Exporters: Product Churning, Uncertainty and Export Discoveries," *The Economic Journal*, 120, 481–499.

Ju, J. (2003): "Oligopolistic Competition, Technology Innovation, and Multiproduct Firms," *Review of International Economics*, 11, 346–359.

Klette, J. and S. Kortum (xxxx): "Innovating Firms and Aggregate Innovation," *xxxx*, xx, yy–xx.

Malik, H. J. and R. Trudel (1982): "Probability density function of quotient of order statistics from the pareto, power and weibull distributions," *Communications in Statistics-Theory and Methods*, 11, 801–814.

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

Nocke, V. and S. Yeaple (2015): "GLOBALIZATION AND MULTI-PRODUCT FIRMS," *International Economic Review*, 55.

Roberts, M. J., D. Y. Xu, X. Fan, S. Zhang, et al. (2012): "The role of firm factors in demand, cost, and export market selection for chinese footwear producers," *NBER Working Paper*, 17725.

Timoshenko, O. A. (2015): "Product switching in a model of learning," *Journal of International Economics*, 95, 233 – 249.

Topalova, P. and A. Khandelwal (2011): "Trade liberalization and firm productivity: The case of India," *Review of economics and statistics*, 93, 995–1009.

6 Appendix

A Proof of Prediction 4

A.1 Pareto distribution

Let S_{fk} be iid draws from the Pareto distribution with cdf:

$$F(S_{fk}) = 1 - \left(\frac{S_{fk}}{a}\right)^{-v} \tag{3}$$

where a is the location parameter such that S_{fk} and $a > 0$, $v > 0$ is the shape parameter. Let $X_{j:n}$ and $X_{i:n}$ be the i^{th} and the j^{th} order statistic from the random sample of size n from $F(S_{fk})$. With $i < j$ the ratio $Z_{i:j} = \frac{X_{i:n}}{X_{j:n}}$. Using the inverse Mellin's transform Malik and Trudel (1982) show that the distribution of the ratio is

$$h(z_{i:j}) = \frac{v z_{i:j}^{v+vn-vj-1}}{B(j-i, n-j+1)} * (1 - z_{i:j}^{v})^{j-i-1} \tag{4}$$

where $0 \leq z \leq 1$ and $0 < i < j \leq n$.

Letting $j = n$ and the $i = n - 1$, the expression in 4 reduces to

$$h(z_{n-1:n}) = v z_{n-1:n}^{v-1} \tag{5}$$

The expression $h_{n-1,n}(z)$ is independent of the sample size, and so will be the expected value of $E[ln(1/z_{i:j})]$

A.2 Weibul distribution

Let S_{fk} be iid draws from the Weibul distribution with cdf:

$$F(S_{fk}) = 1 - e^{-\frac{S_{fk}^{\alpha}}{\theta}} \tag{6}$$

where $S_{fk} > 0$, $\alpha > 0$, $\theta > 0$. Let $X_{j:n}$ and $X_{i:n}$ be the i^{th} and the j^{th} order statistic from the random sample of size n from $F(S_{fk})$. With $i < j \leq n$ the ratio $Z_{i:j} = \frac{X_{i:n}}{X_{j:n}}$. Using the inverse Mellin's transform Malik and Trudel (1982) show that the distribution of the ratio is

$$h(z_{i:j}) = \frac{n!}{(i-1)!(j-i-1)!(n-j)!} \sum_{r=0}^{j-i-1} \sum_{s=0}^{i-1} (-1)^{r+s} \frac{\alpha z^{\alpha-1}}{[(n-j+r+1)+(j-i-r+s)z^{\alpha}]^2} \quad (7)$$

where $0 \leq z \leq 1$ and $0 < i < j \leq n$.

