CPB Memorandum

CPB Netherlands Bureau for Economic Policy Analysis

Sector	:	International Economics
Unit/Project	:	S6-148 Productiviteit van bedrijven en handelsprestaties
Author(s)	:	Hugo Rojas-Romagosa
Number	:	250
Date	:	28 May 2010

Firm-level exports and productivity in The Netherlands:

Report on the SFGO database

This is a background document for the CPB Discussion Paper 143: Exports and productivity selection effects for Dutch firms. The SFGO database is a firm-level dataset for The Netherlands that was used for that study.

Here we report the procedure used to analyze the dataset, and we present the main characteristics of the data together with a series of descriptive statistics. In addition, we run panel data regressions to test for the presence of export and MNE productivity premia and dynamic regressions to test the self-selection and learning-by-exporting hypothesis.
Our results are consistent with the broad findings of other country-specific firm-level datasets. Hence, we find that there are statistically significant export and MNE productivity premia, there is empirical support to the self-selection hypothesis but no evidence for the learning-by-exporting hypothesis.

1 Characteristics of the Database

This database includes only large firms.[1] It has a set of identifying variables and then balance sheet data. The SFGO is complemented by the SFKO, which is a sample of small firms.[2] These characteristics of SFGO and SFKO deter the analysis of the whole population of firms. However, it still provides useful information that can be compared with other studies from the heterogeneous firms literature.

To process the SFGO we built a series of STATA do-files. Originally, there was a database by year, from 1997 to 2005, but we aggregated the data into a single master file. From there we work with successive STATA do-files to process the data from the balance sheet, and the loss and profit statements, we estimate three different productivity variables, capital stocks and depreciation rates, create summary information by firm and analyze the main variables. Finally, we merge the SFGO with the ABR database and run the panel regressions. The description of these files is given in the Appendix.

1.1 Information included

Each observation has the following identifiers:

- FIRM_ID: onderneming identificatie
- TOP_ORG: juridische eenheid van hoogste moeder van de onderneming
- YEAR: statistiekjaar
- SBI_ID: 2-digit industry code
- MNE_ID: onderneming met buitenlandse dochtermaatschappijen (modified later)
- MNE_SUBS: deelnemingspercentage van buitenlandse onderneming (modified later)
- LAND: land from ownership (we adjust this variable later)
- UBO: ultimate beneficial owner
- EXPANSION_FACTOR: used to correct for non-respondent firms.
- F_START: if the firm started operations that year
- F_STOP: if the firm closed during that year

The main information of the database is a detailed account of the balance sheet and other accounting statements. There is data on the balance sheet (assets, liabilities and equity), profit and loss statements (information on sales, exports, production costs, profits and taxes), and balance sheet changes.

[1] The selection criteria for large firms is to have a balance sheet total asset value of more than €23 million.
[2] The SFKO, however, has less balance sheet information and no expansion factor variable to make the database representative of the whole population of small firms.

1.2 Number of observations and firms

It is important to distinguish the number of observations from the number of firms. An observation is the data of a firm for a given year. Thus, firms may have several observations for different years.

Due to the presence of non-responses to the questionnaires, the CBS includes an expansion factor correction so the SFGO is representative of all big firms. There are 2440 different firms in the database.

Using the information from the EXPANSION_FACTOR we can deduce the total number of firms in the sample (universe) and the percentage of firms represented in the SFGO. In total, 10.470 observations have an expansion factor greater than one (81% of total observations).

Table 1.1 Number of firms and coverage

Year	Firms in SFGO	Total firms	SFGO coverage
1997	1,333	1,528	87.2%
1998	1,406	1,619	86.8%
1999	1,462	1,755	83.3%
2000	1,502	1,832	82.0%
2001	1,684	1,839	91.5%
2002	1,597	1,762	90.7%
2003	1,386	1,691	82.0%
2004	1,308	1,640	79.8%
2005	1,245	1,673	74.4%
Total	12,923	15,340	84.2%

Table 1.2 Number of observations per firm

Number of years	Number of firms	Frequency
1	320	2.5%
2	488	3.8%
3	726	5.6%
4	732	5.7%
5	1,250	9.7%
6	1,446	11.2%
7	1,624	12.6%
8	1,720	13.3%
9	4,617	35.7%
Total	12,923	100.0%

In Table 1.2 we present the information on the frequency of firm observation. The majority of firms have information for most years. Around 50% of firms have 8 or more observations, while only less than 30% of firms have less than 60 observations. This provides a good starting point for panel analysis.

