Preliminary draft prepared for the Comparative Analysis
of Enterprise (Micro) Data Conference, Chicago, IL
September 18-19, 2006

Plant Size and Plant Function

Thomas J. Holmes* and John J. Stevens†

September 8, 2006

Abstract

Relatively little is known about plant-level shipping activity. We use detailed data on plant-level shipments to document two new facts. First, not only are large plants more likely to be exporters, they are also more likely to ship far domestically. Second, small plants tend to have higher unit values. We provide two possible explanations for these results. First, high productivity plants may be more likely to invest in distribution channels that facilitate exporting and longer shipping distances; however, this explanation has ambiguous implications for prices. Alternatively, small plants, even within narrowly defined industries, may be performing more customized or retail-like activities that make them fundamentally different from large plants within the same industry. This difference in function can explain both facts.

*University of Minnesota, Federal Reserve Bank of Minneapolis, and NBER. Address: Dept. of Economics, 1035 Heller Hall, University of Minnesota, Minneapolis, MN 55455, holmes@econ.umn.edu

†Board of Governors of the Federal Reserve System. Address: Mail Stop 82, 20th & C Streets N.W., Washington, D.C. 20551, john.j.stevens@frb.gov

The views expressed herein are those of the authors and not necessarily those of the Federal Reserve Bank of Minneapolis, the Federal Reserve Board, or the Federal Reserve System. This research was conducted while the authors were Special Sworn Employees of the U.S. Census Bureau. Results have been reviewed to prevent disclosure of confidential information. The authors thank Joshua Louria of the Federal Reserve Board for excellent research assistance and Shawn Klimek of the Census Bureau's Center for Economic Studies for assistance with the microdata files.

1 Introduction

In any industry there is tremendous variation in plant sizes. A substantial quantity of recent research has linked plant size to things that matter, like productivity and export behavior. Larger plants are more productive and more likely to export. These findings have led to an influential theoretical literature including Melitz (2003), Eaton and Kortum (2002), Bernard, Eaton, Jensen, and Kortum (2003), and Eaton, Kortum, and Kramarz (2005).

This paper uses Census Bureau micro data to relate plant size to two new plant charactersitics. The data are from a detailed survey of individual shipments leaving a plant. The information collected includes the destination, value, and weight of each shipment. The first chacteristic we examine is a plant's tendency to ship to distant locations within the United States as opposed to nearby locations. The second characteristic is the unit price of a plant's shipments (defined as price per pound).

We report two findings. First, holding fixed a plant's narrowly defined industry, larger plants ship longer distances. This result is analogous to the previously reported result for exports. Our finding suggests that whatever the forces at work determining why larger plants export, the same forces may be at work in determining where they ship within the United States. Second, holding fixed a plant's narrowly defined industry and the distance the plant tends to ship, larger plants have a lower unit price.

We provide two interpretations of our findings. We begin with a simple extension of the Melitz (2003) model to include transportation costs for domestic trade as well as export trade. A key idea in Melitz is that high productivity plants will tend to invest in distribution channels that permit them to export, while low productivity plants do not find it profitable to make such an investment. This difference explains why larger plants export. Our paper suggests that this property may not be unique to export markets. Investment in distribution channels may be just as important for selling in distant states as it is for shipping to distant countries. If so, then larger plants will ship further within the country.

In addition to presenting this extension to the Melitz model, we provide an alternative explanation. We argue that even in narrowly defined industries, there may be variation in plant function. Small plants tend to do different things than large plants; in particular, they specialize in more custom work or retail-like activity. These types of activities are often more efficiently undertaken in small plants. A mention of NAICS industry 336111, "Automobile manufacturing," makes this point clear. There are 197 establishments in the 1997 Economic Census. Of these, 26 plants have 1,000 or more employees, a size one might expect to find in a typical automobile plant. But there are also 92 establishments in this industry with one to four employees. These tiny "automobile plants" make race cars, stretch limos, conversion vans, etc., an obviously different function than that performed by large automobile plants. Our idea is that these retail-like plants tend to ship locally and given they sell only to a local market, they tend to be small.

Our explanation is in one way similar and in another way different from the Melitz explanation. The key similarity is that the small plant has higher transportation costs for shipping to distant locations than the large plant. The key difference lies in the reason why transportation costs are higher for small plants. In the Melitz model, it is because the small, inefficient plant chooses not to invest in building a distribution channel. In our explanation, the difference is exogenous because the small and big plant do different things. Our explanation and the Melitz theory have different policy implications. If tariffs are reduced, small plants shut down in the Melitz model. In our explanation, small plants may survive because they are performing activities for which trade is difficult.

We expect that the endogenous investment explanation of Melitz and our alterative explanation likely both play some role in explaining plant-level shipping behavior. The question is, what is the relative importance of each explanation? Looking at the variations in unit prices is a first step in this direction. While a specific version of the Melitz model is consistent with larger plants having lower prices, this need not be true. Generally, if larger plants have lower costs then they pass these lower costs along to consumers in the form of lower prices. But larger plants may also produce higher quality products, in which case this quality differential can be passed along to consumers in the form of higher prices. What is key for the theory is that larger plants are more productive—no general statements can be made about price. In our explanation we think it is natural to expect that the small plants performing custom activities will be charging higher unit prices than the mass production plants. Our price results match this implication. We recognize that this result is not a sharp way to distinguish the Melitz explanation from our explanation, because the Melitz model implies an ambiguous relationship between plant size and price, and is therefore not inconsistent with the result. Rather, we view this analysis as the first step in a longer project.

Some additional literature needs to be cited in addition to that mentioned above. Our data come from the Census Bureau's Commodity Flow Survey (CFS). Hillberry and Hummels (2005) are the first economists to use the CFS micro data. Their research emphasizes the empirical importance of local shipments. They find, for example, that shipments are much more likely to go someplace in the same zip code as the plant as opposed to an arbitrary zip code located quite a distance away. An important way in which our research differs from theirs is that our focus is on the characteristics of plants that do ship a great distance. Another way that our research differs from theirs is that we use the unit price information available in the survey.

