New Measurement of Export Participation in U.S. Manufacturing*

Christoph E. Boehm
University of Texas at Austin
and NBER

Aaron Flaaen
Federal Reserve Board of
Governors

Nitya Pandalai-Nayar
University of Texas at Austin
and NBER

January 2023

Abstract

We measure export participation rates in the U.S. manufacturing sector using a new administrative dataset and compare them to participation rates constructed from the commonly used Census of Manufacturers (CM). Both at the establishment and firm level export participation rates are near 40 percent in the administrative data, almost twice as high as in the CM. The discrepancy appears to result predominantly from under-counting of small exporters in the CM. Our findings call for reconsidering the conventional wisdom that around 20 percent of manufacturing firms export.

*Email: Boehm-chris.e.boehm@gmail.com; Flaaen- aflaaen@gmail.com; Pandalai-Nayar- npnayar@utexas.edu. We thank our discussant Michael Peters as well as conference participants at the ASSA Annual meetings for helpful comments. Any opinions and conclusions expressed herein are those of the authors and do not necessarily represent the views of the Board of Governors or its research staff. All results have been approved for disclosure by the Census Bureau.

1. Introduction

The concept of export participation has played an important role for the interplay between empirical and theoretical work in international trade over the past several decades. As new micro-level datasets became available in the late 1990s and early 2000s, researchers discovered that at the relevant unit of analysis—the plant or firm—exporting was a relatively uncommon phenomenon: Only about 20 percent of manufacturing firms sold their goods abroad. At the time, prevailing trade theories could not account for these new empirical facts on the extensive margin of exporting.[1] The new measures of export participation played an important role in the heterogeneous firm revolution in international trade theory exemplified by Melitz (2003), Chaney (2008), Bernard et al. (2004), Bernard, Redding and Schott (2007), and Arkolakis (2010), among many others. As a result, there are few stylized facts that are as familiar to trade economists as the proportion of firms that engage in exporting activity.

As is the case with any statistic, the measurement of export participation is subject to limitations and measurement issues specific to the underlying data source. While appropriate qualifications were widely acknowledged at the start, the challenges associated with measuring the conceptual object in available data have since largely been ignored as the stylized facts became embedded in the collective conscience of the literature. In this paper we take a fresh look at the measurement of export participation, leveraging a new data resource recently developed in Boehm et al. (2021) that, for the first time, links administrative records of U.S. export transactions to individual establishments. Accurately identifying the proportion of exporting establishments and firms is important for a variety of reasons such as estimating fixed costs of exporting, evaluating the welfare benefits of new varieties when accounting for firm-by-product differentiation, and understanding how firm revenues are impacted by foreign shocks. Our paper provides new measurement of this key statistic, addressing some of the limitations in previous work.

Our main finding is that export participation in manufacturing in the United States is more common than previously thought. In contrast to the "conventional wisdom" that 20 percent of manufacturing firms export, our new estimates find export participation is near 40 percent. We find that most of these previously unidentified exporters are responsible for small volumes of exports, which suggests that prior estimates based on the Census of Manufacturing may suffer from respondent bias when exports are a small fraction of overall

[1] For instance, see Bernard et al. (2007*a*) (pages 109-110): "... while old trade theory can explain why a country is a net importer in one set of industries and a net exporter in another set, it cannot explain why some firms export and others produce solely for the domestic market, or how the firm-level decision to export interacts with comparative advantage."

shipments.

2. Prior Work Documenting Export Participation

The first comprehensive estimates of plant export participation in the U.S. come from Bernard and Jensen (1995), who exploited a survey question from the 1987 Census of Manufacturers (CM) to highlight that 14.6 percent of manufacturing plants engaged in exporting.[2] Using the equivalent census in 2002, Bernard et al. (2007*a*) cite 18 percent of manufacturing *firms* were exporting, while Bernard et al. (2007*b*) provides a similar estimate of 20 percent of *plants* exporting in that year. These 2002 data have been the most highly influential in cementing the stylized fact of this statistic in the literature.

The second source for statistics on firm-level export participation comes from a linkage of administrative data on export transaction records to the firm level data from the U.S. Census Bureau. The resulting dataset, called the Longitudinal Foreign Trade Transactions (LFTTD) dataset, was highlighted by Bernard et al. (2007*a*) and Bernard, Jensen and Schott (2009) and provided a more complete picture of firm-level trade than was previously possible.[3] Focusing on all firms in the U.S. economy, Bernard, Jensen and Schott (2009) used the LFTTD to document that 4.2 percent of U.S. firms were exporters in 2000, up from 3.5 percent in 1993.

Each of these data sources for measuring export participation has both conceptual and practical issues. As a survey the CM relies on accurate reporting from respondents who may have imperfect knowledge of the wide array of statistics being requested. For the case of the export variable, the survey instructs the respondent: "Report the value of products shipped for export. Include shipments to customers in the Commonwealth of Puerto Rico, and U.S. possessions, as well as the value of products shipped to exporters or other wholesalers for export. Also include the value of products sold to the U.S. Government to be shipped to foreign governments. Exclude products shipped for further manufacture, assembly, or fabrication in the United States."[4] Importantly, these instructions ask for reporting of exports

[2] In follow-up work Bernard and Jensen (2004) indicate that export participation of manufacturing plants rose from 21 percent in 1987 to 30 percent in 1993, although both of these statistics were based on a limited sample—plants with at least 20 employees—due in part to limitations with the 1992 CM. Under the same sample restriction Lincoln and McCallum (2018) document an almost 20 percentage point increase in export participation between 1987 and 2006. For information on export participation by plant size, see the supplementary material of Lincoln and McCallum (2018).