1.3 Distribution of firms by size

To measure the size of a firm, we use the standardised codes from the EUROSTAT and classify firms by the number of employees. From Table 1.3 we see that 98% of the firms in the SFGO have 20 or more employees. While roughly one third of the firms are concentrated in the size code 8: 200 to 499 employees.

Since only 7% of the observations have less than 50 employees, SFGO is roughly representative of firms with 50 or more employees.

Table 1.3 Number of firms by size, measured by number of employees

Eurostat Size code	Number of employees	observations	Percentage	ISGEP size code	Percentage
0	0	2	0.0	1	7.0
10	1	32	0.3	1	
21	2	14	0.1	1	
22	3-4	29	0.2	1	
30	5-9	44	0.3	1	
40	10-19	145	1.1	1	
50	20-49	643	5.0	1	
60	50-99	1,090	8.4	2	34.2
71	100-149	1,238	9.6	2	
72	150-199	1,088	8.4	2	
81	200-249	1,003	7.8	2	
82	250-499	3,135	24.3	3	24.3
91	500-999	2,031	15.7	4	34.5
92	1000-1999	1,110	8.6	4	
93	2000 or more	1,319	10.2	4	
Total		12,923	100.0		100.0

1.4 Land of ownership and sectoral coverage

There where 21% of the observations without a Land identifier. We filled in the Land identifier for firms with enough information to interpolate data. There remained 3% of non-identified observations. Dutch-owned firms represent 54% of the observations, while 22% of firms are owned by US, UK and German MNEs.

Table 1.4 Land of ownership and sectoral distribution of firms

Land	Observations	Percentage	Sector (1-digit)	Observations	Percentage
Australia	16	0.1	0	120	0.9
Austria	17	0.1	1	1,216	9.4
Belgium	367	2.8	2	2,916	22.6
Bermuda	14	0.1	3	706	5.5
Canada	25	0.2	4	1,118	8.7
Taiwan	17	0.1	5	4,208	32.6
Denmark	121	0.9	6	1,135	8.8
Finland	111	0.9	7	1,116	8.6
France	480	3.7	8	32	0.2
Germany	819	6.3	9	356	2.8
Ireland	65	0.5	Total	12,923	100.0
Israel	26	0.2			
Italy	58	0.5			
Japan	400	3.1			
Luxembourg	294	2.3			
Netherlands	6,961	53.9			
Netherlands Antilles	236	1.8			
Norway	38	0.3			
Spain	30	0.2			
Sweden	193	1.5			
Switzerland	364	2.8			
United Kingdom	822	6.4			
United States	1,239	9.6			
Other	43	0.3			
Not identified	167	1.3			
Total	12,923	100.0			

2 Identifying MNEs and exporters

2.1 MNEs and FDI

The original MNE_ID variable is a string variable with two values: "J" and "N". We control for the reliability of this values with several checks: if the firm has FDI, foreign sales and/or intermediate inputs by subsidiaries.

We create the variable MNE_MIN to denote minority holdings by foreign firm. MNE_MIN=1 if 10% ≤ MNE_SUBS > 50%, and 0 otherwise. However, only 1.3% (167 observations) are by minority holdings. This can be a result of the reporting procedure, since there are many missing observations for MNE_SUBS and firms are only obliged to report when MNE_SUBS >=50%.

If MNE_SUBS>=50%, we take the firm to be a MNE subsidiary, and we transform the MNE_SUBS to 1 if subsidiary and 0 otherwise.