Roberts and Supina (1996) show that for a few selected industries in which unit values can be obtained using the product data from the Census of Manufactures, larger plants have lower prices, just like we report. An important benefit of using the CFS microdata is that we are able to conduct an analysis of the entire manufacturing sector, not just a few selected industries. Moreover, in our analysis of price, we are able to control for distance shipped.

2 Theory

We present a simple model to articulate the two candidate explanations for the patterns in the data we report.

We conduct a partial equilibrium analysis. There are three locations, 0, 1, and 2. We look at plants at location 0 and determine the locations to which they ship. Think of location 0 as some interior location like Minneapolis. Location 1 is a domestic location that is a port, e.g. New York. And location 2 is a foreign country. Shipments destined for location 0 are local shipments, those bound for location 1 are distant domestic shipments, and those headed to location 2 are exports.

A plant at location 0 has marginal cost c to produce one unit and deliver it at location 0.

There is an iceberg transportation cost to ship to location 1. The level $\tau_1(x_1)$ of the iceberg cost to ship to 1 depends upon an investment x_1 that the firm takes to lower the cost, with $\tau_1(x_1)$ weakly decreasing in x_1. Assume that to ship to location 2, the firm must go through location 1 and that an analogous investment x_2 can be made to lower the cost $\tau_2(x_2)$ of shipping from 1 to 2. In summary, to deliver one unit to location 0 costs c, to deliver one unit to 1 costs $c\,[1+\tau_1(x_1)]$, to deliver one unit to location 2 costs $c\,[1+\tau_1(x_1)]\,[1+\tau_2(x_2)]$.

Allowing for investment captures a key element of Melitz. Melitz considers an extreme case where $\tau=\infty$ if there is no investment and τ drops to 0 if a particular fixed investment x^* is made. This is a special case of a more general structure where $\tau_1(x_1)$ and $\tau_2(x_2)$ are (weakly) decreasing functions x_1 and x_2. Arkolakis (2006) considers an analogous structure.

Firms vary in the quality, γ, of their product. A firm of quality γ faces a demand curve $D_i(p,\gamma)$ at location i. Assume all locations have the same constant elasticity demand $\eta>1$ and that γ enters in a way proportionate to price. Then the quantity demanded at location i is given by

$$D_i(p_i,\gamma)=\lambda_i\left(\frac{p_i}{\gamma}\right)^{-\eta}.$$

The paramer λ_i scales the level of demand. The problem of the firm is to pick investments x_1 and x_2 and prices p_0, p_1, and p_2 to maximize

$$\begin{aligned}&[p_0-c]\,D_0(p_0,\gamma)+[p_1-c\,[1+\tau_1(x_1)]]\,D_1(p_1,\gamma)\\&+[p_2-c\,[1+\tau_1(x_1)]\,[1+\tau_2(x_2)]]\,D_2(p_2,\gamma)-x_1-x_2.\end{aligned}$$

2.1 Melitz

To illustrate the Melitz explanation, suppose plants vary in two dimensions, marginal cost c and quality γ. Following the usual arguments, for fixed investments x_1 and x_2, the profit maximizing price is a constant markup over margainal cost. For a plant at location 0, the price is

$$p_0=c\frac{\eta}{\eta-1} \tag{1}$$

and the variable profit is

$$\pi_0=c^{-(\eta-1)}\gamma^{\eta}\lambda_0\eta^{-\eta}\,(\eta-1)^{(\eta-1)}. \tag{2}$$

The price ratios are

$$\frac{p_1}{p_0} = [1+\tau_1(x_1)] \tag{3}$$
$$\frac{p_2}{p_0} = [1+\tau_1(x_1)]\,[1+\tau_2(x_2)]\,.$$

The profit maximizing ratio of sales between location 1and location 0 sales will be

$$\frac{Q_1}{Q_0} = \frac{\lambda_1}{\lambda_0}\frac{1}{[1+\tau_1(x_1)]^{\eta}}. \tag{4}$$

The ratio of exports to domestic shipments will be

$$\frac{Q_2}{Q_0+Q_1} = \frac{\lambda_2}{\lambda_0\,[[1+\tau_1(x_1)]\,[1+\tau_2(x_2)]]^{\eta} + \lambda_1\,[1+\tau_2(x_2)]^{\eta}}. \tag{5}$$

Note that for fixed x_1 and x_2, the cost and quality parameters c and γ drop out of the shipment ratio. However, c and γ do matter because they affect x_1 and x_2. Let $x_1^*(c,\gamma)$ and $x_2^*(c,\gamma)$ be the optimal investments given marginal cost draw c. It turns out that the optimal investment levels depend on an index θ of c and γ,

$$\theta = \frac{\gamma^{\eta}}{c^{\eta-1}}.$$

Both investment functions are weakly increasing in θ. Therefore, everying else fixed, firms with higher quality (i.e., higher γ) or lower marginal cost (i.e., lower c) will invest more to reduce their transportation costs. Now with greater investments in x_1 and x_2, the ratio of distant domestic shipments to local shipments will increase (4) and the ratio of exports to total domestic shipments will increase (5). It is also immediate that higher θ firms will have higher total revenue. In addition, because variable costs are a constant proportion of total revenue, higher θ firms will have higher costs.

We conclude that firms with higher θ will have the following properties:

1. Higher revenues and higher total costs (variable costs plus x_1 and x_2).

2. A higher ratio of exports to domestic shipments (equation 5).

3. A higher ratio of distant domestic shipments to local shipments (equation 4).

If all heterogeneity is driven by differences in θ, then larger plants (as measured by total revenues or by total costs) will have more exports. This is exactly the finding in Bernard and Jensen (1995). These plants will also have a larger share of distant domestic shipments. This result is exactly what we find in our empirical work discussed later.

