Decomposing Firm-level Sales Variation*

Jakob R. Munch and Daniel X. Nguyen†
University of Copenhagen
May, 2014

Abstract

Recently, much of the trade literature has been focused on using firm-specific productivity to explain export heterogeneity. This study provides evidence for the importance of incorporating firm-destination-specific effects such as demand shocks in theories of exporter heterogeneity. Our study estimates the proportion of firm-level sales variation within a product-destination market that can be explained by firm-specific effects such as productivity. We use a highly detailed dataset comprising firm–product-destination-specific exports and correct for truncation as modeled by recent trade theories. We find that the contribution of firm-specific effects varies greatly across products and that it is 45% for the median product. That is, within-destination sales variation is primarily explained by firm-destination-specific heterogeneity for the majority of products.

Keywords: Firm heterogeneity, firm-level export data, truncation correction.

JEL Codes: F12, C24

---

*We would like to thank participants at Purdue University, University of Copenhagen, DIEW Aarhus, ETSG, Midwest Meetings, EIIT, LMDG conference on 'Labor Economics and International Trade', Sandbjerg and NOITS for helpful suggestions and comments. In particular, we would like to acknowledge David Hummels, Chong Xiang, and Mette Ejrnæs for suggestions that have greatly improved the direction and exposition of this paper. Financial support from the Tuborg Foundation is gratefully acknowledged.

†Address: Department of Economics, University of Copenhagen, Øster Farimagsgade 5, bygning 26, 1353 København K, Denmark, E-mails: jakob.roland.munch@econ.ku.dk, daniel.x.nguyen@econ.ku.dk.
1 Introduction

There is substantial variation in firm-level export sales. Recent theoretical works have attributed this sales variation to heterogeneity in firm productivities.\(^1\) These theories were motivated by earlier empirical studies that identified differences between firms that do export and firms that do not:\(^2\) on average, exporters produce more, hire more labor, pay higher wages, and exhibit higher productivities as measured by either total factor productivity or value added per worker. The contrasts between exporters and nonexporters supported the story that productivity and exporting status were linked.

This paper decomposes firm-level sales variation within a product-destination market. We estimate the proportion of sales variation that can be explained by a firm specific component. This firm specific component comprises all firm characteristics that would affect firm sales, such as productivity, quality, or economies of scale. Since productivity is only one of many firm specific characteristics, the sales variation explained by the firm specific effects is an upper bound of that explained by productivity. After accounting for firm-specific effects, the remaining variation must be explained by a firm-destination specific component. This firm-destination specific component can be thought of as shocks that affect the firm differently in different markets. These could be demand shocks or firm-destination specific cost shocks.\(^3\)

Our empirical approach is in part inspired by the product-level sales decomposition of Hummels and Klenow (2005). That study decomposes a country’s export sales into the number of products and the sales per product. It shows how much of a country’s export sales can be explained by models concentrating on intensive trade margins (Armington 1969) rather than models concentrating on extensive trade margins (Krugman 1979). We also employ a decomposition method to discriminate between trade theories. We show


how much within-destination sales variation can be explained by a model concentrating on firm specific effects such as Melitz (2003) rather than firm-destination specific effects.

Our paper adds to a small but growing literature examining the destinations to which firms export. The lack of work in the area is due primarily to the dearth of firm-destination specific export observations. Eaton, Kortum, and Kramarz (2004) find that most French firms export to only one destination (the mode being Belgium), and that the entry of French firms into a market accounts for two-thirds of the growth of the French share of market sales. Newer studies show that firms supply domestically for several years before exporting, that they usually begin exporting to one destination country, and that many stop exporting activities soon after they begin.\(^4\) Since productivity is realized before supply to any destination and applies to all destinations, these studies present empirical patterns unreconciled by the productivity heterogeneity models.

The current study adds to and extends this literature by utilizing a highly disaggregated and detailed dataset – we observe destination specific shipment values for the universe of Danish exporters in 2001 to 2003. This disaggregation level allows us to identify the firm specific component and a firm-destination specific component of an export. We can estimate the contribution of each using an empirical technique that accounts for the truncation of export data due to the self selection of firms into destinations. We show that this mechanism is consistent with modern trade theory. Adding destination specific effects to the standard productivity heterogeneity model by Melitz (2003) or multiproduct firm versions such as Bernard, Redding and Schott (2011) or Mayer, Melitz and Ottaviano (2013) yields a decomposition of firm-level sales that we take to the data.

Three contemporary studies have goals related, but not identical, to our own. Eaton, Kortum and Kramarz (2011) structurally estimate the contribution of firm specific productivity to both the probability of entering a destination and the variance of sales conditional on entry. They find that the variance of firm specific effects can account for 57% of the variation of entry into a destination and 39% of the variation of sales in a desti-

nation conditional upon entry into that destination. Kee and Krishna (2008) examine Bangladeshi exports of textiles to the US and EU. They find that a textile firm’s market share in EU cannot predict its market share in the US: the correlation between the two is not statistically different from zero. Lawless and Whelan (2008) use firm-destination data from a survey of 676 Irish-owned exporters to explain to where and how much firms export. They use OLS estimates to find that firm-year specific effects account for 41% of firm-destination-year sales variation.

In contrast to Kee and Krishna (2008) and Lawless and Whelan (2008) our empirical approach accounts for the truncation of data consistent with the broad class of CES trade models with firm heterogeneity. On the other hand our approach is more empirical than Eaton, Kortum and Kramarz (2011) in the sense that we only use the theoretical framework to show how firm-level sales may be decomposed into its specific components. Other than that we do not impose much structure on the estimations. This structure requires that we carefully account for the effects that self selection has on our data: only the most productive firms will export and this must be accounted for to avoid biased estimates. Finally, our study uses the most detailed dataset of the related studies. The data cover the universe of Danish firms and uniquely identifies exports by destinations at the eight digit product level. This level of disaggregation at the product level is not available in the three other studies. It turns out to be very important as we document substantial cross product variation in the contribution of firm specific productivity to sales variation.\(^5\)

In our main results, we estimate the contribution of firm specific heterogeneity to overall 2003 Danish export sales variance by HS6 product category. For half of Danish exported products, the contribution is lower than 45%. The mean firm specific contribution for our sample is marginally higher at 49%. In summary, firm specific effects play a nontrivial role, but firm-destination specific effects matter more in explaining firm-level sales variation for most products. As robustness checks, we look at different product

\(^5\)In Section 6 we offer a more thorough discussion of how our approach and results relate to the literature.
aggregation levels and different years. We also remove small trade flows and new trade flows. Our results consistently point towards firm-destination specific effects as an important driver of sales variation. This suggests that it is relevant to incorporate and work out the implications of firm-destination specific effects such as demand shocks in theories of exporter heterogeneity.

In the next section, we present the Danish firm-level export data including an illustrative example of how firm specific effects may or may not drive firm sales across destinations. Next, we show how firm-level sales may be decomposed in a standard heterogeneous firm trade model and how truncation biases standard estimation procedures. Section 4 outlines our empirical strategy to overcome this bias. Section 5 presents our estimation results, and we conclude with a discussion of our results in contrast to previous estimates.

2 Danish firm-level data

The Danish External Trade Statistics provides product-level destination specific export data for the universe of Danish firms. Exports are recorded according to the eight-digit Combined Nomenclature (CN) product code which encompasses approximately 10,000 different product categories. While all trade flows with non-EU countries are recorded by customs authorities (and so the coverage rate in the data is higher than 95 percent), there is not a similar system in place for intra-EU trade. However, intra-EU trade is recorded through the Intrastat system, where firms are obliged to report trade data on a monthly basis. One source of inaccuracy in this system is that some predominantly small firms appear not to report data to the system. Also, data on intra-EU trade is censored in a way such that only firms exporting goods with a total annual value exceeding a certain threshold\(^6\) are recorded in the files. No such data limitations exist for trade out of the EU. As a result the coverage rate in the Intrastat system is lower but still in the range

\(^6\)For the years considered, this threshold was DKK 2.5 million corresponding to approximately USD 500,000.
This study examines Danish manufacturing exports in 2003 (for some robustness checks below we also include data from 2001 and 2002). We select all manufacturing firms with positive inputs of labor and capital and with positive export sales. Also, we consider only manufacturing products by selecting products in one-digit SITC categories 5, 6, 7 and 8. With these restrictions our 2003 dataset comprises 155,426 firm-destination-product sales observations by 4,304 firms in 5,339 eight-digit CN8 products to 223 destination countries, see Table 1. The aggregate value of all these trade flows totals 182 billion Danish kroner (DKK), which in 2003 roughly corresponded to USD 28 billion.

The product level dimension of our data allows us to examine the contribution of firm and destination specific effects in sales variation within each product.\textsuperscript{7} To identify the contribution of firm and destination specific effects, we need positive degrees of freedom for each product, such that the number of firm-destination trade flows exceeds the number of firms plus the number of destinations. Therefore, in the following, we will only consider products satisfying this requirement. With this restriction we end up with just 1,075 CN8 products or 880 HS6 products. However, they constitute 74 or 81\% of the overall trade volume respectively, see Table 1. For the restricted samples there is not much difference between the HS6 and CN8 levels. In the following we focus on the HS6 level as it covers the largest portion of the total Danish export volume, but we also report results for the CN8 level.