[3] Indeed, Bernard et al. (2007*b*) includes a LFTTD-based export participation measure for firms in the 1997 Census of Manufacturers (see Table 7) which is considerably higher (27 percent) than the corresponding 2002 statistic based on the Census of Manufacturers alone. While in line with the main point of emphasis in this paper, Bernard et al. (2007*b*) do not elaborate on what could be behind these discrepancies.

[4] See https://bhs.econ.census.gov/2012forms/MA10000.pdf, last accessed 12/19/2022.

that are routed through a separate exporter or wholesaler firm, but not any products that are indirectly exported following additional manufacture (presumably because those subsequent exports would be captured by the reporting of another manufacturing establishment). There is still the possibility of respondent confusion, such as for products for which the export process is initiated by other, administrative, establishments of the firm (those not considered a separate "exporter" or wholesale firm). Perhaps most important, though, is whether or not the survey respondent has accurate information on which shipments of the firm are ultimately destined for export markets. Of course, the advantage of the CM is a tight link to all other data in the survey (sales, employment, etc.) at the establishment level as the unit of observation.

The LFTTD has its own shortcomings as a source for measuring export participation. Most prominently, the identifier on export transactions is a tax identifier. While less precise than the address of an establishment, this identifier can be linked to the firm identifiers used by the U.S. Census Bureau. Therefore *plant*-level export participation cannot be measured directly based on the LFTTD. With other surveys available only at the establishment level, researchers are left in the awkward position of reconciling establishment-level variables with firm-level export information. The LFTTD also has the opposite problem of the CM: the firm associated with the export transaction may be a pure wholesaler or retailer, and not the one engaged in actual production. The advantages of the LFTTD are improved data quality from the link to administrative sources (Customs records), as well as much richer detail (such as transaction-level records with information on product and destination countries, among other things, at the daily frequency).

3. New Estimates of Export Participation in U.S. Manufacturing

In recent work, Boehm et al. (2021) develops a new data resource that links the transaction-level trade data in the LFTTD to individual establishments, thereby alleviating one of the principle disadvantages of that data product for international trade researchers. A full description of this data is available in Boehm et al. (2021); in short, they combine several sources of information—including on the organizational structure of the firm, geographical information associated with the export transaction, product-industry concordances, and survey evidence—to identify the establishment that is most likely the origin of each export transaction recorded in the LFTTD. This new data resource allows researchers to take a fresh look at export participation at the establishment level, which was previously not possible

Table 1: Establishment/Firm-Level Export Participation, by Source

Source	*Percent reporting positive exports*			
	Establishment Level[1]	Firm Level	Establishment Level[1]	Firm Level
	2007		*2012*	
CM (survey)	21%	19%	25%	23%
LFTTD (administrative)	41%	36%	44%	39%

[1] Establishment level exports are constructed as described in Boehm et al. (2021).
Notes: The table reports the fraction of overall plants and firms identified as exporters based on the source of data. Columns (1) and (2) show statistics for 2007 and columns (3) and (4) for 2012.

with U.S. Census administrative data.

In our comparison of export participation by establishments and firms in the CM and LFTTD, we focus attention on the manufacturing sector. To do this, we take the set of establishments and firms in the CM and match them to the exporting indicators in the LFTTD. We focus attention on manufacturing to be consistent with the stylized facts on exporting in the literature and since the goods-producing sector is typically the appropriate analog in models of trade. One challenge with this approach is that limiting export records tied to manufacturing establishments will fail to capture exporting undertaken through an intermediary, as these transactions in administrative records could show up either as linked to a non-manufacturing establishment of the firm, or linked to a different firm altogether.

Table 1 shows export participation measures for two Census years, 2007 and 2012. For the year 2012 we find that 44 percent of manufacturing establishments and 39 percent of manufacturing firms engage in exporting based on the administrative (LFTTD) source, considerably higher than the participation rates from the survey-based (CM) source. It is also notable to observe that export participation increases between 2007 and 2012, regardless of data source, or unit of analysis.

4. Exploring Sources of Misalignment

A striking feature of Table 1 is the discrepancy in magnitudes of export participation across the two data sources. Export participation is higher in the administrative data regardless of whether the unit of analysis is the establishment or the firm, suggesting that the source of the discrepancy is not an artifact of the method used in Boehm et al. (2021) to assign firm-level

LFTTD exports to individual establishments. Before exploring the reasons for misalignment between these two data sources, we first discuss the potential reasons:

1. **Respondent Misreporting**: The survey respondent may misreport an establishment's exports. While this could lead to an over-estimate or under-estimate of export *values*, from the perspective of *participation* it appears more likely that a respondent in the CM would fail to report small values of an establishment's exports, hence leading to an under-estimate of participation.[5]

2. **Exporting via Intermediaries**: Following the instructions in the CM highlighted above, the survey respondent could report goods as exports that are "shipped to exporters or other wholesalers for export." When focusing on manufacturing establishments and firms, this could increase survey-based participation relative to the administrative-based participation measures. The reason is that in the administrative data the subsequent export transaction would be recorded as linked to a different establishment or firm, which may not be in the manufacturing sector (e.g., a wholesaler).