Next consider price. If plants only vary in c, with γ fixed at 1, then the low cost plants will have low prices. Hence larger plants will have lower prices. But we get something different if the low cost plants also have lower quality. Suppose, for example, that quality differences are proportionate to cost differences, $c/\gamma = 1$. Then

$$\theta = \frac{\gamma^\eta}{c^{\eta-1}} = \frac{\gamma^\eta}{c^\eta}c = c.$$

Then the higher cost, higher quality plants will be the ones investing in distribution channels. These plants will be the largest in terms of revenue and costs, and they will tend to export and ship far domestically. In addition, these plant will also tend to have higher prices.

2.2 Plant Size and Plant Function

We keep the same basic structure as above, but now imagine that the heterogeneity arises for different reasons. We also eliminate investment in reducing transportation cost.

There are two types of plants: mass-production plants and custom/retail-like plants. Index these types by m and r. Both types of plants are classified in the same narrowly defined industry, but they perform different functions. For simplicity, assume the two types of plants face the same demand curves.

The plants differ in their marginal costs of production. Assume $c^m < c^r$. The retail-like plants have to do extra steps for their custom work, which raises the cost. Also, with small production batches, the plants have little incentive to pay the fixed costs necessary to routinize particular tasks and lower the marginal cost.

We also assume that the plants differ in their transportation costs. Specifically, we assume $\tau_1^m < \tau_1^r$ and $\tau_2^m < \tau_2^r$. The retail-like task may involve face-to-face communication and the costs of moving people are high. In addition, the extra steps of production necessary to customize production may make the product too bulky to ship far.

It is staightforward to show the manufacturing plants compared to custom/retail-like plants will have

1. Higher revenues and higher total costs.
2. A higher ratio of exports to domestic shipments.
3. A higher ratio of distant domestic shipments to local shipments
4. *Lower prices.*

Properties 1, 2, and 3 are the same as in the Melitz case discussed above. Property 4 is new.

3 Data

The Census Bureau's Census of Manufactures (CMF) collects establishment-level data for each six-digit NAICS manufacturing industry. For a subset of the establishments, detailed product-level data corresponding to 10-digit NAICS product categories are also collected. Despite this fine level of product detail, very little is known about the actual shipments. Is the product being shipped 1,195 miles from Minneapolis to New York or 9.6 miles from Minneapolis to St. Paul. Are the boxes being sent one-by-one on a FedEx truck, or are dozens of boxes being packed into shipping containers and transported by rail and then by truck? At what price are the products sold? To what extent do the answers to these questions depend upon the industry and the size of the establishment?

The Census Bureau's Commodity Flow Survey (CFS), which is conducted in cooperation with the United States Department of Transportation, sheds some light on the answers to these questions. This survey collects information on individual plant-level shipments for a sample of establishments. Specifically, each plant is requested to report details on a one-week sample of its shipments in each quarter of the year. The information includes the origin and domestic destination (from which shipping distances are calculated), weight, value, modes of transport, and a flag for whether or not the shipment was exported. For exports, the domestic destination is the U.S. water port, airport, or border crossing from which the shipment left the country; the destination city, country, and mode of export are also requested. In addition, respondents note whether the shipment was containerized. The commodity codes attached to a shipment are Standard Classification of Transported Goods (SCTG) codes, which resemble SIC or NAICS codes. However, we do not use these codes as we link the CFS with the CMF at the plant level and are therefore able to use the detailed NAICS code that was assigned to the establishment.[1] The current analysis focuses on data from 1997.[2]

The 1997 CFS contains about 64,000 establishments in manufacturing, mining, wholesale trade, and selected retail industries. We restrict attention to the manufacturing establishments that we can match with the CMF. Our base sample has 30,148 establishments. The aggregate properties of the CFS match the CMF reasonably well. As shown in table 1, our CFS manufacturing sample does not quite cover total manufacturing shipments and employment as measured by the CMF. For example, our CFS sample (when weighted with the sample weights in the survey) has about $3.3 trillion in shipments versus $3.8 trillion in the CMF. The difference for employment is similar. However, establishments in the CFS sample tend to be larger, with an average employment of 95 versus 46 for the CMF overall. Likewise, CFS plants ship roughly twice as much on average. Table 2 compares the overall CMF with the subset of CMF plants that we are able to match to the CFS. Unweighted averages further demonstrate that the CFS plants are larger, more likely to export (about three times more likely), and tend to export a larger share of their production. We adjust our sample weights to align the shipments of each plant in the CFS with its reported shipments in the CMF. This aligning is why the shipments per worker is identical in table 3 when comparing the CFS with the weighted CMF subset. Importantly, shipments per worker (both overall and by plant size) are similar to the overall CMF because of this re-weighting. In contrast, the unweighted subset has considerably higher shipments per worker. The export shares for

[1] Because we collapse shipments down to a single establishment-level observation, using the establishment's NAICS code makes more sense than using its SCTG codes.

[2] The Commodity Flow Survey was conducted five times between 1963 and 1983, and, more recently, in 1993, 1997, and 2002. The sample size for 2002 was one-half the size of the 1997 sample, therefore we felt more comfortable using the 1997 data for this initial anyalsis. As a robustness check, we plan to replicate our analysis for 1993 and 2002 in the near future.

the CFS are nonetheless a bit higher than in any of the CMF aggregations. However, the qualitative relationship between the export share and plant size is similar.

4 Results

As noted earlier, we have two key findings. Large plants tend to ship further, and large plants tend to have lower prices. The first result is shared by both the Melitz model and our explanation based on plant function. The later is true under our explanation but need not hold in the Melitz model.

