Is Publicly-Reported Firm-Level Trade Data Reliable? Evidence from the UK*

Holger Breinlich†, Patrick Nolen‡ and Greg C. Wright§

February 3, 2017

Abstract

In this paper we compare firms' self-reported overseas sales, as reported in a commonly used UK financial reporting dataset, with their actual exports, as reported by Her Majesty's Revenue and Customs (HMRC). Finding that these flows are in several dimensions quite different, we then explore the implications of these differences more formally. Since several studies within the international trade literature report findings based on the self-reported export values in financial datasets, we discuss these findings in light of the departure of financial dataset-based exports from "true" (HMRC) export values.

Key Words: financial datasets, export data
JEL Codes: F10, F14

1 Introduction

Firm-level datasets have been increasingly used to explore questions related to international trade. The most common sources for these data are customs records and surveys by national government statistics agencies, while the next most commonly used data come from the self-reports of firms in their end-of-fiscal-year financial reports. These reports are made available to researchers and others by several private data providers, and are available for many countries and regions including the U.S. (CompuStat), the U.K. and Ireland (FAME), Germany (dafne), India (Prowess), the Americas and Asia (Orbis), Russia, Ukraine and Kazakhstan (Ruslana), and Europe (Amadeus), to name just a few. Here we explore the reliability of these financial datasets for international trade research, focusing on the UK and the Bureau van Dijk dataset FAME. In particular, we are motivated by the possibility that these self-reported data may be systematically misreported, which may be very important to the extent that policy decisions are informed by estimates derived from these datasets.

While financial datasets are used as a source of export information for several countries, the UK's FAME database is one of the most widely used.[1] Researchers have used FAME to explore a range of questions, many of which address fundamental issues within the international trade literature, highlighting the need to understand the extent to which the data accurately reflect the UK economy. Recent work has explored the impact of exporting on R&D (Girma, Görg and Hanley, 2008); the impact of the financial crisis on exporting (Görg and Spaliara, 2013); the relationship between the financial health of a firm, exporting, and firm survival (Görg and Spaliara, 2014; Greenaway, Guariglia and Kneller, 2007); the relationship between exporting and agglomeration economies (Greenaway and Kneller, 2003); the role of exchange rate uncertainty in the export decision (Greenaway, Kneller and Zhang, 2008); the magnitude of learning-by-exporting (Greenaway and Kneller, 2007); the contribution of exporting to UK productivity growth (Harris and Li, 2008); the relationship between exporting and firm exit (Harris and Li, 2010); and firm heterogeneity in barriers to exporting (Kneller and Pisu, 2011). FAME has also been used extensively for the evaluation of export promotion policies (e.g., Rogers and Helmers, 2010; Breinlich et al., 2012).

2 Data

We compare patterns of overseas sales across two data sources, restricting our analysis to the manufacturing sector in accordance with most of the literature. The first source is the UK's FAME data, a financial reporting dataset produced by Bureau van Dijk Electronic

---

[1]This is in part because UK customs data have only recently become accessible to researchers and the UK's main firm-level production survey (the ABI/ABS) did not contain any information about goods exports until 2011 (and only a binary indicator for export status since then).

Publishing, which includes balance sheet information for nearly all UK firms. In addition to reporting a long list of variables related to firm performance and firm finance, FAME also reports "overseas turnover", a variable that primarily captures export sales but also includes the local (overseas) sales associated with the foreign affiliate of a UK firm. This is the variable used as a proxy for export status in the studies listed above and we will refer to it as either "overseas turnover" or simply "exports" throughout this note.

We compare and contrast these FAME-reported values with those reported by another source of export information: the universe of UK transaction-level exports, collected and housed by HMRC. These data are derived from customs declaration forms associated with the physical shipment of goods across borders and should provide a more accurate picture than self-reported exports. In addition, they are not contaminated by the inclusion of local affiliate sales, which are conceptually different from exports.