The final MNE identifier variable is MNE_CID, which has three values:

- 0 = local firm (23.7%).
- 1 = MNE (or Dutch MNE, 31.6% of the observations),
- 2 = MNE subsidiary (or Foreign MNE, 44.6%), and

These values already include the correction on MNE status using the FDI check. We define FDI flows as the change in non-portfolio international fixed assets and FDI stocks as the mid-period value of these assets. FDI_ST_ID=1 identifies the firms with FDI stocks (58.2% of the observations have positive FDI stocks). If FDI_stock≥100.000 euros, and MNE_ID="N" we change the status of the firm to being a MNE. Moreover, if the land of ownership is the Netherlands, MNE_CID=1, if not, MNE_CID=2. This is the only change done to the MNE identifier variable. With this adjustment, MNE status changes for 2320 observations (18% of total observations).

We create the MNE_START (171 or 7% of firms) and MNE_STOP (158 or 6% of firms) binary variable to identify firms that become or cease being a MNE. In addition, the MNE_SWCH (66 or 3% of firms) variable identifies firms that have both started and stopped being a MNE.

2.2 Exporters

We identify exporters with the variable X_ID=1, where 55% of the observations are exporters, and 45% are non-exporters. This 55% represents the export participation rate, which has been growing in time from 52% in 1997 to 57% in 2005.[3] We also estimate the export intensity, defined as the share of export in total sales. The mean of this variable is 19%, but the median is only 2%, reflecting the concentration of exports in large firms.

We also identify firms that start and stop exporting with X_START (466 or 19% of firms) and X_STOP (515 or 21% of firms) , while X_SWCH (289 or 12% of firms) reports firms with both variables present.

In Table 2.1 we combine MNE_CID with X_ID to analyze the relationship between exporters and MNEs. The most remarkable issue is the high number of Dutch MNEs (MNE_CID=1) that do not export (32%). This can be due to foreign sales being done by affiliates.[4]Foreign subsidiaries (MNE_CID=2) have a high proportion of non-exporters (37%), while roughly half (48%) of Dutch firms (local firms and Dutch MNEs) export.

Table 2.1 Combination of export and MNE identifiers

	Local firms	Dutch MNEs	Foreign MNEs	Total
Non-exporter	2,399	1,317	2,157	5,873
Exporters	666	2,771	3,613	7,050
Total	3,065	4,088	5,770	12,923

When we use the selection criteria of ISGEP (2007), we obtain a smaller sample but the information is comparable across the countries that participate in this project. Table 2.2 summarizes the information on participation rates and export intensity for three different years. We also present the export concentration rates, which are defined as the percentage of total exports that are given by the top 1, 5 and 10 per cent of exporting firms when they are ranked by their export values. These data are in accordance with the findings from other countries, where a small percentage of firms have the majority of export value. Mayer and Ottaviano (2007) refer to these small percentage of firms as “the happy few”.

[3] The reporting of exports by SFGO firms is problematic. Over the entire period 1997-2005 the reporting was as follows: Positive exports 55%, zero exports 9%, and non-available export data 36%. We do not know for sure whether the missing observations are positive or zero. Statistics Netherlands (2009) using detailed customs data for exports found for 2007 that 47% of the SFGO firms are exporters. This is substantially less than the 55% what we found. On the basis of this auxiliary information we felt confident in assuming that our observations with non-available export data were in fact zero export cases.

[4] This can also be a problem with tax havens (i.e. Dutch firms that are registered abroad for tax purposes, but have local sales). However, most of the Dutch MNEs that do not export are registered as Dutch firms (61%) or without land of ownership identification (39%) , so we cannot say if they are registered abroad.

Table 2.2 Exporter participation rate, export intensity and the share of exports for top exporters

Country / year	Participation rate	Export intensity	Share of exports		
			Top 1 per cent	Top 5 per cent	Top 10 percent
Netherlands (1997)	77.5	31.3			
Netherlands (2000)	77.0	28.6			
Netherlands (2005)	79.6	29.9			

Note: Results are for manufacturing firms with at least 20 employees. Participation rate is the percentage share of exporting firms in the total number of firms. Export intensity is the average percentage share of exports in total sales for exporting firms.
Source: SFGO and own estimations.

3 Analysis of the variables from the balance sheet

In this section we analyze the construction and statistical characteristics of the main variables obtained from SFGO.