4.1 Large Plants Ship Further

Plants ship remarkably long distances, and this varies in striking ways with plant size. Table 4 provides summary statistics that characterize the shipping behavior of manufacturing establishments. Among all establishments, 20 percent ship less than 50 miles on average and nearly twice that amount ship more than 500 miles on average. However, these shares vary dramatically with the employment size of an establishment. In establishments with 0 to 19 employees, 46 percent tend to ship within 50 miles and only 22 percent tend to ship 500 miles of more. In constrast, only 10 percent of establishments with at least 500 employees have an average shipping distance of less than 50 miles while 47 percent average at least 500 miles.

As noted at the outset of this paper, previous researchers have documented the fact that large establishments are more likely to export and tend to export a larger share of their production. In our data, the average export share for large establishments is 11 percent while the average for small establishments is only 3 percent. These statistics may overstate the difference in export shares if small establishments are more likely to sell to an establishment specializing in the aggregation of small shipments for export. Nonetheless, the export shares paint an interesting picture of the shipping patterns of large and small establishments when viewed in conjuction with the domestic shipping distances. Namely, small establishments are not only less likely to ship abroad, they are less likely to ship from Minneapolis to New York.

Because the Commodity Flow Survey data contains information on both the origin and destination of shipments, we can further characterize shipping patterns by examining the zip codes to which an establishment ships. On average, small establishments ship 12 percent of their production within the same zip code while large establishment only ship 2 percent within the same zip code. However, shipping patterns do exhibit a fair degree of regularity. Among all establishments, the zip code receiving the largest share of an establishment's shipments accounts for nearly 30 percent of its production. This share varies from an average of 42 percent for small establishments to 25 percent for the largest establishments. Table 5 provides counts of establishments based on the share of shipments traveling less than 100 miles. Interestingly, less than 1 out of every 4 plants ships 50 percent or more of its production within 100 miles (which can be seen by summing the last six values in the "all" column), and the modal plant ships less than 10 percent within 100 miles.

The distances that plants tend to ship may seem surprising given Hillberry and Hummels (2005) finding that most shipments are highly localized. They argue that observed shipping patterns, in the aggregate, primarily reflect the supply of intermediate products to downstream stages of production, as the co-location of stages of production results in a high number of short shipments. It is important to remember that although we are using the same microdata, we are looking at the data from a different angle. Our unit of analysis is the establishment rather than a shipment. Because a large share of shipments are highly localized, the focus on individual shipments obscures some of the establishment characteristics. Thus, our results are complementary to those of Hillberry and Hummels in that we characterize establishments while they characterize aggregate shipments.

In our work, by treating establishments as the unit of analysis, we effectively downweight the influence of intermediate products. For example, suppose there are two locations located 1000 miles from each other and that there are also two establishments, one of which supplies the other. Because of transportation costs, the two plants have an incentive to co-locate. They each have the same number of shipments, but the upstream establishment ships everything locally while the downstream establishment ships half of its products locally and half to the distant location (i.e., final demand is uniformly distributed). A plot of the count of shipments by distance will show 3/4 being local (shipped 0 miles) and 1/4 being distant (shipped 1000 miles). However, when aggregated to the establishment level, a similar plot will show 1/2 being local (shipped 0 miles) and 1/2 being distant (shipped 500 miles $= 0.5 \times 0$ miles$+0.5 \times 1000$ miles). Using establishment-level distances reduces the maximum distance, but it also reduces the influence of intermediate goods.

We were concerned that our results, while striking and appealing theoretically, may reflect the mix of industries. Industries that tend to ship far may have establishments that look different from industries that tend to be local. Table 6 looks at the share shipped less than 100 miles, hereafter referred to as the "local share," as well as the share that is exported, on an industry-by-size basis. Most industries exhibit a fairly steep decline in the local share as we move from small to large establishments, though the textiles and apparel industries and the computer and electronic products industry are notable exceptions. Thus, it does not appear that the industry-heterogeneity was the primary factor driving the relationship between the more highly aggregated shares.

The local shares also vary widely by industry, holding size fixed. Among establishments with under 20 employees, the local share ranges from 15 percent for the computer and electronic products industry to 81 percent for petroleum and coal and nonmetallic mineral products industries. Among large establishments, the local share ranges from 5 percent (miscellaneous) to 65 percent (petroleum and coal products), though the latter is something of an outlier as the next highest share is 30 percent (apparel). Similarly, export shares for small establishments range from 0 percent to 10 percent, and for large establishments they range from 1 percent to 24 percent; in both cases the lowest share is for printing and related support activities while the largest share is for computer and electronic product.

Despite the local nature of a large share of shipments, it is still striking how far shipments travel. Conditional on shipping more than 100 miles on average, the share that is shipped at least 500 miles or exported rises markedly. As shown in table 7, among establishments with at least 95 percent of their shipments going further than 100 miles, the average share going further than 500 miles is about 70 percent. Furthermore, this share shows little relationship to plant size. However, as the share shipped 100 miles or more decreases, the share shipped 500 miles or more also decreases, and the difference between large and small plants reemerges, though to a lesser degree than was the case in table 4. In table 4, the share shipped 500 miles or more (or exported) was 25 percent for small establishments (22 percent more than 500 miles and 3 percent export) and 58 percent for large establishments (47 percent more than 500 miles and 11 percent exported). Conditioning on the share shipped 100 or more miles greatly reduces the difference between small and large establishments, particularly when this share is high.

Although table 6 suggests that plant size matters for shipping patterns even within 3-digit industries, it may still be the case that there is industry-level heterogeneity that is driving our plant size results. To check this possibility we regressed each shipping share from table 4 on an intercept and plant size dummies (0-19 employees is the excluded category), the results of which are summarized in table 9. The signs are as expected, monotonic in plant size, and statistically significant. For each dependent variable we also ran the regression including 6-digit industry controls. These controls reduced the magnitude of the coefficients on the size variables, which suggests that some of what we are seeing is indeed industry-level heterogeneity. Nonetheless, the coefficients on the size dummies remain large and continue to show the same monotonic patterns as the regressions without the industry controls.