We merge monthly HMRC transaction-level exports covering the period 2007 to 2010 with FAME, using a common firm identifier. The merged dataset contains two export variables: overseas sales from FAME and data on actual exports from HMRC. Throughout the analysis we also exploit additional firm-level variables such as assets, employment or sales reported by FAME.[2]

3 Comparing the FAME and HMRC data – Export Status and Export Values

We begin by asking how well FAME captures some basic facts about export activity. Figures 1 through 4 illustrate the extent to which FAME self-reported exports deviate from the true distribution of exports by comparing the value of exports and number of exporters reported in both FAME and HMRC, by quartile of firm total assets. Total assets is the only variable available for the universe of firms in FAME and is used as a proxy for size throughout. Figure 1 simply illustrates the fact that there are a greater number of HMRC-reporting (actual) exporters relative to FAME-reporting exporters, and this is true for each year in the sample and also true across the firm size distribution. There seems to be a particularly large absence of FAME-reporting exporters among the smallest firms. Figure 2 then narrows the focus to the top percentiles, where the largest disparities are again among the smallest of the large firms. This suggests that the FAME data vastly under-represent the number of exporters in all categories of firm size except

---

[2]One issue is that HMRC exports are associated with a trader identification number, which in 26 percent of cases is associated with more than one FAME identifier (HMRC trade flows need to be aggregated to the enterprise group level to be matched to FAME, and these groups often encompass several enterprises). We therefore perform our analysis with a sample that aggregates FAME variables up to the level of each unique trader identification number. We also performed the analysis on the sample of unique FAME-to-HMRC matches (the 74 percent of cases), with very similar results.

for the largest 1 percent.[3]

Across most of the distribution of export volumes there is little difference between FAME-reported values and true HMRC values. However, for the top quartile of firms as measured by assets, FAME-reported export sales vastly overstate both total UK exports as well as the importance of large firms in total exports (Figure 3). Furthermore, Figure 4 shows that the overstatement of exports among large firms in FAME is entirely driven by the concentration of export value among the very largest firms (the top 1 percent). Given the well-documented concentration of large multinational enterprises among the largest firms, the most likely explanation for this pattern is that the inclusion of local affiliate sales in FAME leads to a substantial overestimate of export values at the top of the firm size distribution.

4 Implications of Mismeasurement – Determinants of Export Status and Exporter Premia

While export status and export values are likely to be severely mismeasured in FAME, this does not necessarily invalidate the key results from the studies mentioned earlier. The two principal goals of these and other studies of export behavior are i) to understand the determinants of export status; and ii) to establish whether exporting has a positive and (possibly) causal association with firm performance indicators ("export premia"). To see whether and how the measurement error introduced by misreporting of exports in FAME changes existing insights, we replicate standard export status and premia regressions for our FAME and HMRC datasets and compare the results. Table 1 reports OLS regression results in which export status (1,0) is regressed on several firm variables. Columns (5)-(8) include year and industry fixed effects while columns (3), (4) and (7), (8) add lagged export status. First, in our preferred specifications, columns (7) and (8), both assets and turnover are positive and highly significant when applied to the HMRC export data, a result that is consistent with the literature. In contrast, these firm size proxies are near zero and not significant when applied to the FAME self-reported exports. And second, the HMRC data show a strong positive relationship between labor productivity and exporting, also consistent with the literature, which is not found in FAME. To summarize, the regressions that adopt the FAME export status variable suggest that firm size and labor productivity play no role in determining whether a firm exports or not. However, the regressions that adopt the HMRC export status variable indicate that larger firms, and more productive firms, are much more likely to export.

---

[3]Note that exporters are only required to report intra-EU exports to HMRC if they exceed an annual threshold (£250,000 in 2016). This implies that HMRC might also underestimate the number of actual exporters, suggesting that the disparity between FAME and the true figure may be even greater than reported here.

In Tables 2 through 4 we estimate export premia by regressing firm capital investment, turnover, and wages on export status. For each case we estimate OLS specifications with and without controls for assets, a proxy for firm size. There is a consistent pattern throughout, namely that the estimates are not very different from one another for both FAME self-reported exports and HMRC exports.