By focusing on exporters only we are not capturing productivity differences between exporters and non-exporters, and this might underestimate the importance of firm specific effects in our empirical approach. The 4,024 firms in our final estimation sample (the last column of Table 1) employ more workers per firm, exhibit higher total sales, pay higher average wages and are somewhat more productive measured in terms of value added per

\footnote{The next section explains how this approach is consistent with a simple multiproduct firm trade model.}
worker than all the remaining manufacturing firms. However, these 4,024 firms account for the bulk of total sales and employment in the manufacturing sector (85% and 81% respectively), so our results are representative of a large majority of Danish manufacturing.

Table 1 shows some similarities between exporters in Denmark and those in bigger economies. The median number of destinations for firm-product exports is 1, which is in line with the findings for the US (Bernard and Jensen, 1995) and France (Eaton, Kortum, Kramarz, 2004). Clearly, some firms ship their products to many destinations – the mean number of destinations is 4 and the maximum number is 138.

At the disaggregated eight-digit product level most destinations do not have many Danish firms present. The median number of firms is 2 and the mean is 2.9 for the restricted CN8 sample. At the slightly more aggregated six-digit Harmonized System (HS6) level the mean number of firms is 3.3 and the median is 2. Table 1 also shows the extent of truncation in the data. We flag a sales observation for a firm-product variety in a destination as *missing* if there are no sales of that variety in that destination but at least one other Danish firm sells that product to that destination. For the restricted HS6 sample the median product has 87% missing trade flows, and even the product at the 10 percentile has 72% missing trade flows.

### 2.1 An illustrative example

To aid the reader in understanding our goal, we begin with an illustrative example. Suppose Denmark exports to only two destinations: Sweden and Germany. Models based exclusively on productivity heterogeneity predict that firms that sell to both destinations should have relative revenues that are one-to-one correlated. If a Danish brick firm’s sales to Germany are twice the average of all Danish brick firms selling to Germany, we can say that this firm is twice as productive as average. The same firm’s Swedish sales should also be twice the average. The one-to-one correlation predicts that the variation in German relative revenues should completely explain the variation in Swedish relative revenues.

Figure 1 depicts demeaned (log) revenues for Danish firms exporting building bricks
and plastic boxes to Sweden and Germany.\textsuperscript{8} The revenues are relative to the mean Danish firm revenues of the respective product to the respective destination. The straight lines show the slopes from OLS regressions of relative revenues in Sweden on relative revenues in Germany. The results for plastic boxes support a weaker interpretation of models based on productivity heterogeneity: that relative revenues are positively correlated. The slope, although statistically different from one, is still high at 0.84. Another measure of interest is the proportion of variation in Swedish sales that is explained by the variation in German sales, i.e., the $R_2^2$. For plastic boxes, the variation in German relative revenues explains 54\% of the variation in Swedish relative revenues. In contrast, the OLS results for building bricks suggest a different story. The implied correlation is negative and not statistically different from zero. The $R_2^2 = 0.08$ shows that little of the Swedish variation is explained by the German variation.

Insert Figure 1 here

Of course, this simple example ignores several estimation issues that will be carefully dealt with in the main part of the paper. First, not all firms in the sample above export to both Sweden and Germany. For example, of the firms that sell plastic boxes to Sweden or Germany, only around half sell to both. Since it is likely that nonexistent exports to a destination is highly correlated with low potential sales in that destination, this truncation biases the results.

Second, while we examine the elasticities (the slopes) between sales in different destinations more thoroughly below, our primary focus is on how well the variation in firm-specific effects explains sales variation, so we will not rely on the estimated slope as a measure of the contribution of productivity to relative revenues. One reason is that the estimated slope could be negative, which is not interpretable. Another reason is that the

\textsuperscript{8}The two products are more precisely "Building bricks (excl. those of siliceous fossil meals or similar siliceous earths, and refractory bricks of heading 6902)" (CN8 product 69041000) and "Boxes, cases, crates and similar articles for the conveyance or packaging of goods, of plastics" (CN8 product 39231000). Sweden and Germany are the two most popular destinations for Danish exporters.
slope does not tell us the contribution of productivity variation to the total variation. To illustrate this point, consider Figure 2, which presents three graphs of simulated relative revenues in two destinations. All three scatters have fitted slopes of 1, but different $R^2$ values. In the top left graph there is almost a perfect one-to-one correspondence between relative revenues in the two destinations (the $R^2$ is set to 0.999), while in the two other graphs deviations from a one-to-one correspondence are introduced (the $R^2$ is set to 0.5 and 0.12 respectively). If productivity heterogeneity is the sole source of the variation, we should expect to see scatters similar to that in the upper left panel of Figure 2. As the contribution of firm-destination specific heterogeneity rises, the scatters begin to look more like the ones in the upper right and bottom left of Figure 2. Therefore, estimated slopes close to unity are misleading confirmations of models based on firm-level productivity differences. Instead we will focus on a statistic similar to $R^2$ as our measure of the contribution of firm productivity to sales variation.

Insert Figure 2 here

### 2.2 Preliminary evidence

In this section we provide some preliminary evidence for the importance of firm effects and firm-destination effects in explaining the variation in export sales using the full sample. The illustrative example showed a product with a negative correlation between firm-level sales in two destinations. To investigate the prevalence of such negative correlations in our sample, we first run product-specific regressions of firm-destination export flows on firm-level total domestic sales. Figure 3 displays a histogram of the product-specific slopes. It is seen that a substantial mass of slopes are around 1 as predicted. However, 33% of exported product flows, accounting for 30% of total exports, have a slope significantly different (at the 1% significance level) from 1, suggesting that something other than firm

---

9The simulation in Figure 2 is based on 2,000 instances of variables $x$ and $e$, each drawn from a standard normal distribution. To produce the $R^2 = 0.999$ cloud, we generated a variable $y = x + 0.01e$ and plotted $y$ against $x$. To produce the $R^2 = 0.5$ cloud, we generated $y = x + e$. Finally, to produce the $R^2 = 0.12$ cloud, we generated $y = x + 2.666e$. 

9
productivity – for example a firm-destination component – must play a role for some products. Even more incongruous with the single-source firm heterogeneity models are the 16.5% of exported product flows which exhibit negative slopes. These flows account for 6% of total exports and are entirely counterintuitive to the predictions of Melitz (2003).\footnote{Approximately 4\% of the flows (accounting for 1\% of the export volume) have slopes that are significantly negative.}

Insert Figure 3 here

As a first attempt to address the importance of the firm-destination component in a simple way, we run the following two regressions for each product:

\begin{align}
  r_{\omega jt} &= \alpha_{jt} + \alpha_{\omega} + e_{\omega jt} \tag{1} \\
  r_{\omega jt} &= \beta_{jt} + \beta_{\omega j} + u_{\omega jt}. \tag{2}
\end{align}

In equation (1) firm-destination-year specific export sales are regressed on destination-year fixed effects and firm fixed effects. In (2) the firm fixed effects are replaced with firm-destination fixed effects. The firm-destination-year effects $e_{\omega jt}$ and $u_{\omega jt}$ are residuals from this naive regression. Since the set of covariates in equation (2) expands on the set in (1), the estimation of $R^2$ from (2) is weakly greater than from (1), and the difference in $R^2$s between the two equations can be naively interpreted as the contribution of the firm-destination component over the firm component in the variation in export sales. To estimate the equations we need to use data for multiple years, 2001-2003, and to mitigate the truncation issues discussed above, we use unlogged export sales, which allows us to include zero export flows. For comparison we also run regressions excluding the zero export flows.

The top half of Table 2 shows the distribution across products of the $R^2$s from running these two equations with zeros included. The median $R^2$ in equation (1) is 0.166 and the median $R^2$ in equation (2) is 0.924. We also calculate the difference between the two $R^2$s for each product, and the median difference is 0.721, while even at the 25th percentile the
difference is almost two thirds. This means that the firm-destination component typically explains the majority of the variation in firm-level exports.

The bottom half of Table 2 shows the results when we drop zero trade flows. It is evident that the explanatory power of model with only firm fixed effects in equation (1) increases substantially such that the difference between the two models narrows down. However, much sales variation is still explained by the inclusion of firm-destination fixed effects in model (2). These results suggest that the firm-destination component is important and that truncation potentially plays an important role.

The results reported so far indicates that the tight relationship between the firm component and export sales in a given destination predicted by standard trade theory frequently is violated. This raises the question whether certain firm and destination characteristics tend to account for these violations. A simple way to address this question is to calculate a residual measuring if trade is higher or lower than predicted and then correlate this residual with firm and destination characteristics. We calculate the residual in the following way:

$$RESI_{\omega j} = (\log r_{\omega j} - \log \tau_j) - (\log rt_{\omega} - \log \bar{r}\bar{t}_j),$$

where $\tau_j$ is the average revenue in destination $j$ among all firms selling in that destination, $rt_{\omega}$ is firm $\omega$’s total export revenue across all destinations and $\bar{r}\bar{t}_j$ is the average total revenue among all firms selling in destination $j$. If the residual is positive, the firm is selling more in the destination than predicted by its point in the distribution of total export sales, and vice versa if it is negative.