4.2 Large Plants Have Lower Prices

The differences in shipping distances suggest that small and large establishments within narrowly defined industries may be performing different activities. Another way to infer that plants within a narrowly defined industry are performing different activities is to look at the prices of what they are shipping. Plant-level prices are typically hard to come by, but the microdata in the Commodity Flow Survey afford us the unique opportunity to construct a proxy for these plant-level prices by taking the ratio of each shipment's value to its weight. We think of this "price-per-pound" measure as the unit value of a shipment. These prices are noisy indicators, both within and across plants. Therefore we use robust measures of the central tendency of these prices. Specfically, for each plant, we take the median of its unit values and call the resulting number the plant's price. We then calculate a trimmed mean of the plant-level prices, measured as the average of the plant-level prices in a band centered about the median plant.

Before discussing the results, it is worthwhile to note some of the benefits of these prices as well as some caveats. First, an important advantage to using these prices is that they are easily compared across industries. Other work has had to focus on individual industries where the products are sufficiently homogeneous that physical measures of output can be used thereby avoiding prices altogether (e.g., Syverson, 2004), or has come up with a means to strip out differences in measurement units (e.g. Roberts and Supina, 1996). Second, because we effectively obtain multiple reported prices for each plant, we are able to construct proxies for plant-level prices that are relatively insensitive to outliers. Third, these prices seem reasonable. For example, we when looked at narrowly defined industries with relatively homogenous products, the prices per pound generally matched what we would imagine the price per pound would be if we collected price data in the marketplace. Nonetheless, these prices are not perfect. For plants producing a wide array of heterogeneous products that vary significantly in value, an average price for the plant will necessarily be less informative than in an industry where plants produce fairly homogeneous goods. That said, we think these prices are a promising tool for thinking about unobserved heterogeneity across plants within an industry.

Turning to our results, in the top half of table 10 we report the trimmed mean price by plant size and the share shipped less than 100 miles. Small plants tend to have uniformly higher prices, although the prices drop off rapidly as plant size increases. For most of the distance categories, prices drop sharply as we move from plants with less than 20 employees to plants with 20 to 99 employees. They fall a bit further for plants with 100 to 499 employees. Interestingly, the prices often edge back up for the largest plants, but to a level typically well below that of the smallest plants. (However, regression results given below suggest this may reflect, in part, industry heterogeneity.) The difference between the large and small plants is greater for those plants that tend to ship more of their output longer distances (i.e., looking down columns in table 10). Taken together, these results suggest that small plants are performing a different function than large plants. When the plants do not tend to ship far, they are more likely to be competing directly, and hence exhibit more similarity in prices. However, if both plants ship far, then the small plant must be doing something sufficiently different in order to explain the significantly higher price.

The lower half of table 10 repeats the analysis but restricts the sample to plants for whom the Commodity Flow Survey collected data on at least 20 shipments. This robustness check has little effect on the results, with the exception of the cell corresponding to the smallest plants in the shortest distance category; this price falls a good bit when this restriction is imposed.

Table 11 looks at the prices for 3-digit industries, and again the bottom half of the table includes the restriction to plants with 20 or more shipments. The highest prices tend to be for computer and electronic product, apparel, leather product, machinery, and electrical equipment, appliance, and components, which are all valuable relative to their weights. In contrast, the lowest prices are in commodities (or near commodities) such as nonmetallic mineral product, petroleum and coal product, wood product, food, and beverage and tobacco product. The computer prices rise with establishment size, while the other industries decrease (at least weakly).

Similar to the exercise we performed with the shipping shares, we regressed our price measures on dummy variables for plant size and shipping distance, both with and without 6-digit industry controls. We also repeated the analysis with the data restricting to plants with 20 or more shipments. These results are summarized in table 12. Price rises uniformly with distance in each regression. However, most of the increase reflects industry-level heterogeneity. Including 6-digit industry controls dramatically reduces the effect of distance, as the coefficients on the distance dummies are reduced by more than 80 percent. In contrast, the effect of establishment size becomes more pronounced once detailed industry controls are included. In addition, without industry controls, price appears to be U-shaped in employment size, but with industry controls prices decrease monotonically.

As noted earlier, the fact that large plants have lower prices emerges unambiguously from our explanation of small plants tending to be more retail-like and producing customized goods. While this price result does not overturn the Melitz model, it does raise the distinct possibility that there is more going on than just productivity differences and fixed costs of exporting.

5 Conclusions

This project is still in the early stages. One direction that we are still investigating is the means by which manufacturers deliver their goods to the market. It may be possible to more sharply delineate between the Melitz model explanation and the plant function explanation of the patterns documented in this paper by examining shipping choices. For example, the decision to containerize or the decision to use parcel delivery rather than rail may be important. In addition, the size of shipments may shed light on the degree of customization.

References

[1] Arkolakis, Costas. "Endogenous Market Access Costs and the New Consumers Margin in International Trade." Working paper, University of Minnesota, June 2006.

[2] Bernard Andrew B.; Eaton, Jonathan; Jensen, J. Bradford and Samuel Kortum. "Plants and Productivity in International Trade." *American Economic Review*, 2003, 93(4), pp. 1268-90.

[3] Bernard, Andrew B. and J. Bradford Jensen. "Exporters, Jobs, and Wages in U.S. Manufacturing, 1976-1987." *Brookings Papers on Economic Activity: Microeconomics*, 1995, pp. 67-119.

[4] Dunne, Timothy; Roberts, Mark J. and Larry Samuelson. "The Growth and Failure of U.S. Manufacturing Plants." *Quarterly Journal of Economics*, November 1989, 104(4), pp. 671-98.

[5] Eaton, Jonathan and Samuel Kortum. "Technology, Geography, and Trade." *Econometrica*, 2002, 70(5), pp. 1741-79.

[6] Eaton, Jonathan; Kortum, Samuel and Francis Kramarz. "An Anatomy of International Trade: Evidence from French Firms." Working paper, 2005.