Finally, following the literature we estimate the productivity premia associated with export starters, export stoppers and continuing exporters, for each measure of export status (HMRC versus FAME). Formally, we regress the change in each firm's labor productivity between periods t and t+1 on a set of indicators for whether the firm started exporting, stopped exporting, or continued to export between t and t+1. Table 5 presents the results, where columns (2) and (3) control for firm assets (size) and column (3) also adds industry and year fixed effects. We see that the (negative) premium associated with export stopping is nearly identical in both cases, a result that has also been identified throughout the literature (e.g., Wagner, 2007). On the other hand, the true HMRC results suggest no statistically discernible impact of starting or continuing to export, while FAME reports a positive and significant effect of export starting, and a negative and significant effect of continuing to export. The HMRC results are more consistent with the literature (and of course reflect the true behavior of UK firms), which has typically found that firms self-select into exporting, such that the act of beginning to export has little causal impact on productivity levels. With respect to continuing exporters, the evidence from the literature is mixed as to whether there is so-called "learning-by-exporting" – i.e., rising productivity over the export tenure. However, to our knowledge there is no evidence in the literature suggesting that there is a decrease in productivity over the export tenure, as is indicated by the FAME-based result in Column (3).

5 Concluding Remarks

In this note we have explored the extent to which the export values reported in a widely used U.K. financial dataset, FAME, reflect the true export behavior of those firms. Financial datasets are a commonly used source of export information, and our results should therefore be informative in interpreting existing studies as well as in directing future work that utilizes these data.

Our analysis centers around a comparison of the export values reported in FAME with the true export values collected by HMRC. We conclude with a summary of our findings and some comments on their implications:

- Small (and, possibly, medium-sized) firms often report no exports in FAME when, in fact, they have exported. As a consequence, FAME is unreliable for estimating the total number of exporting firms.

- Export values derived from FAME substantially overstate exports for the largest firms. As a consequence, total exports reported by FAME across industries or economy-wide are not reliable.

- The determinants of export status are not very well captured by FAME. In particular, the relationships between size and exporting, and productivity and exporting, are inconsistent with the HMRC data as well as the existing literature.

- The premia associated with export status are captured fairly well by FAME. One exception is that FAME overestimates the productivity effects associated with starting to export while overstating the losses associated with continuing to export.

References

[1] Baldwin, John R., and Wulong Gu. "Export market participation and productivity performance in Canadian manufacturing." Canadian Journal of Economics 36.3 (2003): 634-657.

[2] Bernard, Andrew B., J. Bradford Jensen, and Peter K. Schott. "Falling trade costs, heterogeneous firms, and industry dynamics. No. w9639. National Bureau of Economic Research (2003).

[3] Bernard, Andrew B., and Joachim Wagner. "Exports and success in German manufacturing." Weltwirtschaftliches Archiv 133.1 (1997): 134-157.

[4] Breinlich, Holger, Giordano Mion, Patrick Nolen and Dennis Novy. "Intellectual Property, Overseas Sales, and the Impact of UKTI Assistance in Entering New Overseas Markets," Report for UK Trade and Investment (2012).

[5] Clerides, Sofronis K., Saul Lach, and James R. Tybout. "Is learning by exporting important? Micro-dynamic evidence from Colombia, Mexico, and Morocco." Quarterly Journal of Economics (1998): 903-947.

[6] Girma, Sourafel, Holger Gorg, and Aoife Hanley. "R&D and exporting: A comparison of British and Irish firms." Review of World Economics 144.4 (2008): 750-773.

[7] Gorg, Holger, and Marina-Eliza Spaliara. "Financial Health, Exports and Firm Survival: Evidence from UK and French Firms." Economica 81.323 (2014): 419-444.

[8] Gorg, Holger, and Marina-Eliza Spaliara. "Export market exit, financial pressure and the crisis." CEPR Working Paper (2013).

[9] Greenaway, David, Alessandra Guariglia, and Richard Kneller. "Financial factors and exporting decisions." Journal of International Economics 73.2 (2007): 377-395.

[10] Greenaway, David, and Richard Kneller. "Exporting, productivity and agglomeration." European Economic Review 52.5 (2008): 919-939.

[11] Greenaway, David, and Richard Kneller. "Industry differences in the effect of export market entry: learning by exporting?" Review of World Economics 143.3 (2007): 416-432.

[12] Greenaway, David, Richard Kneller, and Xufei Zhang. "The effect of exchange rates on firm exports: The role of imported intermediate inputs." The World Economy 33.8 (2010): 961-986.

[13] Harris, Richard, and Qian Cher Li. "Evaluating the contribution of exporting to UK productivity growth: some microeconomic evidence." The World Economy 31.2 (2008): 212-235.