3. **Issues specific to the method for assigning export transactions to establishments in Boehm et al. (2021)**: We consider below whether aspects of the methodology behind the new dataset utilized in this paper affects our plant-level export participation calculations.

To assess these potential reasons for misalignment we study the matrix of export participation by source, as shown in Table 2. Panel A of Table 2 shows that in 2012, 88 percent (22/25) of those establishments recording positive exports in the CM also record positive exports in the LFTTD-based export measure. On the other hand, only 50 percent (22/44) of establishments recording positive exports in the LFTTD also record positive exports in the CM. Hence, it appears that respondent under-reporting is more problematic than accounting for exports via intermediaries.

This pattern also holds at the firm-level, as shown in Panel B of Table 2. 83 percent (19/23) of firms reporting positive exports in the CM also record positive exports in the LFTTD measure of exports. Conversely, only 49 percent of firms recording positive exports in the LFTTD also record positive exports in the CM.

To further explore potential respondent misreporting in the CM, we consider a measure of export intensity, calculated as the value of exports (from each source) as a fraction of CM-identified shipments. Figure 1 plots kernel density estimates of this statistic by source

[5]Misreporting is less of a concern in the LFTTD administrative side, since reporting is required by law.

Table 2: Plant/Firm-Level Export Participation, by Source

Panel A: Establishment-Level

LFTTD Exports \ CM Exports	No	Yes	
No	53%	3%	56%
Yes	22%	22%	44%
	75%	25%	

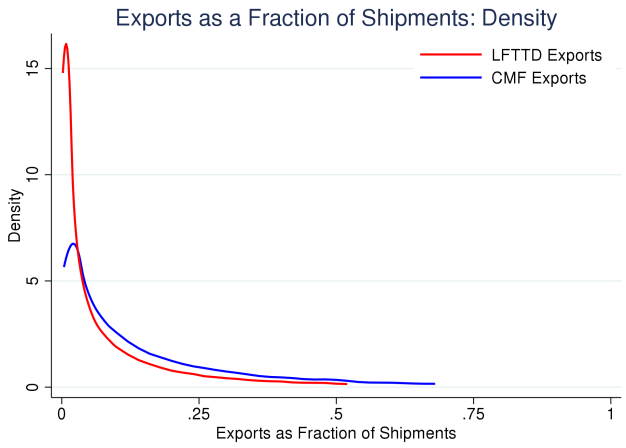
Panel B: Firm-Level

LFTTD Exports \ CM Exports	No	Yes	
No	58%	3%	61%
Yes	20%	19%	39%
	77%	23%	

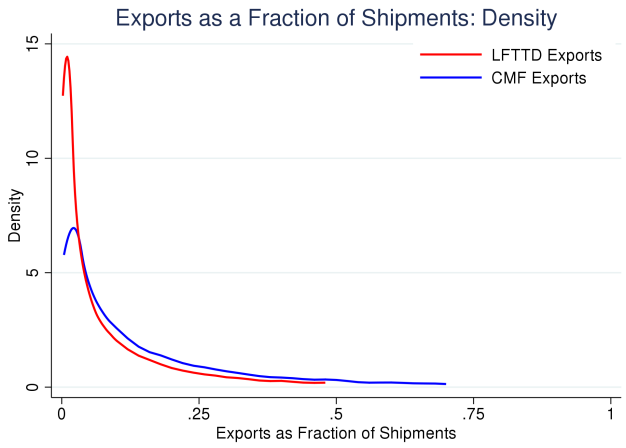
Notes: The table reports the fraction of overall plants and firms identified as exporters based on the source of data. The statistics are for Census year 2012. Appendix Table A1 reports analogous numbers for Census year 2007.

Figure 1: Export Share of Shipments: Distribution by Source of Exports

(a) Establishment-Level



(b) Firm-Level



Sources: Author's calculations using LFTTD and CM as explained in the text.
Notes: Both density estimates are constructed separately and integrate to unity. They exclude establishments or firms that report zero exports. Per Census disclosure rules, the top and bottom 5 percent of each distribution is truncated. Density estimates are for Census year 2012. Results for Census year 2007 are reported in Appendix Figure A2.

of the export variable.[6] The figure shows that the distribution based on the LFTTD source of exports has a substantially greater mass of small export shares than the CM-source, a feature consistent with the presence of under-measurement of exporting in the CM for establishments with low levels of export activity. This is true both at the establishment and firm level of aggregation.

[6]The density functions are estimated among establishments/firms with positive exports, and hence do not capture those reporting zero exports. The fact that the red line lies below the blue line for higher export intensities therefore does not indicate that there is a smaller absolute number establishments (or firms) with high export intensities when exports are measured based on the administrative data of the LFTTD.

As a further check on how intermediaries may affect measurement in the LFTTD, we first removed the exports of non-manufacturing establishments in our establishment-based LFTTD measure before summing up to the firm-level.[7] The resulting alternative firm-level measure of participation excludes cases in which a firm has manufacturing establishments but exports other firms' goods as an intermediary. We find such cases to be quite rare: the alternative firm-level export participation rate in 2012 is 38.5 percent, nearly identical to the 39 percent in the LFTTD.[8]

We next evaluate whether aspects of the method developed in Boehm et al. (2021) for assigning export transactions to individual establishments may affect the alignment. While the method largely assigns export transactions uniquely to establishments (see Appendix B for details), in some rare cases the transactions are assigned proportionally to employment across several establishments when other underlying data sources are not informative. This could artificially inflate export participation in the LFTTD establishment data, though it would not affect the corresponding firm-level statistics in Table 1. To check this, we instead reassign all of these exports to the largest likely establishment and recompute the statistics in Table 1. The fraction of establishments reporting exports in the LFTTD in 2012 falls marginally to 40.5%, but clearly this explanation cannot account for the participation discrepancy across sources.