[7] Hillberry, Russell and David Hummels. "Intranational Home Bias: Some Explanations." *Review of Economics and Statistics*, November 2003, 85(4), pp. 1089-92.

[8] Hillbery, Russell and Hummels, David, "Trade Responses to Geographic Frictions: A Decomposition Using MicroData," working paper, January 2005.

[9] Holmes, Thomas J. and John J. Stevens. "Geographic Concentration and Establishment Scale." *Review of Economics and Statistics*, November 2002, 84(4), pp.682-90.

[10] Holmes, Thomas J. and John J. Stevens. "Geographic Concentration and Establishment Size: Analysis in an Alternative Economic Geography Model." *Journal of Economic Geography*, June 2004a, 4(3), pp. 227-50.

[11] Holmes, Thomas J. and John J. Stevens "Spatial Distribution of Economic Activities in North America," *Handbook on Urban and Regional Economics*, North Holland: (2004b)

[12] Hummels, David and Alexandre Skiba, "Shipping the Good Apples Out? An Empirical Confirmation of the Alchian-Allen Conjecture." *Journal of Political Economy* 112 (2004) 1384-1402.

[13] Melitz, Marc. "The Impact of Trade on Intra-Industry Reallocations and Aggregate Industry Productivity." *Econometrica*, November 2003, 71(6), pp. 1695-1725.

[14] Piore, Michael J. and Charles F. Sabel. *The Second Industrial Divide*, New York: Basic Books, 1984.

[15] Roberts, Mark and Dylan Supina, "Output Price, Markups, and Producer Size," *European Economic Review Papers and Proceedings,* Vol. 40, Nos. 3/4 (April 1996), pp. 909-921.

[16] Scherer, F. M. (1980),*Industrial market structure and economic performance* 2nd Edition, Chicago : Rand McNally.

[17] Syverson, Chad, "Market Structure and Productivity: A Concrete Example," *Journal of Political Economy*, 112:6 (December 2004) 1181-1222.

Table 1. Comparing Aggregate Manufacturing Statistics from the Commodity Flow Survey and Census of Manufactures

	Employment	Shipments
Sum		
Commodity Flow Survey	14,350,751	3,324,513,064
Census of Manufactures	16,896,010	3,847,165,343
Average		
Commodity Flow Survey	95	21,951
Census of Manufactures	46	10,553

Note. Manufacturing is defined according to the North American Industry Classification System.

Table 2. Average Values of Selected Statistics from the Census of Manufactures

	All data	Unweighted subset
Total Employment	46	263
Total Value of Shipments	10,553	76,596
Shipments per Employee	133.13	271.70
Exports per Employee	4.16	17.02
Export Indicator	0.129	0.450
Export Share	0.021	0.062

Note. The subset includes those manufacturing establishments that were matched with the Commodity Flow Survey.

Table 3. Comparing the Commodity Flow Survey and the Census of Manufactures

		Total Employment		
	All	**0 to 99**	**100 to 499**	**over 500**
Shipments/Employment				
CFS, weighted	231.66	161.05	208.43	310.69
CMF, all	227.70	150.33	213.44	317.87
CMF, weighted subset	231.66	161.05	208.43	310.69
CMF, unweighted subset	291.43	272.19	253.83	324.90
Exports/Shipments				
CFS, weighted	0.12	0.06	0.10	0.15
CMF, all	0.08	0.04	0.07	0.11
CMF, weighted subset	0.09	0.05	0.07	0.12
CMF, unweighted subset	0.09	0.05	0.07	0.11

Note. The weighting is designed to match the level of each plant's shipments in the Commodity Flow Survey with its shipments in the Census of Manufactures. As a result, the "CFS, weighted" and "CMF, weighted subset" lines match for shipment per worker. Export shares are larger for the CFS than for the CMF subset, as the weights are not designed to mimic exports.

Table 4. Average Establishment-Level Shipping Shares

		Total Employment			
	All	**0 to 19**	**20 to 99**	**100 to 499**	**over 500**
Share < 50 miles	0.20	0.46	0.25	0.14	0.10
Share < 100 miles	0.27	0.54	0.34	0.21	0.14
Share < 500 miles	0.55	0.75	0.62	0.51	0.42
Share >= 500 miles	0.38	0.22	0.33	0.41	0.47
Export share	0.07	0.03	0.05	0.08	0.11
Same zip code share	0.04	0.12	0.05	0.02	0.02
Max zip code share	0.29	0.42	0.29	0.27	0.25

Table 5. Establishment Counts for Share Shipped Less than 100 Miles

		Number of Reported Shipments	
	All	**Less than 10**	**At least 10**
All	30148	4155	25993
Share shipped <100 miles			
.0	3859	1153	2706
.0 to .1	11123	419	10704
.1 to .2	3499	173	3326
.2 to .3	2178	126	2052
.3 to .4	1491	108	1383
.4 to .5	1248	108	1140
.5 to .6	1014	101	913
.6 to .7	929	105	824
.7 to .8	831	136	695
.8 to .9	882	157	725
.9 to 1.0	1305	377	928
1.0	1789	1192	597