[14] Harris, Richard ID, and Qian Cher Li. "Export Market Dynamics and the Probability of Firm Closure: Evidence from the UK." Scottish Journal of Political Economy 57.2 (2010): 145-168.

[15] Isgut, Alberto. "What's different about exporters? Evidence from Colombian manufacturing." Journal of Development Studies 37.5 (2001): 57-82.

[16] Kneller, Richard, and Mauro Pisu. "Barriers to exporting: what are they and who do they matter to?" The World Economy 34.6 (2011): 893-930.

[17] Rogers, Mark and Christian Helmers, "Intellectual property, exporting, and UKTI support for export capability building," Report for UK Trade and Investment. (2010).

[18] Wagner, Joachim. "Exports and productivity: A survey of the evidence from firm-level data." The World Economy 30.1 (2007): 60-82.

Figure 1.

HMRC vs FAME Number of Exporters
by Year and Size (Asset) Quartile, Aggregated Sample

Figure 2.

HMRC vs FAME Number of Exporters
by Year and Top Size (Asset) Quartile, Aggregated Sample

Figure 3.

HMRC vs FAME Total Reported Exports (Bil £)
by Year and Size (Asset) Quartile, Aggregated Sample

Figure 4.

HMRC vs FAME Total Reported Exports (Bil £)
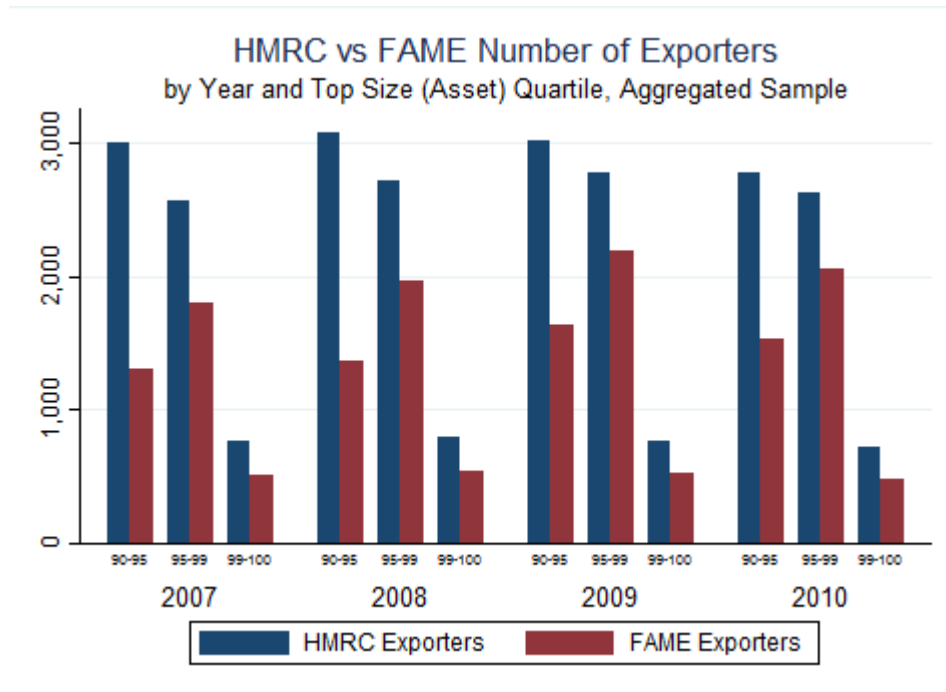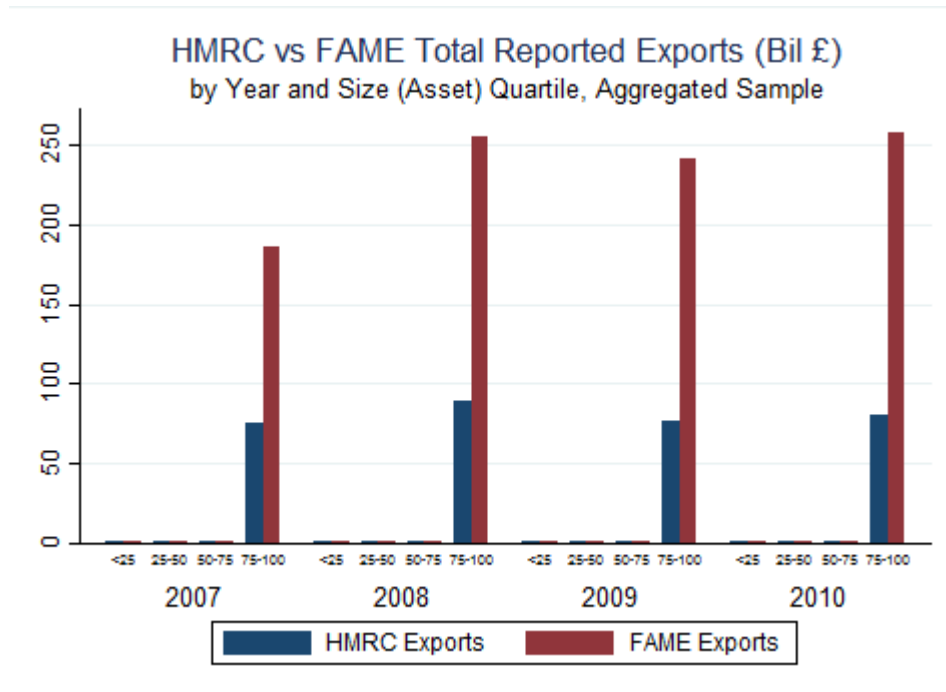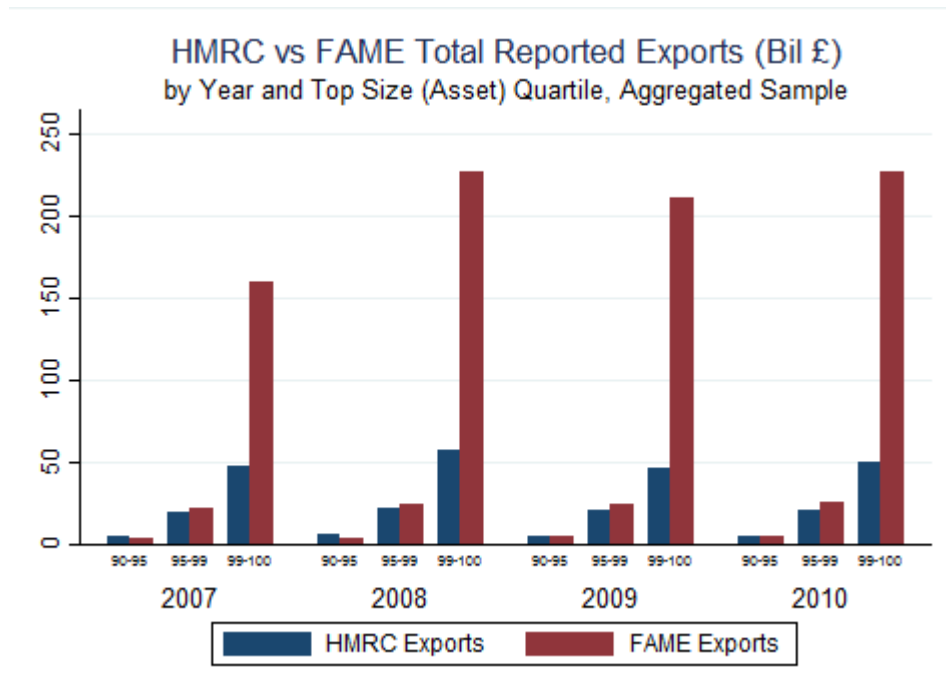by Year and Top Size (Asset) Quartile, Aggregated Sample
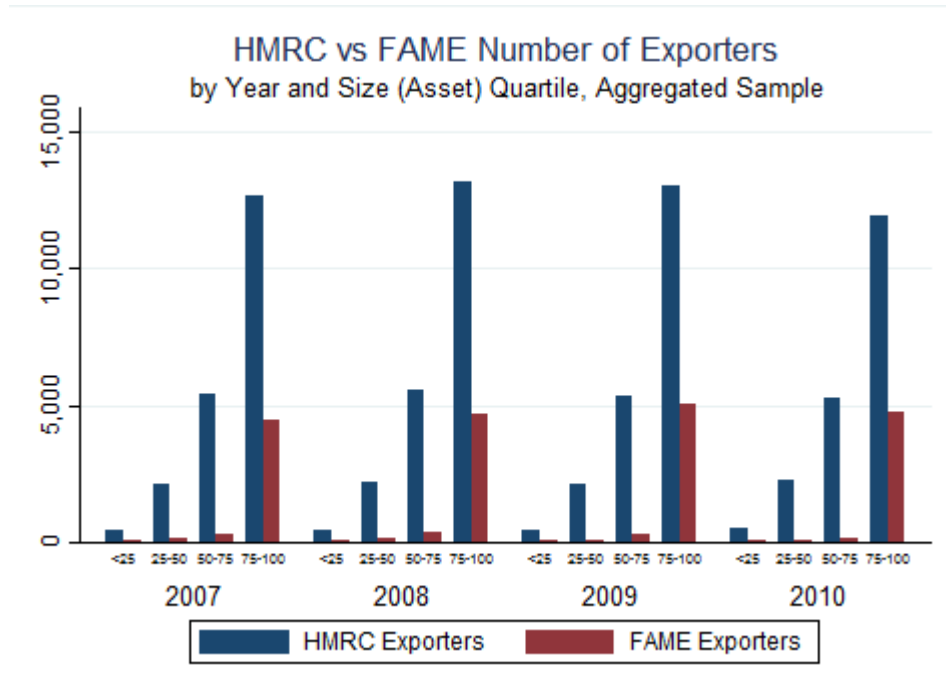
Table 1. Determinants of Export Status, FAME vs HMRC