Other aspects of the data serve to confirm the validity of the establishment-level export measure constructed in Boehm et al. (2021) and the associated export participation statistics. For instance, Figure 1 shows that unrealistic levels of export intensity (those above one) are exceedingly rare.[9] Moreover, among establishments with positive exporting in both sources, the correlation of exports-per-shipment in 2007 is high at 0.9 and 0.78 for firms and establishments, respectively.[10]

Finally, changes in measured export participation between sources and across two Census years—2007 and 2012—suggest that the higher LFTTD-based participation measure is not due to occasional exporters. As we show in Appendix Table A3, conditional on survival, solely LFTTD-identified exporters in 2007 are substantially more likely to continue to export in 2012 than solely CM-identified exporters (by 11 p.p. for establishments and by 18 p.p. for firms).

[7] While the method in Boehm et al. (2021) prioritizes manufacturing establishments, it occasionally assigns exports to non-manufacturing establishments depending on features of the transaction or firm.

[8] Such cases would not be characterized under the "factoryless goods producers" (FGPs) definition as in Bernard and Fort (2015) since they do have a manufacturing presence. Nevertheless, the effect of FGPs on the scale of manufacturing activity over time likely plays an important role in export participation.

[9] Census disclosure rules require suppression of the top and bottom 5 percent of the distribution.

[10] In 2012, the correlation of exports-per-shipment is 0.85 and 0.65 for firms and establishments respectively.

5. Concluding Remarks

This paper explores several measurement issues relating to export participation at the establishment and firm-level in the United States. Survey and administrative data each suffer from limitations; combining these two sources reveals possible sources of misalignment. On balance the evidence points to higher export participation in the manufacturing sector than was previously appreciated. Approximately 40 percent of firms and plants export. These estimates are based on data, which—unlike data from the CM—do not capture possible exporting through wholesalers.[11] In that sense, export participation from either source alone should be a lower bound on the fraction of exporting manufacturing establishments in the U.S.

This new, higher, estimate of export participation in manufacturing is important for researchers calibrating models featuring heterogeneous firms. For instance, our new estimates substantially raise the mass of traded varieties in models in which firms offer differentiated products. Beyond the export participation rate, these data could help answer other new research questions, such as whether plants within a firm specialize in serving domestic or specific foreign markets.

References

Arkolakis, Costas. 2010. "Market Penetration Costs and the New Consumers Margin in International Trade." *Journal of Political Economy*, 118(6): 1151–1199.

Bernard, Andrew B., and J. Bradford Jensen. 1995. "Exporters, Jobs, and Wages in U.S. Manufacturing: 1976 – 87." *Brookings Papers on Economic Activity*, 1995: 67–112.

Bernard, Andrew B., and J. Bradford Jensen. 2004. "Entry, Expansion, and Intensity in the U.S. Export Boom, 1987-1992." *Review of International Economics*, 12(4): 662–674.

Bernard, Andrew B., and Teresa C. Fort. 2015. "Factoryless Goods Producing Firms." *American Economic Review Papers and Proceedings*, 105(5): 518–523.

Bernard, Andrew B., Jonathan Eaton, J. Bradford Jensen, and Samuel S. Kortum. 2004. "Plants and Productivity in International Trade." *American Economic Review*, 93(4): 1268–90.

[11] Another approach, which we leave for future work, would be to take a more comprehensive exploration of the goods sector and look jointly at manufacturing, wholesale, and retail firms, as in Bernard et al. (2010).

Bernard, Andrew, Bradford Jensen, and Peter Schott. 2009. "Importers, Exporters, and Multinationals: A Portrait of Firms in the U.S. that Trade Goods." In *Producer Dynamics: New Evidence from Micro Data*. , ed. Timothy Dunne, J. Bradford Jensen and Mark J. Roberts, 513–552. Chicago:University of Chicago Press.

Bernard, Andrew B., Stephen J. Redding, and Peter K. Schott. 2007. "Comparative Advantage and Heterogeneous Firms." *Review of Economic Studies*, 74(1): 31–66.

Bernard, Andrew, J. Bradford Jensen, Stephen J. Redding, and Peter K. Schott. 2007*a*. "Firms in International Trade." *Journal of Economic Perspectives*, 21(3): 105–130.

Bernard, Andrew, J. Bradford Jensen, Stephen J. Redding, and Peter K. Schott. 2007*b*. "Firms in International Trade." NBER Working Paper 13054.

Bernard, Andrew, J. Bradford Jensen, Stephen J. Redding, and Peter K. Schott. 2010. "Wholesalers and Retailers in U.S. Trade." *American Economic Review Papers and Proceedings*, 100(2): 408–413.

Boehm, Christoph E., Aaron Flaaen, Nitya Pandalai-Nayar, and Jan Schlupp. 2021. "The Local Area Incidence of Exporting." Working Paper.

Chaney, Thomas. 2008. "Distorted gravity: The intensive and extensive margins of international trade." *The American Economic Review*, 98(4): 1707–1721.