Letting $j = n$ and the $i = n-1$, the expression in 7 reduces to:

$$h(z_{n-1:n}) = n(n-1) \sum_{s=0}^{n-2} (-1)^s \binom{n-2}{s} \frac{\alpha z^{\alpha-1}}{[1+(s+1)z^{\alpha}]^2} \quad (8)$$

The expression $h_{n-1,n}(z)$ depends on the size of the sample from which the ordered statistics are calculated, and so will be the expected value of $E[ln(1/z_{i:j})]$. The expected value of $E[ln(1/z_{i:j})] = E[ln(1/z_{i:j})]$

$$\begin{aligned}
E[ln(1/z_{n-1:n})] &= \int_0^1 ln\left(\frac{1}{z}\right) * h(z_{n-1:n}) dz_{n-1:n} \\
&= \int_0^1 ln(1) * h(z_{n-1:n}) dz_{n-1:n} - \int_0^1 ln(z) * h(z_{n-1:n}) dz_{n-1:n} \\
&= -\int_0^1 ln(z) * h(z_{n-1:n}) dz_{n-1:n} \\
&= -\int_0^1 ln(z) \left(n(n-1) \sum_{s=0}^{n-2} (-1)^s \binom{n-2}{s} \frac{\alpha z^{\alpha-1}}{[1+(s+1)z^{\alpha}]^2} \right) dz_{n-1:n}
\end{aligned}$$

Interchanging summation and integration order obtain:

$$-n(n-1) \sum_{s=0}^{n-2} (-1)^s \binom{n-2}{s} \left(\int_0^1 \frac{ln(z) \alpha z^{\alpha-1}}{[1+(s+1)z^{\alpha}]^2} dz_{n-1:n} \right)$$

The integral

$$\int_0^1 \frac{ln(z)\alpha z^{\alpha-1}}{[1+(i+1)z^\alpha]^2} dz_{n-1:n} = -\frac{ln(2+s)}{(s+1)\alpha}$$

$$\begin{aligned} E[ln(1/z_{n-1:n})] &= -n(n-1)\sum_{s=0}^{n-2}(-1)^i \binom{n-2}{s}\left(-\frac{ln(2+s)}{(s+1)\alpha}\right) \\ &= \frac{1}{\alpha}\sum_{s=0}^{n-2}(-1)^i \binom{n}{s}\frac{ln(2+s)}{(s+1)} \end{aligned}$$

The expected value of the ratio drawn from the Weibull distribution is given by the oscillating sum and in general cannot be shown to be monotonically decreasing or increasing in the size of the sample n. However for small value of n it can be shown to be monotonically decreasing in n.

A.3 Log-normal distribution

Let S_{fk} be iid draws from the log-normal distribution with cdf:

$$F(S_{fk}) = \frac{1}{2} + \frac{1}{2}Erf\left(\frac{-(lnS_{fk}-\mu)^2}{2}\sigma^2\right) \tag{9}$$

where $S_{fk} > 0$, $-\infty \leq \mu \leq \infty$, $\sigma > 0$. $X_{j:n}$ and $X_{i:n}$ are the i^{th} and the j^{th} order statistic from the random sample of size n from $F(S_{fk})$. With $i < j \leq n$ the ratio $Z_{i:j} = \frac{X_{i:n}}{X_{j:n}}$. In the case of log-normal distribution it is not possible to derive the distribution of the ratio for an arbitrary n. Instead we adopt numerical integration approach and calculate the expected log-ratio for various values of μ and n.