	(1) Export > 0	(2) Export > 0	(3) Export > 0	(4) Export > 0	(5) Export > 0	(6) Export > 0	(7) Export > 0	(8) Export > 0
HMRC EXPORTS								
Log Assets	0.0714*** (0.002)	0.0448*** (0.004)	0.0043*** (0.002)	0.0031 (0.002)	0.0684*** (232.45)	0.0487*** (0.004)	0.0057*** (0.002)	0.0060*** (0.002)
Log Wage		0.0397*** (0.004)	0.0009 (0.002)	0.0092*** (0.002)		0.0361*** (0.003)	0.0006 (0.002)	0.0101*** (0.002)
Log Turnover			0.0133*** (0.002)				0.0153*** (0.003)	
Log Lagged Export Status			0.8077*** (0.008)	0.8137*** (0.008)			0.7714*** (0.013)	0.7765*** (0.013)
Labor Productivity				0.0137*** (0.002)				0.0130*** (0.003)
Year FE	N	N	N	N	Y	Y	Y	Y
Industry FE	N	N	N	N	Y	Y	Y	Y
N	409162	55031	37284	28791	409162	55031	37284	28791
R-Squared	0.200	0.214	0.736	0.701	0.281	0.360	0.752	0.721
FAME EXPORTS								
Log Assets	0.0332*** (0.002)	0.00191*** (0.004)	-0.0030* (0.002)	-0.0044* (0.002)	0.0300*** (0.002)	0.0214*** (0.003)	-0.000 (0.002)	0.0007 (0.003)
Log Wage		0.0486*** (0.004)	0.0223*** (0.002)	0.0150*** (0.002)		0.0471*** (0.03)	0.0215*** (0.002)	0.0155*** (0.002)
Log Turnover			-0.0011 (0.002)				0.0009 (0.002)	
Log Lagged Export Status			0.8039*** (0.005)	0.8158*** (0.005)			0.7752*** (0.006)	0.7836*** (0.006)
Labor Productivity				0.0037* (0.002)				0.0021 (0.002)
Year FE	N	N	N	N	Y	Y	Y	Y
Industry FE	N	N	N	N	Y	Y	Y	Y
N	409162	55031	37284	28791	409162	55031	37284	28791
R-Squared	0.145	0.152	0.705	0.675	0.217	0.282	0.723	0.697

Note: t-statistics in parentheses. * p<0.05, ** p<0.01, *** p<0.001. Dependent variable is Export Status (1,0) in the HMRC or FAME datasets.