Lincoln, William F., and Andrew H. McCallum. 2018. "The Rise of Exporting by U.S. Firms." *European Economic Review*, 102: 280–297.

Melitz, Marc J. 2003. "The Impact of Trade on Intra-Industry Reallocations and Aggregate Industry Productivity." *Econometrica*, 71(6): 1695–1725.

Pierce, Justin, and Peter Schott. 2012. "A Concordance Between U.S. Harmonized System Codes and SIC/NAICS Product Classes and Industries." *Journal of Economic and Social Measurement*, 37(1-2): 61–96.

A. Appendix A: Sources of U.S. Data on Firm and Establishment Level Exports

The original source for firm and establishment level export data from the U.S. Census Bureau comes from the quinquennial Census of Manufacturers (as well as the annual supplement, the Annual Survey of Manufacturers (ASM)). This survey asks establishments to report the dollar value of their shipments that are destined for foreign countries. The advantage of this survey-based question is the tight mapping between export values and a particular manufacturing plant ostensibly involved with the actual production of that exported product. There are numerous disadvantages of this source, however. There is no product or destination-level detail, it is only an annual measure of total exports, there are concerns of reliability due to the survey basis of the reporting, there is no information on the import side, and the longitudinal nature of the data is limited.

Detailed information on trade transactions are compiled by U.S. Customs for purposes of enforcing trade laws, documenting trade flows, and monitoring border security. The administrative documentation attached to these transactions offers rich detail, including the value, quantity, detailed product, country of export/import, port location, type of transaction, and more. Beginning with Bernard, Jensen and Schott (2009), the U.S. firm associated with these trade transactions was linked to other Census datasets, thus improving the information on firm-level trade for study by economists. The principal identifier for linkage is the employer identification number (EIN), as it was recorded by Customs Bureau documents accompanying individual shipments and also as part of the establishment/firm register in the Census Bureau.[12] The resulting Linked/Longitudinal Foreign Trade Transactions (LFTTD) database has been a very useful resource for trade economists studying import/export patterns by U.S. firms. One disadvantage, however, is that the EIN-based matching does not allow exports or imports to be assigned to individual establishments (plants) of the firm. The drawbacks of this limitation are discussed further below.

The final major resources for firm-level trade are the surveys of multinational firms collected by the Bureau of Economic Analysis. The level of aggregation is a mix between the firm and establishment (affiliate), depending on the direction of trade, size of the firm, or whether a particular year falls under the BEA's benchmark survey period. Like the survey-based information from the Census Bureau, the BEA data do not have extensive information

[12]On the export side, the EIN is listed on the "Shippers Export Declaration. The exceptions are shipments to Canada, which do not contain EINs but rather a field listing the firm name. On the import side, the Customs Forms (7501 and 7503) record the EIN representing the "ultimate consignee" of the imported goods.

on products or high-frequency detail of shipments. And while the focus on multinational firms provides for useful splits between arms-length and related-party trade, the BEA data do not have any information on non-multinational firms.

B. Appendix B: Details on Establishment-Level Allocation Methodology

This Appendix provides a shortened summary of the method developed by Boehm et al. (2021) on the construction of a dataset that assigns the firm-level LFTTD export transactions to individual establishments. Please refer to Boehm et al. (2021) for further details on their approach.

Their method utilizes information on products and industries, on the geography of where export transactions originate, and (to a limited extent) the CM and ASM. We describe each of these sources below, and then give a rough summary of how they combine this information to find the best establishment within a firm for each export transaction.

B.1 Industry-Based Matching

The goal of this industry-based measure is to establish a mapping between a particular exported product and a set of likely industries that could have produced that product. It relies on specific supplementary survey data from the Census of Manufacturers, as described below.

Every five years as part of the Census of Manufacturers, the Census Bureau surveys establishments on their total shipments broken down into a set of NAICS-based (6 digit) product categories.[13] Each establishment is given a form—specific to its industry—with a list of pre-specified products. There is also additional space to record other product shipments not included in the form. The resulting product trailer file to the CM allows a researcher to construct the set of industries that are primary producers of a given product.

There are several data issues that must be addressed before using the CM-Products file to infer information about the relative value of product-level shipments by a particular firm. First, the trailer file contains product-codes that are used to "balance" the aggregated product-level value of shipments with the total value of shipments reported on the base CM survey form. We drop these product codes from the dataset. Second, there are often codes that do not correspond to any official 7-digit product code identified by Census. (These

[13]The 1992 version of the CM used SIC-based product codes.

are typically products that are self-identified by the firm but do not match any of the pre-specified products identified for that industry by Census.) Rather than ignoring the value of shipments corresponding to these codes, Boehm et al. (2021) attempt to match at a more aggregated level. Through an iterated process they try to find a product code match at the 6, 5, and 4 digit product code level, and use the existing set of 7-digit matches as weights to allocate the product value among the 7-digit product codes encompassed by the more aggregated level.

Finally, the link between the Harmonized Commodity Description and Coding System (or Harmonized System, HS) codes and Standard Industrial Classification System (SIC) and North American Industrial Classification System (NAICS) product codes is referred to as a SIC base or NAICS base, depending on which CM year is being used. These basecodes are up to 8 alphanumeric characters long, with shorter basecodes representing more highly aggregated products. Given linkage between either SIC or NAICS, the first four to six digits of the basecodes are called the baseroot. Each HS code has a single baseroot, while a baseroot might be associated with multiple HS codes. They use the NAICS (or SIC) to HS concordance from Pierce and Schott (2012), to map the information from the CM-Products file to the LFTTD trade data.