We relate the residual to firm and destination characteristics by running a regression of the residual on firm size, value added per worker, the share of high skilled workers in the firm, destination country GDP, population and distance, see Table 3. It is evident that
larger, more productive, and more skill intensive firms tend to have lower residual trade values, implying that "better" firms actually have smaller trade flows than their overall distribution would suggest. In addition, residual trade values are negatively correlated with GDP and population and positively correlated with distance. Thus, residual trade values run counter to the predictions of the gravity model. One interpretation of these results is that the scope of a firm’s exports outweigh its scale. That is, a firm’s exports flows are dominated by small flows. For a “good” exporter, its low-volume export flows to tiny or distant destinations are more numerous than its high-volume export flows, which may be to one or two larger or closer destinations. We emphasize that since we calculate the residual trade value off of just the firms in the destination, these results suffer from truncation bias. In the remainder of the paper, we introduce and use an estimation strategy that controls for this truncation bias. A primary contribution of this paper is to recognize and mitigate the truncation bias introduced by the self-selection of exporters into markets, as described in the following section.

Insert Table 3 here

3 Theory and sales variance decomposition

This section specifies a standard productivity heterogeneity model, which we extend with destination specific effects such that firm-level sales may be decomposed into a firm-specific component and a firm-destination specific component. This decomposition allows us to construct our statistic of interest, i.e., a $R^2$-based measure for the contribution of firm-specific effects to overall sales variation. It will be shown that equations (1) and (2) from the simple way of assessing the contribution of firm-specific effects above are not consistent with the standard trade model, and the model also shows how to properly account for truncation.

Consider a small country exporting $N$ products. For each product $n \in N$, there are
W_n firms each producing a unique variety \( \omega \) to sell to \( J_n \) foreign country destinations.\(^{11}\) Not all varieties are exported to all destinations; only \( W_{nj} < W_n \) firms supply to destination \( j \in \{1..J_n\} \). The selection of varieties into destinations are determined by the variety’s potential destination specific sales \( r^*_{\omega jn} \). A simple extension of Melitz (2003) or re-characterization of Bernard Redding and Schott (2011) decomposes the log potential sales \( \ln r^*_{\omega jn} \) into a destination specific component \( a_{jn} \) and a firm specific component \( b_{wn} \). The residual \( x_{\omega jn} \) comprises the firm-destination specific effects that can be explained by neither \( a_{jn} \) nor \( b_{wn} \):\(^{12}\)

\[
\ln r^*_{\omega jn} = a_{jn} + b_{wn} + x_{\omega jn}.
\]

The destination specific component \( a_{jn} \) captures all characteristics that affect sales in that destination, such as distance and GDP.

The firm specific effect \( b_{wn} \) represent those firm characteristics that contributes the corresponding variety’s sales in all destinations. These include lower costs (Melitz 2003), higher quality (e.g. Baldwin and Harrigan 2011, Kugler and Verhoogen 2012), better capability (Johnson 2012), or any other firm specific characteristic discussed in the recent firm-heterogeneity trade literature. In the literature and this model, firm specific effects are drawn from exogenous and independent distributions with mean \( \bar{b}_n \) and variance \( s_{bn}^2 \).

Finally, the residual \( x_{\omega jn} \) captures sales variation that cannot be explained by gravity-style country specific effects \( a_{jn} \) or firm-heterogeneity-type firm specific effects \( b_{wn} \). Within each product \( n \), this residual represents firm-destination specific characteristics. Recent theoretical literature has offered explanations for the source of this residual. Das, Roberts and Tybout (2007) model foreign demand shocks that could be due to fluctuating foreign income or real exchange rates. In addition to market specific demand shocks Eaton, Kortum and Kramarz (2011) also incorporate market and firm-specific entry costs in their

\(^{11}\)A firm may produce multiple products, but following e.g. Bernard, Redding and Schott (2011) we assume independence of production and demand across products, which allows us to treat each firm-product as a unique variety.

\(^{12}\)In this model, multi-product firms can be characterized as a collection of single-product firms by allowing the productivity to vary across products as in Bernard, Redding Schott (2011). \( b_{wn} \) can be interpreted as a firm-product specific component in the framework of Bernard, Redding Schott (2011). The model details are relegated to Appendix A.
model. Nguyen and Schaur (2012) and Vannoorenberghe (2012) posit that firms’ cost functions exhibit increasing marginal costs and that they face market-specific demand shocks. In such a setting firms will react to a positive demand shock in one market by expanding sales here while reducing sales in other markets. Related to this Blum, Claro and Horstmann (2013), Blum, Claro, Horstmann and Tombe (2013), Rho and Rodrigue (2014) and Soderbery (2014) show that similar linkages between markets may arise if firms are capacity constrained and exposed to market specific demand shocks. Finally, in Nguyen (2012) firms face destination specific perceived quality draws. Here firms will use realized sales in supplied markets to forecast sales in unsupplied markets. In this framework it is shown that firms delay exporting to some markets and that firms sequentially enter more and more export destinations. This current study does not distinguish among these possible explanations but quantifies their joint power in explaining firm-level sales variation.

Given that \( a_{jn} \) and \( b_{\omega n} \) captures all sales variation specific to the destination and to the firm, any non-zero central tendency in the firm-destination effect \( x_{\omega jn} \) is swept up by \( a_{jn} \) and \( b_{\omega n} \). Since we are interested in the variance and not the mean of the firm-destination effect, we do not separately identify \( \bar{x}_{jn} \) from \( a_{jn} \) or \( \bar{x}_{\omega n} \) from \( b_{\omega n} \); we make the simplifying assumption that \( \bar{x}_{jn} = \bar{x}_{\omega n} = 0 \). Our decomposition methodology also requires structure on the distribution of \( x_{\omega jn} \), and we assume \( x_{\omega jn} \) is normal with variance \( s_{xn}^2 \). Our normality assumption is supported by the distribution of domestic revenues of Danish firms presented in Figure 4 and is consistent with previous studies of firm size distributions (Cabral and Mata 2003) and export selection.\(^{13} \)

\(^{13}\) A growing theoretical literature approximates firm productivity with the Pareto distribution. This is to some extent driven by the analytical tractability of this distribution, see e.g. Chaney (2008), Helpman, Melitz and Yeaple (2004) and Eaton, Kortum and Kramarz (2011) as well as some empirical literature (Axtell 2001). The latter two use log-normal error terms to better fit the data. In these studies, the sales distributions deviate from Pareto in a way that is indicative of a truncated log normal distribution: first, the curvature is concave, not convex like Pareto, and second, the mass of firms with very low sales is too large to fit the Pareto distribution.
Consistent with the assumptions of the standard linear regression model, the firm-destination-specific draw, \( x_{\omega jn} \), is constructed to be orthogonal to the destination specific component, \( a_{jn} \), and the firm specific component, \( b_{\omega n} \).\(^{14}\) Given our normality assumption on \( x_{\omega jn} \), and equation (3), the variance of the within-destination sales distribution can be expressed as the sum of the variances of the firm specific effect and the firm-destination specific effect, \( s_{b n}^2 + s_{x}^2 \). Our goal is to estimate the contribution of the firm specific effect to the variance of potential sales for firms within a destination, controlling for destination specific effects. That is, we estimate the statistic

\[
Q_n^2 = \frac{s_{b n}^2}{s_{b n}^2 + s_{x}^2},
\]

(4)

for each Danish product exported in 2003. Notice the resemblance between \( Q_n^2 \) and the adjusted coefficient of determination, \( \bar{R}_n^2 \).

### 3.1 Truncation

If sales were observed for every firm-destination pair, a simple ANOVA of \( \ln r_{\omega jn}^* \) on destination and firm specific effects would consistently decompose the variance into that explained by \( a_{jn} \) and that explained by \( b_{\omega n} \), with the residual being attributed to \( x_{\omega jn} \). However, our dataset is an unbalanced panel where not every firm sells to every destination. The firm-destination sales, \( r_{\omega jn}^* \), is truncated, with the truncation endogenously correlated with \( a_{jn}, b_{\omega n}, \) and \( x_{\omega jn} \). This source of bias precludes the use of ANOVA to measure \( Q_n^2 \) or \( \bar{R}_n^2 \).