Table 2. Capital Export Premia, HMRC vs FAME

	(1) Log Capital	(2) Log Capital	(3) Log Capital	(4) Log Capital	(5) Log Capital	(6) Log Capital
Export Status HMRC	2.445*** (0.090)		0.168*** (0.016)		0.1013*** (0.014)	
Export Status FAME		3.7987*** (0.0820		0.1930*** (0.021)		0.1527*** (0.020)
Log Assets			0.9150*** (0.004)	0.9207*** (0.004)	0.9246*** (0.003)	0.9267*** (0.004)
Year FE	N	N	N	N	Y	Y
Industry FE	N	N	N	N	Y	Y
N	247,324	247,324	247,323	247,323	247,323	247,323
R-Squared	0.163	0.127	0.718	0.718	0.724	0.724

Note: t-statistics in parentheses. * p<0.05, ** p<0.01, *** p<0.001. Dependent variable is Log Capital Investment. The regressor is Export Status (1,0).

Table 3. Turnover Export Premia, HMRC vs FAME

	(1) Log Turnover	(2) Log Turnover	(3) Log Turnover	(4) Log Turnover	(5) Log Turnover	(6) Log Turnover
Export Status HMRC	3.7092*** (0.0174)		0.3973*** (0.036)		0.4191*** (0.034)	
Export Status FAME		3.4730*** (0.140)		0.3479*** (0.024)		0.3663*** (0.023)
Log Assets			0.8735*** (0.008)	0.8836*** (0.008)	0.8664*** (0.006)	0.8756*** (0.006)
Year FE	N	N	N	N	Y	Y
Industry FE	N	N	N	N	Y	Y
N	80,427	80,427	79,510	79,510	79,510	79,510
R-Squared	0.307	0.224	0.874	0.874	0.891	0.890

Note: t-statistics in parentheses. * p<0.05, ** p<0.01, *** p<0.001. Dependent variable is Log Turnover. The regressor is Export Status (1,0).

Table 4. Wage Export Premia, HMRC vs FAME

	(1) Log Wage	(2) Log Wage	(3) Log Wage	(4) Log Wage	(5) Log Wage	(6) Log Wage
Export Status HMRC	2.5000*** (0.105)		0.2787*** (0.028)		0.2683*** (0.024)	
Export Status FAME		2.2681*** (0.103)		0.3495*** (0.023)		0.3457*** (0.019)
Log Assets			0.8525*** (0.009)	0.8532*** (0.009)	0.8746*** (0.010)	0.8744*** (0.009)
Year FE	N	N	N	N	Y	Y
Industry FE	N	N	N	N	Y	Y
N	55,220	55,220	55,031	55,031	55,031	55,031
R-Squared	0.204	0.151	0.819	0.820	0.843	0.844

Note: t-statistics in parentheses. * $p<0.05$, ** $p<0.01$, *** $p<0.001$. Dependent variable is the Log Wage bill for a firm. The regressor is Export Status (1,0).

Table 5. Productivity Premium Associated with Starting, Stopping and Continuing to Export
HMRC vs FAME

	(1) Δ Labor Productivity	(2) Δ Labor Productivity	(3) Δ Labor Productivity
HMRC			
Export Start	0.0119 (0.016)	0.0047 (0.016)	0.0024 (0.018)
Export Stop	-0.0672*** (0.017)	-0.0731*** (0.017)	-0.0799*** (0.019)
Export Continuing	0.0174** (0.007)	0.0078 (0.007)	0.0117 (0.008)
Log Assets		0.0057*** (0.002)	0.0085*** (0.003)
FAME			
Export Start	0.0763*** (0.016)	0.0679*** (0.017)	0.0626*** (0.017)
Export Stop	-0.0628*** (0.018)	-0.0703*** (0.018)	-0.0758*** (0.020)
Export Continuing	-0.0203*** (0.005)	-0.0288*** (0.005)	-0.0336*** (0.006)
Log Assets		0.0074*** (0.002)	0.0111*** (0.003)
Year FE	N	N	Y
Industry FE	N	N	Y
N	26,129	26,076	26,076

Note: t-statistics in parentheses. * $p<0.05$, ** $p<0.01$, *** $p<0.001$. Dependent variable is the change in labor productivity between t and $t+1$. The regressors are indicators (1,0) for whether the firm started exporting, stopped exporting or continued to export between t and $t+1$.