We now describe how Boehm et al. (2021) construct the set of "Production-Associated Industries (PAIs)" associated with a given product. Formally, let x_{pij} denote the value of shipments of product p by establishment i in industry j during a census year. Then the total output of product p in industry j can be written as:

$$X_{pj} = \sum_{i=1}^{I_j} x_{pij},$$

where I_{jp} is the number of establishments producing p in industry j. Total output of product p is then:

$$X_p = \sum_{j=1}^{I_{jp}} X_{pj}.$$

The share of product output accounted for by a given industry j is therefore:

$$S_{pj} = \frac{X_{pj}}{X_p}.$$

Because of reporting errors and aggregation of products, Boehm et al. (2021) designate an industry as a PAI of product p provided that its share S_{pj} passes a certain threshold –

which they set at 5 percent.[14] That is, they define the set of industries for product p for which $S_{pj} > 0.05$ as J_p. Boehm et al. (2021) match individual years of the LFTTD data to the closest available Census year.

To summarize, this procedure allows one to associate a set of industries with each exported product from the LFTTD; the industry identifiers of establishments in the LBD can then be used to match products to establishments.

B.2 Location-Based Matching

Along with industry-based information on production, Boehm et al. (2021) also use geographic information to narrow the set of potential establishments involved with a particular trade transaction. Part of the shipper's export declaration form (now electronically administered via the Automated Exporter System) asks for address information on the USPPI where "goods begin their journey to the Port of Export." Both the zipcode and state information from this entry are included in the LFTTD microdata.

Although uncommon, some trade transactions in the LFTTD record a missing or incomplete zipcode. For these observations, Boehm et al. (2021) fill in missing zip-codes iteratively by replacing missing values with the largest zip-code value within the firm, country, month, HS code and baseroot observation. By attaching a U.S. location to an export transaction, this method may also assist in identifying the relevant establishment of export, though subject to a variety of limitations that are described in greater detail in Boehm et al. (2021). In short, one must be careful because the zipcode could be associated with an export transaction is the location of production or the location of export processing.

B.3 Survey-Based Information

A final set of information used to assign firm-level exports to individual establishments is the export variable included in the CM and ASM. Although the CM should be comprehensive across all manufacturing establishments, one must be more careful in ASM years given that not all of a firm's manufacturing plants may be included in the survey sample. For this reason Boehm et al. (2021) use the export indicator from the CM/ASM as a way of distinguishing between a firm's plants only when the industry/location information specified above yields multiple plants associated with a given transaction.

[14] They note that they have varied this threshold without affecting their results.

B.4 Assignment Procedure

The paragraphs below broadly outline how Boehm et al. (2021) combine these sources of information to assign all firm-level LFTTD exports to the most likely establishment associated with that export. This assignment will not always identify the establishment of export manufacture for several reasons. First, the PPI identified in the export declaration may be a non-manufacturing firm entirely that is solely involved in the export of the good; in such cases it is not possible to identify the establishment of production. Second, if the PPI is the firm of manufacture but processes shipments for export in a separate establishment, then our location information will point to a non-production establishment. Hence, while the assignment procedure described below attempts to prioritize establishments of production over non-manufacturing establishments, the data will often only identify the establishment involved with the export process. It is worth emphasizing that both production and non-production establishments involved with exporting activity will be impacted by trade and export markets.

To retain as much detail as possible, Boehm et al. (2021) take the raw LFTTD export data and aggregate only up to the firm, product, country, month, zipcode, port, and export-method (rail, air, etc) level. Next they make copies of each of these export observations and merge them to all of the firm's establishments. Because the LBD only registers a firm if it existed on March 12th of a particular year, some firms could be trading in the LFTTD but not exist in the LBD. To remedy this issue, Boehm et al. (2021) match the trade data not found in the LBD for that period with samples from the year prior and the year following. Using this large dataset, they retain the most likely establishment for each trade observation according to an iterative set of rules, decreasing in the degree of confidence in the establishment match.

- **Case 1: Single unit firm.** The assignment is a trivial exercise for those firms having only one establishment. Boehm et al. (2021) first remove these transactions, but flag whether these establishments are manufacturing or non-manufacturing, and whether the establishment records positive export shipments in the ASM/CM.

- **Case 2: Unique zip code match: manufacturing.** If a single zip code matches to a unique establishment, then the relevant trade is assigned to that establishment.

- **Case 3: Non-unique zip code match to PAI establishment** If there is only one establishment matching the zip code that also matches based on their PAI criteria, then they allocate all trade to the zipcode match that also aligns with the appropriate industry. Given that an HS code is assigned to multiple PAIs, it is possible for there to be several

establishments matching this case. Absent any distinguishing information on export activity from the ASM/CM, they use employment weights to allocate the exports across establishments.

- At this stage for any unassigned export transactions, Boehm et al. (2021) loop back through cases 3 through 6 but looking for PAI establishments at the NAICS-5, and then NAICS-4 industry basis.

- **Case 4: No zip code match but unique PAI establishment.** In this case, they simply assign all trade to the unique PAI establishment.

- **Case 5: No zip code match but non-unique PAI establishments.** For multiple establishments matching a PAI, they use employment weights to allocate the trade across matching establishments (absent distinguishing information on export status from the ASM/CM).