3.1 Capital stocks and investment

We constructed variables for capital stock of fixed assets (STOCK_FA), divided by tangible and intangible assets. To avoid end-of-period extreme values (i.e. sell outs, closure) we use instead an average between begin-of-period and end-of-period values for the stocks of fixed assets. This average is called mid-period stock. In Table 3.1 we summarize the characteristics of these variables. It is important to note that only 49% of observations have positive values for intangible asset stocks, while the median value is of only 16.000 euros (compared with a mean value of 17 million euros). This suggests a strong concentration of intangible assets among firms.

Investment (INV) is defined as fixed assets purchases, nominal value sales (not book value) and valuation changes due to mergers and consolidations. This definition of INV is not the same as the change in assets between begin-of-period and end-of-period. INV does not include depreciation, value corrections and uses the nominal value of sales.

We used the Perpetual Inventory Method (PIM) from Delgado et al. (2002) to estimate capital stocks. We used a fixed depreciation rate of 5%, and the value of STOCK_FA as the initial capital stock. As price of capital we used the implicit prices of investment in the Dutch National Accounts.

However, the estimation of capital stocks using the perpetual inventory method was not satisfactory. The estimated CAP_STOCK variable performs badly. Some observations have negative values (115, or 1% of total), and depreciation rates are not meaningful (both the median and the mean are above 100%). Therefore, we prefer to use the stock of fixed assets (STOCK_FA) as our capital stock indicator. This variable has no negative values, produces meaningful depreciation rates. Moreover, it has a 97% correlation with the estimated CAP_STOCK.

3.2 Depreciation rates

Depreciation rates are calculated by observation and by firm. It is the ratio of the value of depreciation to the mid-term fixed asset stocks. We also separated depreciation rates between tangibles and intangibles assets. Stocks and depreciation values for tangible and intangible fixed assets where obtained from the balance sheet changes statement.

We have summary statistics for the depreciation rates by type of asset (all, tangible, intangible), sector (2-digit) and year.

All depreciation values are non-negative (a problem found for the plant-level data) and in Table 3.1 we present the average and median rates for total depreciation, and for intangible and tangible assets. We find that the variability of the depreciation rates is large and there are many extreme values. This is specially the case for intangible assets, which has many rates above 100%. Therefore, the median may be a better option than the mean to summarize the depreciation rate by specific groups.

Finally, the correlation rate between depreciation and the mid-term capital stock is 0.82. Since there is no capital-stock information in the plant-level data, this high correlation suggests that the use of depreciation values is a good proxy variable for capital stocks.

Table 3.1 Statistical characteristics of the main variables from the balance sheet, in thousand euros

	Observations (greater than zero)	Proportion of total	Mean	Median	Minimum value	Maximum value
FDI flows	6893	53.3%	6226	0	– 14087588	20934024
FDI stocks	7517	58.2%	87517	48	– 17194898	55942376
Investment in fixed assets	12343	95.5%	23214	3,670	– 7990746	9697000
Fixed assets (stock, mid-period)	12920	100.0%	105519	19,925	0	10924706
Intangible fixed assets	6276	48.6%	16901	16	0	2297074
Tangible fixed assets	12920	100.0%	88617	17,132	0	10101000
Depreciation rates	12766	98.8%	18.6%	14.9%	0.0%	966.2%
Depreciation rates intangibles	6022	46.6%	41.3%	22.7%	0.0%	6653.6%
Depreciation rates tangibles	12761	98.7%	19.1%	15.1%	0.0%	966.2%

4 Analysis of the variables from the profit and loss statement

We have information on revenues, costs and financial results for each observation. From the revenues information we obtain sales and export data (this export data is used to identify firms in X_ID). From the costs information we obtain data on: wages (total labour costs, including social security), depreciation values, intermediate inputs (INT_INP), net subsidies (SUBS) and operational profits. From the financial results we get financial profits and corporate income taxes (CIT). Total profits (PROFITS) are the sum of the operational and financial profits.

We estimate gross value added as: VA=DEPR_VALUE+WAGES+NET_TAX+PROFITS, where NET_TAX=CIT-SUBS. Alternatively we also calculated value added as: VA=SALES-INT_INP. We then checked the consistency between both definitions.

There is also information on number of employees (EMP), which are full time equivalents (FTE). We then combine EMP with WAGES to obtain average wages per employee: WAGE_EMP.