Melitz (2003) suggests that the presence of a firm in a destination is tied to its potential profit in that market. Following Melitz (2003), we assume firm \( \omega \)'s profits, \( \pi_{\omega jn} \), gained

\(^{14}\)If this assumption fails, then that would produce biased estimates of \( b_{\omega n} \), but our goal is not to estimate \( b_{\omega n} \). Any correlation between the firm-specific component and the firm-destination specific component will be absorbed by the firm-specific component. For example, suppose \( b_{\omega n} \) represented a productivity shock and \( x_{\omega jn} \) comprised a firm-destination demand shock. A positive correlation between the two components means that highly productive firms also have high average demand shocks. Variation in the firms’ average demand shocks are indistinguishable from variation in the firms’ productivity shocks, and we treat them both as the firm-specific component.
from supplying product $n$ to destination $j$ are

$$
\pi_{\omega j n} = \frac{1}{\sigma} r^{*}_{\omega j n} - f_{j n}.
$$

where $\frac{1}{\sigma} > 0$ is the portion of firm $\omega$’s sales over its variable cost of supply and $f$ is the fixed cost of supply to destination $j$. Profits are positive when $r^{*}_{\omega j n} > c_{j n}$, where $c_{j n} = \sigma f_{j n}$ is a destination specific sales cutoff for each product. Firms will only export products to destinations where they garner positive profits. Therefore, we only observe $r_{\omega j n}$, where

$$
r_{\omega j n} = \begin{cases} 
    r^{*}_{\omega j n} & \text{for } r^{*}_{\omega j n} \geq c_{j n} \\
    0 & \text{for } r^{*}_{\omega j n} < c_{j n}.
\end{cases}
$$

Equation (6) given (3) is the standard Type 1 Tobit Model with latent effects described in Honoré and Kyriazidou (2000). In an earlier work, Honoré (1992) shows that if the latent effects, which in our case correspond to $a_{j n}$ and $b_{\omega n}$, are correlated with the probability of truncation, then the Heckman (1979) two-step procedure is biased. Honoré’s solution to this problem treats the latent effects as nuisance variables and differences them out. We cannot use this approach for two reasons. First, the method renders the effects immeasurable, while in our study, $b_{\omega n}$ is a parameter of interest. Second, in our case the unobserved heterogeneities are two-dimensional, while Honoré and Kyriazidou (2000) only assumes one-dimensional unobserved heterogeneity.

Instead, for every product, $n$, we estimate $a_{j n}$, for each destination country, $j$, $b_{\omega n}$ for each firm, $\omega$, and $s_{\tau n}^2$ using a Bayesian Markov Chain Monte Carlo (MCMC) procedure.\footnote{We also tried a Monte Carlo Expectation-Maximization Maximum Likelihood Estimation (MCEM) method proposed by Walker (1996), but simulation results clearly favored the Bayesian MCMC approach. In a trade context the Bayesian MCMC method has also been employed by e.g. Das, Roberts and Tybout (2007) and Roberts, Xu, Fan and Zhang (2012).} We then find the sample variance of $b_{\omega n}$ and use that and the estimated $s_{\tau n}^2$ to calculate $Q_{n}^2$.\footnote{We also tried a Monte Carlo Expectation-Maximization Maximum Likelihood Estimation (MCEM) method proposed by Walker (1996), but simulation results clearly favored the Bayesian MCMC approach. In a trade context the Bayesian MCMC method has also been employed by e.g. Das, Roberts and Tybout (2007) and Roberts, Xu, Fan and Zhang (2012).}

16
4 Bayesian Markov Chain Monte Carlo procedure

This section briefly outlines how we estimate the portion of within-country sales variance contributed by firm specific effects using the Bayesian MCMC algorithm with details of the procedure left in Appendix B. The outcome of the algorithm is a similar object to $R^2$ as measured in equation (4).

In Appendix B we first derive the likelihood function based on the model in equation (3) disregarding the truncation issue. This involves determining the conditional density of observing revenues for each firm-destination pair. Next, truncation is introduced by following the truncation model in equation (6). The algorithm then generates a Markov chain of $a_n$, $b_n$, and $s_n$ samples, each of which are approximately drawn from the probability distributions that maximizes the likelihood of observing the set of realized sales. At each link in the chain, the triplet $a_n$, $b_n$, and $s_n$ is iteratively drawn from distributions parameterized by the prior draws of the triplet. This is repeated 100,000 times. After the first 25,000 burn-in draws are discarded, the final parameter estimates are means of the remaining 75,000 draws, such that we can estimate $Q^2_n$ with $\hat{Q}^2_n$:

$$\hat{Q}^2_n = \frac{VAR(\hat{b}^2_n)}{VAR(\hat{b}^2_n + \hat{s}^2_{xn})},$$

where $\hat{b}^2_n$ and $\hat{s}^2_{xn}$ are the means of the parameter draws across the iterations.

The next step is to verify our procedure’s ability to accurately estimate $Q^2_n$ under various conditions.\textsuperscript{16} We simulate a large number of datasets of firm-destination sales $r_{\omega_{jn}}$ using our model given by equation (6) given (3). For each dataset, we impose truncation by dropping sales below a certain cutoff to simulate our truncated population of firm-destination sales data.\textsuperscript{17} For each dataset, we define a true $\tilde{Q}^2$ in the range $(0.1, 0.9)$.

\textsuperscript{16}Again, details of the simulation procedure are found in Appendix B.

\textsuperscript{17}In our base sample the median product has 87\% zero trade flows, and even the product at the 10 percentile has 72\% zero trade flows (see Table 1). To make sure that our BMCMC procedure was in line with the data, we chose parameters to achieve missing trade flows between 62\% and 93\% in our simulation. Our simulations had an average of 82\% missing trade flows, with a standard deviation of 6\%.
With the given $\tilde{Q}^2$ and a fixed total sales variation to be explained we can determine the variances $s^2_b$ and $s^2_x$ from equation (4). With these in hand we draw $b_{w_n}$, $x_{w_n}$ and $a_{j_n}$ from the known distributions, such that we can generate sales and likelihood contributions following equations (6) and (3). Finally we can then calculate the estimated $\hat{Q}^2$ based on the MCMC procedure described above. This allows us to compare the estimated $\hat{Q}^2$ with the known true $\tilde{Q}^2$, and further, we can compare with the estimates obtained by OLS.

The results of our simulations are summarized in Figure 5 below. Our estimated $\hat{Q}^2_n$ tracks the true value $\tilde{Q}^2_n$ (plotted on the 45° line) well: the root mean square deviation of $\hat{Q}^2_n$ from $\tilde{Q}^2_n$ was 6%. On the other hand, the OLS estimate, $\bar{R}^2_n$, was consistently below the true value, with a root mean square deviation of 29%.

5 Estimation results

We use the Bayesian MCMC procedure to obtain an estimate, $Q^2_{MCMC}$, for the contribution of firm specific effects for each product exported by Danish firms in 2003. We also perform OLS dummy regressions of destination-mean-differenced observed revenues $\left( \ln r_{wjn} - \frac{1}{W_{nj}} \sum_{\omega=1}^{W_{nj}} \ln r_{\omega jn} \right)$ on firm fixed effects. From the OLS regressions, we retrieve the adjusted coefficient of determination $\bar{R}^2$ and estimate $Q^2_{OLS}$ by\textsuperscript{18}:

$$Q^2_{OLS} = \max \left\{ 0, \bar{R}^2 \right\}.$$  \hspace{1cm} (7)

$Q^2$ is defined as a positive number, so we treat negative $\bar{R}^2$ values as estimates of 0 for $Q^2_{OLS}$. In the following, we compare the MCMC estimates $Q^2_{MCMC}$ to the OLS estimates $Q^2_{OLS}$, with the main difference being that truncation is unaccounted for in $Q^2_{OLS}$. Our main results are derived at the detailed HS6 product level, which is followed by a number of robustness checks.

\textsuperscript{18}We use the adjusted coefficient of determination to avoid small sample bias. Cramer (1987) shows that the unadjusted $R^2$ is heavily biased upwards for small samples.
5.1 Product level

In this section we use the restricted HS6 product level sample described above. This sample has a total of 119,217 firm-destination-product observations, spanning 4,024 firms in 880 products to 223 destination countries and totalling DKK 147 billion, see Table 1. We estimate $s_b^2, s_x^2,$ and consequently $Q^2_{MCMC}$ for each of these 880 products.

Our estimation procedure resulted in mean and median values of 49% and 45% for $Q^2_{MCMC}$ across the 880 HS6 products, see Table 4. This is considerably higher than OLS estimates, which resulted in mean and median values of 35% for $Q^2_{OLS}$. The downward bias of the OLS estimates is consistent with the simulation results in Figure 4. For comparison, Lawless and Whelan (2008) obtain an $R^2$ of 41% across their sample of Irish exporters.

We also estimate the contribution of firm-specific effects to the volume of Danish exports, by weighting each $Q^2_{MCMC}$ by the value of the total export volume for each product. Table 4 shows that the mean and median values for $Q^2_{MCMC}$ to 53% and 51%, while the OLS estimates remain almost unchanged at 35% and 36%. This shows that the firm-specific component as measured by the MCMC estimator plays a somewhat bigger role in products with relatively high trade volumes. In contrast, $Q^2_{OLS}$ appears not to be correlated with the importance of the product in terms of its export volume.

It should be noted that the product level dimension of the data is important, as the estimated $Q^2_{MCMC}$’s exhibit substantial variation across products. Histograms for the MCMC and OLS estimates are presented in Figure 6 below: The histogram of the MCMC estimates is systematically to the right of that of the OLS estimates again indicating that truncation biases the OLS estimates downwards.

Another way to assess the properties of the MCMC and OLS estimators is to look for
consistency over time. A regression of 2003 export flows at the firm-country-CN8 product level on corresponding 2001 export flows find a beta of 0.80. That is, a 10% increase in the 2001 exports of a product to a country by a firm corresponds to an 8% increase in the exports of the same product-country-firm triplet in 2003. When we restrict the regression by removing the constant, the coefficient increases to 0.98, suggesting strong persistency in the export patterns between 2001 and 2003. Given this consistency of export flows across time, one would expect that the $Q^2$s at the product level are fairly stable over time. To investigate this, we repeat the BMCMC estimation for the year 2001 and find similar results. For the 917 HS6 products that fit our restrictions in 2001, we obtain median estimates of $Q^2_{MCMC} = 49\%$ versus $Q^2_{OLS} = 34\%$.