Table 6. Shipping Shares by Industry and Establishment Size

NAICS		Mean share <100 miles				Mean export share			
		Total Employment				Total Employment			
		0 to 19	20 to 99	100 to 499	over 500	0 to 19	20 to 99	100 to 499	over 500
311	Food	0.53	0.38	0.28	0.22	0.06	0.04	0.04	0.03
312	Beverage & tobacco product	ND	0.41	0.43	0.26	ND	0.01	0.06	0.02
313	Textile mills	0.27	0.29	0.23	0.23	0.08	0.05	0.05	0.06
314	Textile product mills	0.37	0.23	0.20	0.20	0.04	0.05	0.03	0.04
315	Apparel	0.29	0.25	0.26	0.30	0.02	0.03	0.05	0.06
316	Leather & allied product	0.20	0.16	0.21	0.09	0.09	0.10	0.07	0.10
321	Wood product	0.56	0.42	0.22	0.09	0.03	0.04	0.04	0.04
322	Paper	0.47	0.44	0.33	0.09	0.02	0.03	0.04	0.09
323	Printing & related support activities	0.72	0.49	0.32	0.16	0.00	0.01	0.02	0.01
324	Petroleum & coal products	0.81	0.42	0.45	0.65	0.02	0.05	0.02	0.02
325	Chemical	0.38	0.24	0.18	0.19	0.07	0.08	0.13	0.14
326	Plastics & rubber products	0.39	0.29	0.19	0.12	0.03	0.05	0.07	0.10
327	Nonmetallic mineral product	0.81	0.59	0.30	0.10	0.01	0.03	0.06	0.10
331	Primary metal	ND	0.24	0.18	0.14	ND	0.05	0.07	0.08
332	Fabricated metal product	0.61	0.38	0.20	0.09	0.02	0.04	0.07	0.09
333	Machinery	0.42	0.17	0.10	0.07	0.05	0.12	0.16	0.19
334	Computer & electronic product	0.15	0.17	0.11	0.08	0.10	0.13	0.21	0.24
335	Elect. equip., appliance, & component	0.33	0.13	0.10	0.11	0.03	0.10	0.10	0.10
336	Transportation equipment	ND	0.23	0.17	0.16	ND	0.06	0.09	0.14
337	Furniture & related product	0.72	0.34	0.16	0.09	0.01	0.01	0.02	0.03
339	Miscellaneous	0.33	0.18	0.10	0.05	0.03	0.06	0.08	0.12

Note. ND denotes not disclosed.

Table 7. Average Share Shipped 500+ Miles (or Exported) Conditional on the Share Shipped 100+ Miles

		Total Employment			
	All	**0 to 19**	**20 to 99**	**100 to 499**	**over 500**
All	0.56	0.45	0.51	0.58	0.66
Share shipped 100+ miles					
0 to .25	0.31	0.29	0.27	0.35	0.41
.25 to .50	0.35	0.28	0.33	0.37	0.49
.50 to .75	0.43	0.37	0.41	0.44	0.55
.75 to .90	0.54	0.52	0.52	0.54	0.62
.90 to .95	0.61	0.57	0.58	0.61	0.66
.95 to 1.0	0.69	0.65	0.69	0.69	0.71

Note. The share is calculated as the sum of the export share and the share going at least 500 miles, divided by the share going at least 100 miles. Plants with no shipments beyond 100 miles are excluded.

Table 8. Establishment Count for Average Share Shipped 500+ Miles (or Exported) Conditional on the Share Shipped 100+ Miles

		Total Employment			
	All	**0 to 19**	**20 to 99**	**100 to 499**	**over 500**
All	28,359	1,834	8,944	14,030	3,551
Share shipped 100+ miles					
0 to .25	2,602	460	1,163	854	125
.25 to .50	2,359	232	1,016	978	133
.50 to .75	3,712	245	1,389	1,748	330
.75 to .90	4,704	243	1,515	2,415	531
.90 to .95	3,066	143	882	1,647	394
.95 to 1.0	11,916	511	2,979	6,388	2,038

Note. The share is calculated as the sum of the export share and the share going at least 500 miles, divided by the share going at least 100 miles. Plants with no shipments beyond 100 miles are excluded.

Table 9. Mileage Regressions

	Share <50 miles	Share <50 miles	Share <100 miles	Share <100 miles	Share <500 miles	Share <500 miles	Share >= 500 miles	Share >= 500 miles
Intercept	0.46	0.29	0.54	0.36	0.75	0.67	0.22	0.30
	(0.006)	(0.037)	(0.006)	(0.040)	(0.007)	(0.041)	(0.006)	(0.040)
sizedum2	-0.20	-0.13	-0.21	-0.14	-0.13	-0.10	0.11	0.09
	(0.006)	(0.006)	(0.007)	(0.006)	(0.007)	(0.006)	(0.007)	(0.006)
sizedum3	-0.32	-0.21	-0.34	-0.23	-0.24	-0.18	0.20	0.15
	(0.006)	(0.006)	(0.007)	(0.006)	(0.007)	(0.007)	(0.007)	(0.006)
sizedum4	-0.36	-0.22	-0.40	-0.25	-0.33	-0.21	0.25	0.16
	(0.007)	(0.007)	(0.008)	(0.008)	(0.009)	(0.008)	(0.008)	(0.008)
6-digit industry controls	No	Yes	No	Yes	No	Yes	No	Yes
R-Square	0.10	0.33	0.10	0.36	0.07	0.36	0.05	0.29
Adj R-Sq	0.10	0.31	0.10	0.34	0.07	0.35	0.05	0.28
Num. obs.	30148	30148	30148	30148	30148	30148	30148	30148

Note. Standard errors in parentheses. The reported dummy variables for size correspond to total employment of 20 to 99, 100 to 499, and over 500.