The joint distribution of two order statistics $X_{j:n}$ and $X_{i:n}$ such that

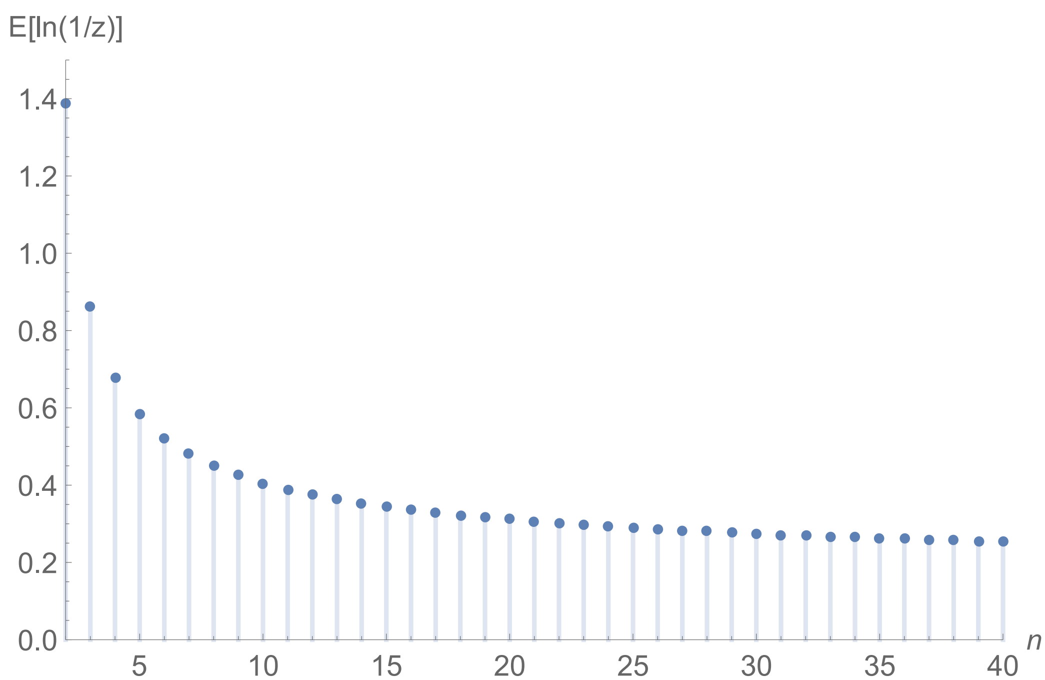


Figure 10 – $E[ln(X_{n:n} - ln(X_{n-1:n})]$ as a function of the size of the ordered sample drawn from Weibull distribution

$0 < X_{i:n} \leq X_{j:n} < \infty$ is given by

$$g(X_{j:n}, X_{i:n}) = \frac{n!}{(i-1)!(j-i-1)!(n-j)!} F^{i-1}(X_{i:n}) \left[F(X_{j:n}) - F(X_{i:n})\right]^{j-i-1}$$
$$\left[1 - F(X_{j:n}\right]^{n-j} f(X_{j:n}) f(X_{i:n})$$

Letting $j = n$ and $i = n-1$ obtain

$$g(X_{n-1:n}, X_{n:n}) = n(n-1)F^{n-2}(X_{n-1:n})f(X_{n:n})f(X_{n-1:n})$$

The expected value of $E\left[ln\left(\frac{X_{j:n}}{X_{i:n}}\right)\right]$ or $E\left[ln(X_{j:n}) - ln(X_{i:n})\right]$

$$E\left[ln\left(\frac{X_{j:n}}{X_{i:n}}\right)\right] = \int_0^\infty \int_{X_{n:n}}^\infty \left(ln(X_{n:n}) - ln(X_{n-1:n})\right) g(X_{n:n}, X_{n-1:n}) dX_{n:n} dX_{n-1:n}$$

While the integral is hard to evaluate analytically we show results of numerical integration in the table below. The expected value decreases with the size of the ordered sample. The decrease is particularly important for small samples.

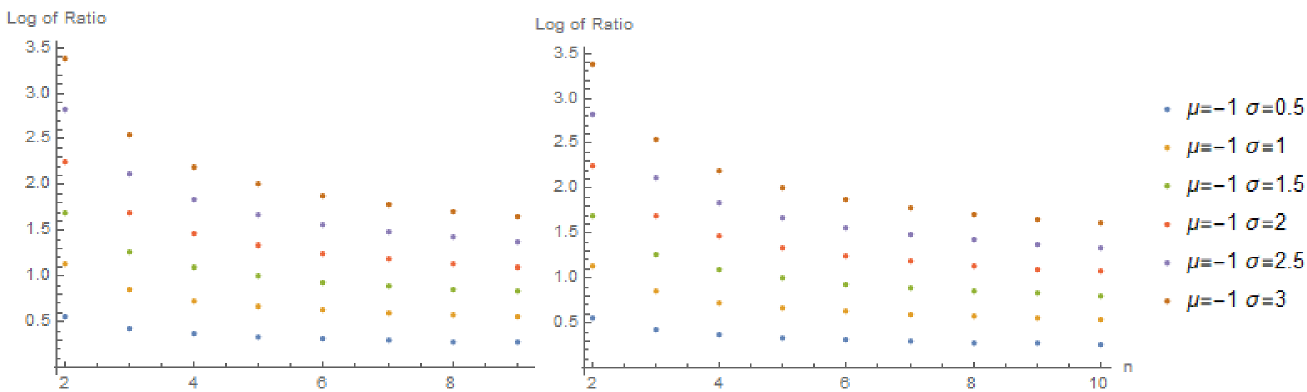


Figure 11 – $E[ln(X_{n:n} - ln(X_{n-1:n})]$ as a function of the size of the ordered sample drawn from log-normal distribution