The $Q^2$ estimates are not only correlated in the aggregate, but at the individual product level. There were 739 HS6 products that passed our estimation restrictions in both 2001 and 2003. The left panel of Figure 7 presents the point estimates for the two years. The strong correlation between $Q^2_{MCMC,2003}$ and $Q^2_{MCMC,2001}$ contrasts with the lack of correlation between OLS estimates $Q^2_{OLS,2003}$ and $Q^2_{OLS,2001}$.

The right panel of Figure 7 shows this lack of consistency across years. We take this exercise as further evidence that our procedure accurately identifies the contribution of firm specific effects, while OLS estimates do not.

Insert Figure 7 here

5.2 Sample restrictions

We now proceed with a number of robustness checks of the Bayesian MCMC estimates. As explained in the data section above, the restricted sample used in the estimations above is the largest possible given that we need positive degrees of freedom for each product, i.e., the number of firm-destination trade flows must exceed the number of firms

\footnote{We regressed $Q^2_{MCMC}$ for 2003 on that for 2001 for the 739 products. Our estimated marginal effect was 0.41 with a standard error of 0.035. That is, a 10% increase in $Q^2_{MCMC,2001}$ corresponded to a 4% increase in $Q^2_{MCMC,2003}$. When we restrict our regression constant to zero we get an estimated marginal effect of 0.97 with a standard error of 0.01. A regression of the two OLS estimates resulted in no significant correlation between the two.}
plus the number of destinations. This way we keep products that contain only very few observations. One may be concerned that our Bayesian MCMC procedure is unable to handle such products very well. To investigate this, the first two rows of Table 5 report the distribution of $Q^2_{MCMC}$ among products with at least 100 observations and with at least 200 observations. These restrictions reduce the number of products substantially, but it is evident that the distribution is almost unchanged. The mean and median values of $Q^2_{MCMC}$ are in both cases 49% and 45% just as for our base specification in Table 4.

Insert Table 5 here

5.3 Aggregation

As briefly mentioned in the introduction, our dataset contains product code information and we have found substantial cross-product variation in the contribution of firm-specific effects to sales variation. This level of disaggregation is not contained in the data used by Eaton, Kortum and Kramarz (2011) or Lawless and Whelan (2008). Instead they report the overall contribution of firm-specific effects to sales variation.

To examine if the level of product aggregation matters for our results, we also estimate $Q^2_{MCMC}$ at the CN8 product level, the most disaggregated level available to us. The results are close to similar to estimates performed at the HS6 level. We obtain a mean and median of 50% and 47% (last row of Table 5), which is slightly higher than in our base results of Table 4.

To better compare our results to Eaton, Kortum and Kramarz (2011) and Lawless and Whelan (2008), we also aggregate our data to broader industries and estimate $Q^2_{MCMC}$ at the HS2 industry level. For the 59 HS2 industries, we obtain mean and median estimates of 48% and 43% for $Q^2_{MCMC}$, see Table 5. This result is in line with our previous estimates at the HS6 product level. Sixteen industries have estimates of $Q^2_{MCMC}$ greater than 80%. There was no obvious pattern why these industries exhibited higher contributions of firm specific effects. These results suggests that, if anything, the estimated contribution of
firm specific effects rises with the level of aggregation.

5.4 Established exports

Nguyen (2012) suggests that much of the export sales variation is due to firms testing destinations in order to determine whether they can be successful exporting to that destination. Therefore, firm-destination specific effects should play a larger role in the first year of exporting. To test that, we restrict our sample to only those firm-product-destination observations in 2003 that were also positive in 2002. That is, only 2003 exports by those firms that exported the same product to the same destination in both 2002 and 2003 were considered. This restriction leaves us with 66,782 observations spanning 698 HS6 products totalling DKK 131 billion, or about 72% of the 2003 trade flows.

The predictions from Nguyen (2012) are supported by the data. For the 698 established exports in 2003, we obtain mean and median values of 52% for $Q^2_{MCMC}$, see the fifth row of Table 5. These values are 8–13% higher than those estimates estimated for the sample which included first time exports. Therefore, firm specific effects appear to be more important for these established exports. By contrast, firm-destination specific effects are more important for the first year of exporting than for established exports.

5.5 Core products

Firms typically export multiple products, and for such firms the within-firm output distribution across products is known to be highly skewed with typically one core product accounting for a major part of firm sales, see e.g. Bernard, Redding, and Schott (2010). If non-core products are more likely to be sold in destinations where fixed costs related to sales of the core product already have been incurred, the proportion of non-core product sales variation explained by the firm component may be smaller.

To investigate this, we repeat our exercise for only the core product of each firm. We define a firm’s core product as the HS6 category constituting the highest export sales. We drop all other products exported by that firm. With this and the forementioned
restrictions, we are left with 25,210 observations spanning 297 products totalling DKK 87 billion, or about half of our total 2003 trade flows.

The MCMC estimates for core products are higher than those for all products. The median $Q_{MCMC}^2$ is 57% for the 297 HS6 categories comprising only core products, see Table 5. For these same 297 HS6 categories, the median $Q_{MCMC}^2$ is 47% when we include all non-core export flows.

5.6 Product characteristics

Since our $Q_{MCMC}^2$’s are product specific, we investigate whether there are any patterns in the $Q_{MCMC}^2$’s across products. Our theoretical model is stylized and does not give us any predictions about how $Q_{MCMC}^2$ varies with product characteristics, so a priori we do not have any expectations about any relationships. However, we regressed our estimates on several product-level characteristics to investigate any possible relationship.

First, we examined the relationship between $Q_{MCMC}^2$ and the elasticity of substitution. A weighted (by total exports) regression of $Q_{MCMC}^2$ on industry level elasticities of substitution taken from Broda and Weinstein (2006) finds a statistically significant elasticity of 0.03. That is, products from more homogenous industries (higher Broda-Weinstein sigmas) have their sales variances explained more by the firm-specific component, rather than the firm-destination-specific component.

We also regressed $Q_{MCMC}^2$ on the output-weighted means and standard deviations of the capital labor ratio\textsuperscript{20}. We find that a doubling of the mean capital labor ratio for a product corresponds to a 3.7% increase in $Q_{MCMC}^2$ (with a standard error of 1.8%). A doubling of the standard deviation decreases $Q_{MCMC}^2$ by 5.45%. Assuming the spread of capital labor ratios within a product is indicative of its degree of differentiation, this result is consistent with our Broda Weinstein (2006) regressions.

Finally, we partitioned our results by the Rauch (1999) classification of product differ-

\textsuperscript{20}For each of the 880 HS6 products $n$, we multiply each firm’s capital labor ratio by its share of of total exports of $n$ to arrive at our output-weighted capital labor ratio. These are summed up to determine the mean capital labor ratio for $n$. 

23
entiation. Of our 880 products, 657 are classified as ‘differentiated’ and 98 are classified as ‘reference priced.’ The ‘differentiated’ products had a median $Q^2_{MCMC}$ of 46% while the ‘reference priced’ products had a median $Q^2_{MCMC}$ of 43%. However, the distributions largely overlapped, so we refrain from speculating about any true differences.

6 Discussion and conclusion

We use a highly detailed dataset for Danish exporters to estimate the contributions of firm specific and firm-destination specific effects to the variation of sales within a product-destination market. We find that the contribution of firm specific effects varies greatly across products, and that it explains less than 45% of the variation for over half of Danish HS6 products. Our results suggest that firm-destination specific heterogeneity, rather than a firm specific effect such as productivity, captures the majority of heterogeneity for most products and is the primary driver of variation in a market.

We also show that OLS estimates tend to underestimate the contribution of the firm-specific component. To consistently estimate firm specific effects, we employ a Bayesian Markov Chain Monte Carlo strategy, and we argue that this method can be employed fruitfully in studies of firm-level exporting with truncation issues.

The Melitz (2003) model or multiproduct firm versions such as Bernard, Redding and Schott (2011) use firm specific productivity to deftly explain variation between exporters and nonexporters. Much of the literature has built upon this idea, and it does indeed offer a tractable way of modelling firm heterogeneity. We would like to stress that our results should not be taken as a refutation of the Melitz (2003) model. After all, our results show that firm specific productivity plays an important role in explaining firm heterogeneity for many products. However, the majority of variation is firm-destination specific for most products, which suggests a new direction based on firm-destination specific effects may better reconcile trade patterns.

21 Rauch (1999) also classifies products according to whether they are traded on organized exchanges, but we had only a handful of products of this type in our sample. This is because our dataset contains only manufacturing products.
For the median product, our estimate of the contribution of the firm-specific component in explaining sales variation is somewhat higher than the estimates of Eaton, Kortum, and Kramarz (2011) and Lawless and Whelan (2008). There are several systematic differences between our approaches that could lead to our disparate results.