Table 9. Mileage Regressions, continued

	Export share	Export share	Same zip code share	Same zip code share	Max zip code share	Max zip code share
Intercept	0.03	0.03	0.12	0.07	0.42	0.33
	(0.003)	(0.020)	(0.003)	(0.020)	(0.005)	(0.032)
sizedum2	0.02	0.01	-0.08	-0.05	-0.12	-0.10
	(0.003)	(0.003)	(0.003)	(0.003)	(0.005)	(0.005)
sizedum3	0.05	0.03	-0.10	-0.07	-0.15	-0.13
	(0.003)	(0.003)	(0.003)	(0.003)	(0.005)	(0.005)
sizedum4	0.08	0.05	-0.10	-0.06	-0.16	-0.16
	(0.004)	(0.004)	(0.004)	(0.004)	(0.006)	(0.006)
6-digit industry controls	No	Yes	No	Yes	No	Yes
R-Square	0.02	0.19	0.04	0.13	0.03	0.15
Adj R-Sq	0.02	0.18	0.04	0.11	0.03	0.14
Num. obs.	30148	30148	30148	30148	30148	30148

Table 10. Trimmed Mean Prices by Establishment Size and Distance Category

Share shipped over 100 miles	Total Employment 0 to 19	20 to 99	100 to 499	over 500
All Establishments				
.75 to 1.0	1.41	0.86	1.48	1.96
.50 to .75	2.61	1.57	1.00	1.45
.25 to .50	1.94	1.51	1.28	1.28
.10 to .25	4.63	1.97	1.69	2.15
.05 to .10	7.25	2.76	2.56	2.41
.00 to .05	8.36	4.95	4.20	4.82
Establishments with 20 or more shipping records				
.75 to 1.0	0.56	0.62	1.36	1.84
.50 to .75	1.87	1.54	1.00	1.45
.25 to .50	1.89	1.48	1.28	1.26
.10 to .25	4.98	1.92	1.67	2.10
.05 to .10	7.09	2.72	2.55	2.41
.00 to .05	9.26	4.95	4.06	4.75

Table 11. Trimmed Mean Prices by Establishment Size and Industry

NAICS		Total Employment 0 to 19	20 to 99	100 to 499	over 500
All Establishments					
311	Food	0.26	0.37	0.78	0.91
312	Beverage & tobacco product	ND	0.30	0.31	0.38
313	Textile mills	6.93	4.01	3.03	2.94
314	Textile product mills	7.87	7.00	4.20	3.53
315	Apparel	25.07	15.97	12.07	9.71
316	Leather & allied product	10.64	10.87	11.18	9.81
321	Wood product	0.25	0.23	0.44	0.84
322	Paper	0.57	0.70	0.55	0.42
323	Printing & related support activities	10.21	5.46	3.74	1.55
324	Petroleum & coal products	0.01	0.11	0.10	0.08
325	Chemical	0.82	0.99	1.27	1.79
326	Plastics & rubber products	3.31	2.11	2.17	2.24
327	Nonmetallic mineral product	0.02	0.06	0.21	0.99
331	Primary metal	ND	1.49	1.22	0.80
332	Fabricated metal product	5.91	2.70	3.01	5.31
333	Machinery	16.17	12.91	12.20	7.43
334	Computer & electronic product	68.40	74.46	72.57	97.05
335	Elect. equip., appliance, & component	29.93	11.30	5.44	4.49
336	Transportation equipment	11.74	5.45	4.32	4.82
337	Furniture & related product	3.10	2.93	2.64	2.69
339	Miscellaneous	19.19	13.06	11.68	14.33
Establishments with 20 or more shipping records					
311	Food	0.17	0.35	0.78	0.92
312	Beverage & tobacco product	ND	0.26	0.30	0.38
313	Textile mills	10.69	4.09	3.05	2.97
314	Textile product mills	8.03	6.77	4.04	3.53
315	Apparel	26.63	18.69	12.45	9.98
316	Leather & allied product	11.15	11.24	10.79	9.60
321	Wood product	0.24	0.22	0.38	0.84
322	Paper	0.63	0.70	0.56	0.42
323	Printing & related support activities	10.29	5.50	3.79	1.61
324	Petroleum & coal products	0.01	0.11	0.10	0.08
325	Chemical	0.87	0.97	1.26	1.78
326	Plastics & rubber products	3.47	2.05	2.18	2.23
327	Nonmetallic mineral product	0.02	0.05	0.20	0.98
331	Primary metal	6.03	1.49	1.22	0.79
332	Fabricated metal product	7.23	2.69	2.97	5.20
333	Machinery	14.33	13.46	12.18	7.29
334	Computer & electronic product	61.47	70.25	72.80	96.76
335	Elect. equip., appliance, & component	29.77	11.12	5.50	4.46
336	Transportation equipment	7.27	5.55	4.22	4.76
337	Furniture & related product	2.62	2.74	2.64	2.67
339	Miscellaneous	19.50	13.79	11.36	13.67

Note. ND denotes not disclosed.

Table 12. Price Regressions

	All estabs	All estabs	Estabs w/ 20+ Shipments	Estabs w/ 20+ Shipments
Intercept	-0.01	-1.15	-0.35	-1.15
	(0.046)	(0.164)	(0.052)	(0.164)
distdum2	0.47	0.02	0.64	0.03
	(0.053)	(0.029)	(0.055)	(0.030)
distdum3	0.64	0.07	0.84	0.08
	(0.047)	(0.026)	(0.048)	(0.027)
distdum4	0.99	0.17	1.19	0.19
	(0.044)	(0.026)	(0.046)	(0.027)
distdum5	1.30	0.23	1.50	0.25
	(0.049)	(0.028)	(0.051)	(0.029)
distdum6	1.81	0.29	2.00	0.30
	(0.038)	(0.023)	(0.040)	(0.025)
sizedum2	-0.23	-0.20	-0.10	-0.19
	(0.047)	(0.026)	(0.053)	(0.029)
sizedum3	-0.19	-0.30	-0.05	-0.28
	(0.047)	(0.026)	(0.052)	(0.029)
sizedum4	0.11	-0.35	0.25	-0.35
	(0.056)	(0.032)	(0.060)	(0.034)
6-digit industry controls	No	Yes	No	Yes
R-Square	0.10	0.75	0.11	0.76
Adj R-Sq	0.10	0.75	0.11	0.76
Num. obs.	30148	30148	28084	28084

Note. Dependent variable in each case log(value/pound). Standard errors are in parentheses. The reported dummy variables for size correspond to total employment of 20 to 99, 100 to 499, and over 500. The reported dummy variables for distance correspond to the share shipped less than 100 miles being in (.5,.75], (.25,.5], (.1,.25], (.05,.1], [0, .05].