- **Case 6: Non-unique zip code match, and no PAI establishments.** They split these cases into those matching manufacturing establishments with those matching non-manufacturing establishments. If there are multiple establishments with the same zip code (which they mention is rare), they continue to use employment shares as weights.

- **Case 7: No zip code match and no PAI establishments.** For this final case, they first assign the export transaction to all manufacturing establishments of the associated firm (using either ASM/CM export share weights or employment weights). They also prioritize establishments in the wholesale (NAICS 42) and transportation/warehousing (NAICS 48) industries, based on the distribution of exports from prior matches. If there are no establishments in any of these industries, then the final step assigns the export transaction to all other establishments based on employment shares.[15]

B.5 Characteristics of the Allocation

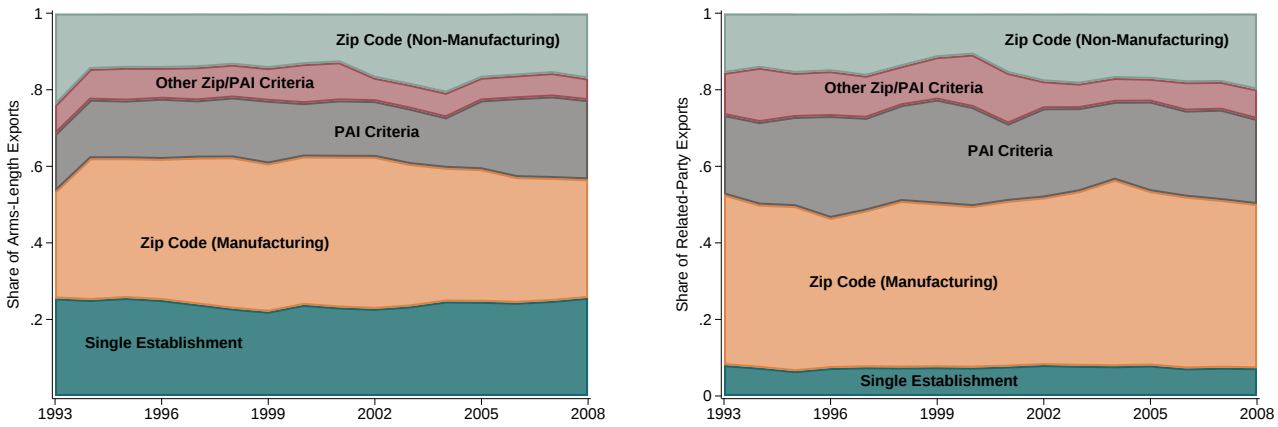
Figure A1 documents the share of overall U.S. exports that is assigned according to the hierarchy of cases as described above.

[15]Note that for these cases they take the top 20 establishments by employment size, as there are some firms with thousands of establishments.

Figure A1: Share of Exports by Type of Establishment Allocation

(a) Non-Related Party Exports

(b) Related Party Exports



Source: Reproduced from Boehm et al. (2021)'s calculations as explained in text.

C. Appendix C: Additional Results

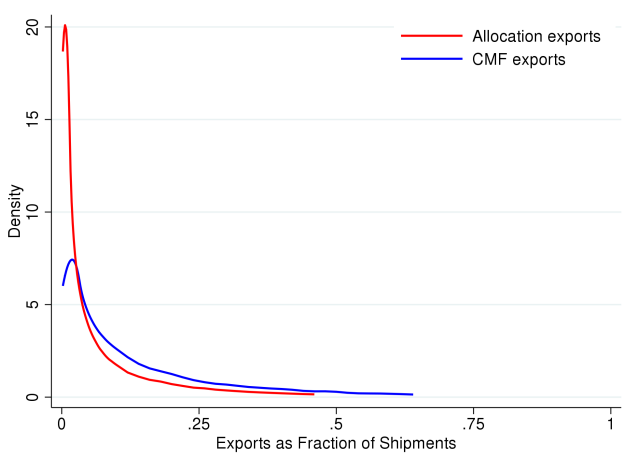
Figure A2 replicates the kernel density estimates of export intensity by source from Figure 1, but for the year 2007. Similarly, we report corresponding full detail of the matrix of alignment between the CM and LFTTD-based measures of export participation for 2007 in Table A1, analogous to Table 2 in the main text.

Changes in measured export participation by establishments and firms over two Census years–2007 and 2012—may also be informative to assess reasons for misalignment. We therefore match manufacturing establishments and firms across the 2007 CM and 2012 CM (while also noting which establishments or firms are missing in each year) and retain the various export participation measures in each year. The results are shown in Table A2. The percentages in each panel of the table sum to 100, and therefore represent the full set of establishments or firms in the two years of analysis. Beyond the issue of measurement, it is possible to construct a transition matrix for exporting, birth, and death in manufacturing over a five-year horizon from this table.