Our approach is different from Eaton, Kortum and Kramarz (2011) in several ways. The most significant difference is that our study is much more empirically focused. Our approach decomposes sales variation and interprets the estimated heterogeneities in a very broad sense: we want to see what proportion of sales is firm specific and what is firm-destination specific. Eaton, Kortum and Kramarz (2011) introduce a model with stricter interpretations of the sources of heterogeneity and then calibrate their model to match the data. In other words, Eaton, Kortum and Kramarz (2011) interpret the results within the structure of their specific model, while we can interpret our results within a broader class of CES trade models with firm heterogeneity.

Eaton, Kortum and Kramarz (2011) separate their results into that concerning entry variation and that concerning sales variation conditional upon entry. They find that firm specific effects account for 57% of the variation in entry and 39% of the variation in sales. Estimating two different percentages is correct if the exporting decision is akin to a type II Tobit, or Heckman (1979) two-stage model. The theoretical firm heterogeneity trade literature (e.g. Melitz 2003 and Bernard, Eaton, Jensen and Kortum 2003) do not separate the two, however. Current theory suggests that entry is driven by the unconditional sales via a type 1 Tobit model. This paper follows the theoretical literature and estimates the contribution of firm specific effects to the unconditional sales variation. The disparity between our results and theirs could arise from this difference: they measure the firm specific effect’s contribution to conditional sales variation, while we measure its contribution to unconditional sales variation.

Finally, the disparity between Eaton, Kortum and Kramarz’s (2011) and our results may depend on the different productivity distributions assumed. They draw their firm specific effects from a Pareto distribution and their firm-destination specific demand and entry shocks from lognormal distributions. This is to some extent driven by the analytical
tractability of this distribution (see e.g. Chaney 2008). We use lognormal distributions for both our firm specific and firm-destination specific distributions. We use the lognormal distribution because the Pareto distribution does not completely characterize the distribution of sales. Even in studies arguing for the Pareto distribution (Axtell 2001), the sales distributions deviate from Pareto in a way that is indicative of a truncated log normal distribution: first, the curvature is concave, not convex like Pareto, and second, the mass of firms with very low sales is too large to fit the Pareto distribution. To account for these discrepancies, studies such as Helpman, Melitz, and Rubinstein (2008) and Eaton, Kortum, and Kramarz (2011) use log-normal error terms to better fit the data.

Our study improves on the methods pioneered by two other studies looking at the importance of firm specific effects. Kee and Krishna (2008) examine Bangladeshi exports of textiles to the US and EU. They find that a textile firm’s market share in EU cannot predict its market share in the US: the correlation between the two is not statistically different from zero. Lawless and Whelan (2008) use firm-destination data from a survey of 676 Irish-owned exporters to explain to where and how much firms export. Using OLS regressions with fixed effects, they find the variation in firm-year and country specific effects accounts for 57 percent of the total variation. By itself, the country specific effects explain 16 percent of the variation, leaving 41 percent of the variation explained by firm-year specific effects. In comparison, we find that firm-specific effects explain around 35 percent of the sales variation for the median product. In addition, we show that truncation issues bias OLS results and must be accounted for. Honoré and Kyriazadou (2000) discuss that the Heckman (1979) two-step procedure cannot correct for this truncation bias when entry into a destination is related to the firm-destination specific demand draws. Instead, this paper uses a monte carlo estimation maximization procedure to consistently account for truncation and the unobserved effects.

While our analysis is a first step in understanding the relative importance of firm (or firm-product) specific components versus firm-destination specific components of sales,

\[22\text{In other specifications they estimate the explanatory power of observed firm characteristics such as value added per employee and sector dummies instead of firm fixed effects.}\]
it can be extended to look at more complicated models. For instance, we rely on the Bernard, Redding Schott (2011) model to simplify our analysis by allowing productivities to vary independently across products within a firm. This approach ignores within-firm economies or diseconomies of scope. In future research, we can break apart that firm-product component into a firm component and a firm-product component, or define some relationship between the number of products a firm produces and the relative importance of that firm-product component.
References


A A model of sales variation

This appendix presents a simple extension of Melitz (2003) from which the firm level sales equation in section 3 may be derived. Consider a small country exporting $N$ products. For each product $n \in N$, there are $W_n$ firms each producing a unique variety $\omega$ to sell to $J_n$ foreign country destinations. Not all varieties are exported to all destinations; only $W_{nj} < W_n$ firms supply to destination $j \in \{1..J_n\}$. The utility gained in destination $j$ from consuming varieties of product $n$ is represented by $u_{jn}$:

$$u_{jn} = \sum_{\omega=1}^{W_{nj}} \exp \left( \frac{x_{\omega jn}}{\sigma} \right) \left( q_{\omega jn} \right)^{\frac{\sigma - 1}{\sigma}},$$  

(8)

where $q_{\omega jn}$ is consumption of variety $\omega$ in $j$ and $\sigma > 1$ is a measure of the substitutability among the different varieties in $n$. The utility function resembles a Dixit-Stiglitz utility function with a demand shifter $x_{\omega jn}$. This demand shifter represents destination $j$ taste for variety $\omega$. Higher $x_{\omega jn}$ corresponds to greater demand for that variety relative to other varieties in the same destination. Destination $j$’s demand for variety $\omega$ can be derived as:

$$q_{\omega jn} = (p_{\omega jn})^{-\sigma} \exp (x_{\omega jn}) \frac{Y_{jn}}{P_{jn}},$$

(9)

$$P_{jn} = \sum_{\omega=1}^{W_{nj}} \exp (x_{\omega jn}) (p_{\omega jn})^{1-\sigma},$$

(10)

where $p_{\omega jn}$ is the price of $\omega$ and $Y_{jn}$ is $j$’s total expenditure on varieties of product $n$. $P_{jn}$ is the corresponding Chamberlainian price index, which is unaffected by the actions of any single firm.

Firms share similar increasing returns to scale production technologies. Firm $\omega$’s cost $c_{\omega jn}$ of supplying $q_{\omega jn}$ units of output to destination $j$ is

$$c_{\omega jn} (q_{\omega jn}) = f + \exp \left( \frac{b_{\omega n}}{1 - \sigma} \right) \tau_{jn} q_{\omega jn},$$

(11)

Nguyen (2012) defines this parameter as ”perceived quality”. We can also think of it as $\omega$’s popularity or appeal in $j$.  

33
where \( f \) and \( \tau \) are fixed and variable costs identical to all firms supplying \( n \) to \( j \). The firm specific \( \exp \left( \frac{b_{\omega n}}{1-\sigma} \right) \) is the firm’s marginal cost of product that is constant across all destinations. The \( b_{\omega n} \) term is a normalized measure of \( \omega \)’s productivity: a higher \( b \) translates to a lower marginal cost for supplying \( n \) across all destinations.

Each firm \( \omega \in \{1, ..., W_{nj}\} \) draws its firm specific productivity \( b_{\omega n} \). In addition, each firm draws a firm-destination specific taste parameter \( x_{\omega jn} \). The two random variables \( b_{\omega n} \) and \( x_{\omega jn} \) determine firm \( \omega \)’s potential sales \( r_{\omega jn}^* \) in destination \( j \), which is presented in log form:

\[
\ln r_{\omega jn}^* = a_{jn} + b_{\omega n} + x_{\omega jn}
\]

\[
a_{jn} = \ln \left( \frac{Y_{jn}^{1-\sigma}}{P_{jn}} \right).
\]  

(12a) \hspace{1cm} (12b)

The firm productivity draws, \( b_{\omega n} \), are drawn an exogenous distribution, as in Melitz (2003). The firm’s costs \( c_{\omega jn} \) can be subtracted from it sales \( r_{\omega jn}^* \) to generate its potential profit \( \pi_{\omega jn} \) gained from supplying \( \omega \) to \( j \):

\[
\pi_{\omega jn} = \frac{1}{\sigma} r_{\omega jn}^* - f.
\]  

(13)

Firms will only supply to profitable destinations. Therefore, observed sales \( r_{\omega jn}^* \) can be characterized by:

\[
r_{\omega jn} = \begin{cases} 
    r_{\omega jn}^* & \text{for } r_{\omega jn}^* \geq \sigma f \\
    0 & \text{for } r_{\omega jn}^* < \sigma f.
\end{cases}
\]  

(14)
B Bayesian MCMC procedure

B.1 Conditional densities

This section derives the likelihood functions used in our Bayesian MCMC algorithm.24 Let us, for a start, disregard the truncation issue; suppose we observe $\ln r^{*}_{\omega j n}$ for each $\omega j$ firm-destination pair. We define the $(W_n \cdot J_n) \times 1$ vector $r^{*}_n = (\ln r^{*}_{11n}, \ldots, \ln r^{*}_{\omega j n}, \ldots, \ln r^{*}_{W_n J_n})$ as the vector of all firm-destination sales for product $n$. Using our model given by equation (3), we can characterize the likelihood of observing $r^{*}_n$ given parameter vectors $a_n = (a_{1n}, \ldots, a_{jn}, \ldots, a_{J_n n})'$, $b_n = (b_{1n}, \ldots, b_{\omega n}, \ldots, b_{W_n n})'$, and $s^2_{x_n}$. The conditional density of $r^{*}_n$ given these parameters is a joint distribution of $(W_n \cdot J_n)$ normally distributed variables:

$$P(r^{*}_n | a_n, b_n, s^2_{x_n}) = \prod_{\omega=1}^{W_n} \prod_{j=1}^{J_n} \left( \frac{1}{\sqrt{2\pi s^2_{x_n}}} \right)^{-1} \exp \left( -\frac{(\ln r^{*}_{\omega j n} - a_{\omega j n} - b_{\omega n})^2}{2s^2_{x_n}} \right). \tag{15}$$