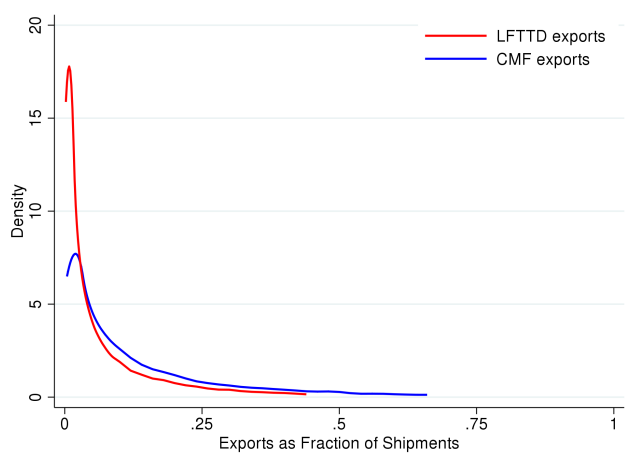
Persistence of exporting. The numbers reported at the end of section 4 are calculated from Appendix Table A3 as follows. For establishments, conditional on having positive exports in the LFTTD and zero exports in the CM in 2007 and conditional on survival, the probability of continuing to export in 2012 according to at least one of the sources (LFTTD, CM, or both) is $25.6\% + 51.4\% + 1.6\% = 78.6\%$. Similarly, conditional having positive exports in the CM and zero in the LFTTD in 2007 and conditional on survival, the probability of

Figure A2: Export Share of Shipments: Distribution by Source of Exports

(a) Establishment-Level



(b) Firm-Level



Sources: Author's calculations using LFTTD and CMF as explained in the text.
Notes: Both density estimates are constructed separately and integrate to unity. They exclude establishments or firms that report zero exports. Per Census disclosure rules, the top and bottom 5 percent of each distribution is truncated. *Allocation* refers to exports measured at the establishment level using the dataset constructed in Boehm et al. (2021). Density estimates are for Census year 2007. Results for Census year 2012 are reported in Table A2.

continuing to export in 2012 according to at least one of the sources (LFTTD, CM, or both) is 33.3% + 14.9% + 19.0% = 67.3%. The difference of 78.6% and 67.3% (= 11.4%) is reported in the text. The calculations are analogous for firms using the numbers in Panel B of Appendix Table A3.

Table A1: Plant/Firm-Level Export Participation, by Source

Panel A: Establishment-Level

LFTTD Exports	CM Exports: No	CM Exports: Yes	
No	56%	3%	59%
Yes	23%	18%	41%
	79%	21%	

Panel B: Firm-Level

LFTTD Exports	CM Exports: No	CM Exports: Yes	
No	61%	3%	64%
Yes	20%	16%	36%
	81%	19%	

Notes: The table reports the fraction of overall plants and firms identified as exporters based on the source of data. The statistics are for Census year 2007. Table 2 reports analogous numbers for Census year 2012.

Table A2: Establishment/Firm-Level Export Persistence Over Time, by Source

Panel A: Shares of Establishments

	Positive LFTTD & CM	Positive LFTTD & Zero CM	Positive CM & Zero LFTTD	Zero CM & LFTTD	Don't Exist
2007			*2012*		
Positive LFTTD & CM	8.7%	2.2%	0.3%	0.6%	2.6%
Positive LFTTD & Zero CM	3.3%	6.6%	0.2%	2.7%	5.6%
Positive CM & Zero LFTTD	0.6%	0.3%	0.3%	0.6%	0.7%
Zero Exports	1.0%	3.3%	0.7%	21.5%	18.7%
Don't Exist	2.2%	3.6%	0.6%	13.3%	

Panel B: Shares of Firms

	Positive LFTTD & CM	Positive LFTTD & Zero CM	Positive CM & Zero LFTTD	Zero CM & LFTTD	Don't Exist
2007			*2012*		
Positive LFTTD & CM	7.0%	1.8%	0.2%	0.4%	3.2%
Positive LFTTD & Zero CM	2.6%	5.1%	0.2%	2.1%	5.5%
Positive CM & Zero LFTTD	0.4%	0.2%	0.3%	0.6%	0.8%
Zero Exports	0.8%	2.7%	0.7%	22.1%	20.9%
Don't Exist	2.8%	3.8%	0.7%	15.1%	

Notes: The table reports the fraction of overall plants and firms identified as exporters based on the source of data that persist as exporters in the data between 2007 and 2012. For details on the data, see this appendix. Establishment level exports are constructed by Boehm et al. (2021). The top panel reports establishment statistics and the bottom panel reports firm statistics.

Table A3: Establishment/Firm-Level transition matrices conditional on survival

Panel A: Transition Matrix for Establishments

	Positive LFTTD & CM	Positive LFTTD & Zero CM	Positive CM & Zero LFTTD	Zero CM & LFTTD
2007		*2012*		
Positive LFTTD & CM	74.1%	18.9%	2.3%	4.8%
Positive LFTTD & Zero CM	25.6%	51.4%	1.6%	21.4%
Positive CM & Zero LFTTD	33.3%	14.9%	19.0%	32.7%
Zero Exports	3.9%	12.6%	2.5%	81.0%
Don't Exist	10.9%	18.3%	3.1%	67.7%

Panel B: Transition Matrix for Firms

	Positive LFTTD & CM	Positive LFTTD & Zero CM	Positive CM & Zero LFTTD	Zero CM & LFTTD
2007		*2012*		
Positive LFTTD & CM	74.3%	19.3%	2.2%	4.2%
Positive LFTTD & Zero CM	25.7%	51.1%	2.0%	21.1%
Positive CM & Zero LFTTD	25.8%	13.2%	21.9%	39.1%
Zero Exports	3.0%	10.2%	2.8%	84.0%
Don't Exist	12.7%	16.9%	3.1%	67.3%

Notes: The table reports the transition matrices conditional on survival between 2007 and 2012 for establishments and firms, where a state is defined as indicated by the row and column labels. The matrices are constructed from the data reported in Table A2.