Using standard Bayesian techniques, we can construct the likelihoods of each parameter in $(a_n, b_n, s^2_{x_n})$ given $r^{*}_n$ and the other parameters:

$$P(a_{jn} | b_n, s^2_{x_n}, r^{*}_n) \sim \left( \frac{1}{\sqrt{2\pi (s^2_{x_n}/W_n)}} \right)^{-1} \exp \left( -\frac{W(a_{jn} - \hat{a}_{jn})^2}{2(s^2_{x_n}/W_n)} \right), \tag{16}$$

$$P(b_{\omega n} | a_n, s^2_{x_n}, r^{*}_n) \sim \left( \frac{1}{\sqrt{2\pi (s^2_{x_n}/W_n)}} \right)^{-1} \exp \left( -\frac{W(b_{\omega n} - \hat{b}_{\omega n})^2}{2(s^2_{x_n}/W_n)} \right), \tag{17}$$

$$P(s^2_{x_n} | a_n, b_n, r^{*}_n) \sim \frac{\theta^{(W_n/2)} \left( \frac{1}{s^2_{x_n}} \right)^{(W_n/2)+1}}{\Gamma(W_n/2)} \exp \left( -\frac{\theta}{s^2_{x_n}} \right), \tag{18}$$

where $\hat{a}_{jn} = \frac{1}{W_n} \sum_{\omega=1}^{W_n} (\ln r^{*}_{\omega j n} - b_{\omega n})$, $\hat{b}_{\omega n} = \frac{1}{J_n} \sum_{j=1}^{J_n} (\ln r^{*}_{\omega j n} - a_{jn})$, $\Gamma(\cdot)$ is the gamma function, and $\theta = \sum_{j=1}^{J_n} \sum_{\omega=1}^{W_n} (\ln r^{*}_{\omega j n} - a_{jn} - b_{\omega n})^2$. Note that $a_{jn}$ and $b_{\omega n}$ have conditional normal distributions and $s^2_{x_n}$ has an inverse gamma distribution.

24Specifically, we use the Gibbs sampler procedure with simulations of the truncated log sales at each iteration. See Casella and George (1992), Robert (1994), de Pooter, Segers and van Dijk (2006), Dorfman (1997) and Greene (2008) Chapter 18 for more information on the Gibbs Sampler and other MCMC techniques. We follow Robert (1994), which examines MCMC under normally truncation variables.
B.2 Truncation

Our conditional densities in (16), (17), and (18) require \( r_n \). However, we only observe the vector \( \mathbf{r}_n = (\ln r_{1n}, ..., \ln r_{\omega jn}, ..., \ln r_{W n \omega n}) \) where \( r_{\omega jn} = 0 \) for \( r_{\omega jn}^* < c_{jn} \). In order to estimate the likelihoods for \( a_{jn}, b_{jn}, \) and \( s_{xn}^2 \), we need to generate a simulated vector \( \mathbf{r}_n^{*(i)} \) using the likelihood function in (15). The reader will notice that (15) requires \( a_n, b_n, \) and \( s_{xn}^2 \). Suppose however, that we had prior estimates of \( a_n, b_n, \) and \( s_{xn}^2 \), which we designate as \( \mathbf{a}_n^{(i)}, \mathbf{b}_n^{(i)}, \) and \( s_{xn}^{2(i)} \). Then we could generate simulated log potential sales for those sales that are unobserved. Based on the likelihood function (15) and truncation model (6), the density of \( r_{\omega jn}^* \) conditional on \( r_{\omega jn}^* = 0 \) is a truncated normal with support \( r_{\omega jn}^* \in (-\infty, c_{jn}) \). Zuehkle (2003) suggests estimating \( c_{jn} \) with the minimum order statistic of the untruncated \( r_{\omega jn} \):

\[
\hat{c}_{jn} = \min \{ r_{\omega jn} | r_{\omega jn} > 0; \omega = 1, ..., W_n \}.
\]

Carson and Sun (2007) prove that \( \hat{c}_{jn} \) converges to \( c_{jn} \) at the rate of \( 1/W_n \). We follow their lead and use \( c_{jn} = \hat{c}_{jn} \). Note that \( c_{jn} \) is best estimated in large samples. Our data includes many instances where \( W_n \) is small. In our results section, we undertake some robustness exercises for various sample sizes and verify that our results are not significantly affected.

Given observed sales \( r_{\omega jn} \), and prior parameters \( \mathbf{a}_n^{(i)}, \mathbf{b}_n^{(i)}, \) and \( s_{xn}^{2(i)} \), we can now generate a potential sales vector \( \mathbf{r}_n^{*(i)} \) using the likelihood function in (15).

B.3 Bayesian MCMC

Now that we have generated \( \mathbf{r}_n^{*(i)} \) using estimated parameters \( \mathbf{a}_n^{(i)}, \mathbf{b}_n^{(i)}, \) and \( s_{xn}^{2(i)} \), we can draw \( a_n \) from the likelihood (16) using \( \mathbf{r}_n^{*(i)}, \mathbf{b}_n^{(i)}, \) and \( s_{xn}^{2(i)} \). We term this draw \( a_n^{(i+1)} \), denoting that our draw of \( a_n \) is the \((i+1)\)th estimate of \( a_n \). We do the same to find \( b_n^{(i+1)} \) and \( s_{xn}^{2(i+1)} \). The drawing of \( \mathbf{a}_n^{(i+1)}, \mathbf{b}_n^{(i+1)}, \) and \( s_{xn}^{2(i+1)} \) using \( \mathbf{a}_n^{(i)}, \mathbf{b}_n^{(i)}, \) and \( s_{xn}^{2(i)} \) completes one iteration in our Bayesian MCMC. Now we can repeat the procedure to generate parameter

\[\text{See Robert (1994) for convergence properties for simulations of truncated normal variables in MCMC estimations.}\]
draws for each iteration \( i \). Since the parameters in each iteration are drawn using the parameters drawn in the prior iteration, the sequence of parameter draws form a Markov Chain. As the number of draws becomes large, the draws \( a^{(i)}_n, b^{(i)}_n, s^{2(i)}_{xn} \) converge to a sample from their joint distribution.\(^{26}\) With enough iterations, we estimate \( Q^2 \) with \( \hat{Q}^2 \):

\[
\hat{Q}^2 = \frac{VAR \left( \hat{b}^2_n \right)}{VAR \left( \hat{b}^2_n + \hat{s}^2_{xn} \right)},
\]

where \( \hat{b}^2_n \) and \( \hat{s}^2_{xn} \) are the means of the parameter draws across the iterations and the \( VAR() \) is the resulting variance of the respective mean (of the firm-specific components).

### B.4 Monte Carlo simulation

Our Monte Carlo simulation procedure consists of the following steps (The Matlab codes is available from the authors upon request):

1. We generate 875 datasets each with 5,000 observations, \((W_n, J_n) = (100, 50)\). We also tried \((W_n, J_n) = (100, 100)\) with similar results.

2. For each dataset we pick a \( \tilde{Q}^2 \) such that they are uniformly distributed over the interval \((0.1, 0.9)\).

3. Obtain \( s^2_b \) and \( s^2_x \) by solving the two equations: \( \frac{s^2_b}{s^2_b + s^2_x} = \tilde{Q}^2 \) and \( s^2_b + s^2_x = 10 \). We also tried different values of the total sales variation to be explained, \( s^2_b + s^2_x = 5 \) and \( s^2_b + s^2_x = 15 \) with similar results.

4. Draw \( a_{jn} \) from a chi-squared distribution with 2, 3 or 4 degrees of freedom for each of the \( J \) destinations. Set \( c_j \) between 1 and 5. These values were chosen so that between 62 and 93 percent of the observations would be truncated. We experimented with various cutoffs in order to generate a simulated dataset with truncation percentages that mirrored the data.

\(^{26}\)See Geweke (1999) for proof and convergence properties of MCMC. We use OLS estimates for the initial iteration.
5. Draw $b_{wn}$ from a $n(0, s_b^2)$ normal distribution for each of the $W$ firms. Draw $x_{wj}$ from a $n(0, s_x^2)$ normal distribution for each of the $J \times W$ observations.

6. Generate $r^*_{wj}$ following the likelihood function in (5) and $r_{wj}$ according to the cutoff condition in (4).

7. Obtain parameter estimates $\hat{a}, \hat{c}, \hat{s}_x^2$ and $\hat{s}_b^2$ via the MCMC procedure described in section 4. Calculate $\hat{Q}^2 = \frac{s_b^2}{s_b^2 + s_x^2}$.

8. Repeat steps 2-7 for 10 trials.

9. Repeat steps 2-8 90 times. This amounts to 875 datasets.

10. For comparison we also calculate the OLS estimates, $\bar{R}_n^2$, obtained regressing

\[
\left( \ln r_{wn} - \frac{1}{W_{n\omega}} \sum_{\omega=1}^{W_n} \ln r_{\omega jn} \right)
\]

on the set of firm dummies $\{b_{\omega j}\}_{\omega=1}^{W_n}$.
Table 1: Descriptive statistics, 2003

<table>
<thead>
<tr>
<th></th>
<th>CN8 product level</th>
<th>HS6 product level</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Full sample</td>
<td>Restricted sample</td>
</tr>
<tr>
<td>Number of</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Products</td>
<td>5,339</td>
<td>1,075</td>
</tr>
<tr>
<td>Firms</td>
<td>4,304</td>
<td>3,898</td>
</tr>
<tr>
<td>Destinations</td>
<td>223</td>
<td>222</td>
</tr>
<tr>
<td>Observations</td>
<td>155,426</td>
<td>116,924</td>
</tr>
<tr>
<td>Trade volume (billion DKK)</td>
<td>182</td>
<td>135</td>
</tr>
<tr>
<td>Destinations per firm-product</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Mean</td>
<td>3.4</td>
<td>4.1</td>
</tr>
<tr>
<td>Median</td>
<td>1.0</td>
<td>1.0</td>
</tr>
<tr>
<td>Min</td>
<td>1.0</td>
<td>1.0</td>
</tr>
<tr>
<td>Max</td>
<td>138.0</td>
<td>138.0</td>
</tr>
<tr>
<td>Firms per product-destination</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Mean</td>
<td>2.2</td>
<td>2.9</td>
</tr>
<tr>
<td>Median</td>
<td>1.0</td>
<td>2.0</td>
</tr>
<tr>
<td>Min</td>
<td>1.0</td>
<td>1.0</td>
</tr>
<tr>
<td>Max</td>
<td>373.0</td>
<td>373.0</td>
</tr>
<tr>
<td>Truncated (missing) obs., per product</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Mean</td>
<td>0.53</td>
<td>0.84</td>
</tr>
<tr>
<td>Median</td>
<td>0.66</td>
<td>0.88</td>
</tr>
<tr>
<td>Standard Deviation</td>
<td>0.34</td>
<td>0.10</td>
</tr>
<tr>
<td>10 percentile</td>
<td>0.00</td>
<td>0.72</td>
</tr>
<tr>
<td>90 percentile</td>
<td>0.91</td>
<td>0.94</td>
</tr>
</tbody>
</table>

Note: The restricted sample considers only products with positive degrees of freedom such that the number of firm-destination trade flows exceeds the number of firms plus the number of destinations.
Table 2: Summary statistics for $R^2$ of 'simple' OLS regressions

<table>
<thead>
<tr>
<th></th>
<th>Mean</th>
<th>S.d.</th>
<th>P25</th>
<th>P50</th>
<th>P75</th>
<th>N</th>
</tr>
</thead>
<tbody>
<tr>
<td>With zeros:</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Firm fixed effects</td>
<td>.186</td>
<td>.098</td>
<td>.113</td>
<td>.166</td>
<td>.238</td>
<td>836</td>
</tr>
<tr>
<td>Firm-destination fixed effects</td>
<td>.890</td>
<td>.108</td>
<td>.837</td>
<td>.924</td>
<td>.973</td>
<td>836</td>
</tr>
<tr>
<td>Difference</td>
<td>.704</td>
<td>.123</td>
<td>.627</td>
<td>.721</td>
<td>.795</td>
<td>836</td>
</tr>
<tr>
<td>Without zeros:</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Firm fixed effects</td>
<td>.589</td>
<td>.201</td>
<td>.431</td>
<td>.579</td>
<td>.738</td>
<td>880</td>
</tr>
<tr>
<td>Firm-destination fixed effects</td>
<td>.934</td>
<td>.075</td>
<td>.903</td>
<td>.960</td>
<td>.989</td>
<td>880</td>
</tr>
<tr>
<td>Difference</td>
<td>.345</td>
<td>.184</td>
<td>.200</td>
<td>.351</td>
<td>.474</td>
<td>880</td>
</tr>
</tbody>
</table>

Note: The rows labelled "Difference" record summary statistics for the product-level differences between $R^2$ from the firm fixed effects and firm-destination fixed effects regressions.

Table 3: Trade residuals and firm and destination characteristics

<table>
<thead>
<tr>
<th></th>
<th>Coef.</th>
<th>Std. Err.</th>
</tr>
</thead>
<tbody>
<tr>
<td>Log (size)</td>
<td>-0.4531***</td>
<td>0.0085</td>
</tr>
<tr>
<td>Log (value added per worker)</td>
<td>-0.2285***</td>
<td>0.0292</td>
</tr>
<tr>
<td>Share of high skilled workers</td>
<td>-0.4170***</td>
<td>0.0899</td>
</tr>
<tr>
<td>Log (GDP)</td>
<td>-0.1607***</td>
<td>0.0156</td>
</tr>
<tr>
<td>Log (distance)</td>
<td>0.0980***</td>
<td>0.0107</td>
</tr>
<tr>
<td>Log (population)</td>
<td>-0.0264***</td>
<td>0.0068</td>
</tr>
<tr>
<td>Constant</td>
<td>6.3470***</td>
<td>0.4522</td>
</tr>
</tbody>
</table>

Note: Table 3 presents the results from a regression of the trade residual on firm and destination characteristics. The regression also includes industry fixed effects. The data used is the 2003 restricted sample from Table 1 collapsed to the firm-destination level. *** p<0.01.

Table 4: Summary statistics for $Q^2$ of MCMC and OLS regressions

<table>
<thead>
<tr>
<th></th>
<th>Mean</th>
<th>S.d.</th>
<th>P25</th>
<th>P50</th>
<th>P75</th>
<th>N</th>
</tr>
</thead>
<tbody>
<tr>
<td>MCMC</td>
<td>.486</td>
<td>.137</td>
<td>.395</td>
<td>.454</td>
<td>.545</td>
<td>880</td>
</tr>
<tr>
<td>OLS</td>
<td>.345</td>
<td>.195</td>
<td>.196</td>
<td>.345</td>
<td>.474</td>
<td>508</td>
</tr>
<tr>
<td>MCMC, weighted</td>
<td>.533</td>
<td>.140</td>
<td>.421</td>
<td>.507</td>
<td>.633</td>
<td>146,624</td>
</tr>
<tr>
<td>OLS, weighted</td>
<td>.350</td>
<td>.172</td>
<td>.214</td>
<td>.362</td>
<td>.477</td>
<td>123,846</td>
</tr>
</tbody>
</table>

Note: The last two rows report results weighted by product-level export value.
Table 5: Summary Statistics for $Q^2$, MCMC Robustness Checks

<table>
<thead>
<tr>
<th></th>
<th>Mean</th>
<th>S.d.</th>
<th>P25</th>
<th>P50</th>
<th>P75</th>
<th>N</th>
</tr>
</thead>
<tbody>
<tr>
<td>Products with at least 100 obs.</td>
<td>.489</td>
<td>.148</td>
<td>.385</td>
<td>.453</td>
<td>.548</td>
<td>342</td>
</tr>
<tr>
<td>Products with at least 300 obs.</td>
<td>.491</td>
<td>.164</td>
<td>.382</td>
<td>.446</td>
<td>.547</td>
<td>80</td>
</tr>
<tr>
<td>CN8</td>
<td>.499</td>
<td>.143</td>
<td>.402</td>
<td>.471</td>
<td>.575</td>
<td>1,075</td>
</tr>
<tr>
<td>HS2</td>
<td>.483</td>
<td>.179</td>
<td>.351</td>
<td>.428</td>
<td>.542</td>
<td>59</td>
</tr>
<tr>
<td>Continuing products</td>
<td>.524</td>
<td>.136</td>
<td>.431</td>
<td>.515</td>
<td>.611</td>
<td>698</td>
</tr>
<tr>
<td>Core products</td>
<td>.555</td>
<td>.159</td>
<td>.483</td>
<td>.572</td>
<td>.670</td>
<td>297</td>
</tr>
<tr>
<td>Core products extended</td>
<td>.504</td>
<td>.141</td>
<td>.409</td>
<td>.473</td>
<td>.568</td>
<td>297</td>
</tr>
</tbody>
</table>

Note: Continuing products are products exported in 2003 that were also exported in 2002. Core products are the top export of each firm only. The last row records the results for the 297 core products but considering all products of the firms.

Figure 1: Sales relative to other Danish firms in Sweden and Germany for Danish Exporters of plastic boxes (left panel) and building bricks (right panel). The axes show relative revenues calculated as a firm’s log revenue minus the average log revenue of all firms in the destination. Statistics for the lines with fitted values: Left panel: slope = 0.84, std.err. = 0.08, $R^2 = 0.54$. Right panel: slope = −0.22, std.err. = 0.12, $R^2 = 0.08$. 

41
Figure 2: Simulated relative sales in two destinations for varying values of $R^2$. 

$R^2 = 0.999$  

$R^2 = 0.5$  

$R^2 = 0.12$
Figure 3: Firm-destination specific exports and domestic sales
Figure 4: The distribution of log domestic sales for Danish manufacturing firms, 2003.
Figure 5: Simulation results
Figure 6: Distributions of $Q^2$ across products
Figure 7: Correlations between $Q^2$s in 2001 and 2003