We construct several productivity variables. Labour productivity is calculated from value-added, as LP=VA/EMP, and from sales as LPS=SALES/EMP, where EMP is the number of full-time equivalent employees. We also constructed a TFP indicator, using the simple formulation from the Cobb-Douglas production function employed in ISGEP (2007). In particular: $\log(TFP)=\log(Y/L)+(1-\alpha_s)*\log(K/L)$, where α_s is the industry-level share of wages in value-added (Y). To estimate this share, we extracted net taxes from value added and obtain an average α_s across sections of 0.65.

However, contrary to the results in ISGEP (2007) we find rather low correlations between these productivity measures. The correlation between labour productivity obtained from sales and from value-added is of 58%, while the correlation between TFP and both labour productivity measures is below 30%.

To analyze this data we use summary statistics (mean, median, minimum and maximum values) and for some cases, histograms. We divided the information by firm, by observation and by characteristic of the firm (i.e. exporter, non-exporter).

4.1 Data organized by observation

First, we present the data by observation (mixing years and firms) for all firms in Table 4.1. Two main observations can be derived from this table. First, the distribution of these variables is skewed to the right, with a small number of large firms with extreme values, although there are also outliers to the left. All variables fit a log-normal distribution, where the median value is always below the average value.

Table 4.1 Summary statistics for main variables of the profit and losses statement, in thousand euros

	Observations (greater than zero)	Proportion of total	Mean	Median	Minimum value	Maximum value
Labour productivity (value-added)	12761	98.7%	123	68	– 3475.25	50776
Labour productivity (sales)	12922	100.0%	900	273	0	157554
TFP	12923	100.0%	25	18	0	1646
Wages by employee	12916	99.9%	47	44	0	600
Capital per worker	12920	100.0%	330	60	0	209050
Number of employees (FTE)	12921	100.0%	1001	306	0	90445
Exports	7050	54.6%	64279	1494	0	11832494
Sales	12792	99.0%	318862	88983	0	28015962
Profits	10714	82.9%	14147	2763	– 628042	4516342
Value added	12625	97.7%	74087	21136	– 336835	8669474

The same data, together with the Balance Sheet data is also presented for export and non-export firms in Table 4.2. This provides valuable information to analyze the export premia. The main characteristics of export firms are that they are larger (measured by sales and value added), have more profits and pay better salaries. However, it is not completely clear from this table that they are more productive (measured by value-added by employee) and have more employees and capital. The mean value for these variables is lower for exporters, but the median is higher. This reflects the larger amount of big firm outliers that are exporters. However, exporters are more productive when we use sales by employee and TFP. Finally, it is important to note that non-exporters have more capital per worker.

When we divide the firms between MNEs and local firms in Table 4.3, we find similar results. MNEs are more profitable, larger (when measured by value-added, sales or number of employees), pay better wages and are more productive (using our three productivity indicators). However, it is not clear that they have more capital and capital per worker is lower for MNEs than for local firms.

Table 4.2 Summary statistics, by exporting and non-exporting firms, in thousand euros

	Observations (greater than zero)	Proportion of total	Mean	Median	Minimum value	Maximum value
Labour productivity (value-added)						
Non-exporters	5807	44.9%	124	64	– 3074	21550
Exporters	6954	53.8%	122	70	– 3475	50776
Labour productivity (sales)						
Non-exporters	5872	45.4%	734	246	0	157554
Exporters	7050	54.6%	1034	294	24	123283
TFP						
Non-exporters	5873	45.4%	22	16	0	1646
Exporters	7050	54.6%	26	19	0	1352
FDI flows						
Non-exporters	1143	8.8%	1365	0	– 3403617	2669874
Exporters	2954	22.9%	10275	0	– 14087588	20934024
FDI stocks						
Non-exporters	2047	15.8%	29268	0	– 2486493	6286774
Exporters	4801	37.2%	136041	1650	– 17194898	55942376
Investment in fixed assets						
Non-exporters	5499	42.6%	22981	3581	– 3801714	9697000
Exporters	6844	53.0%	23409	3700	– 7990746	4495193
Fixed assets (stock, mid-period)						
Non-exporters	5870	45.4%	108900	20134	0	10101000
Exporters	7050	54.6%	102702	19729	1	10924706
Depreciation rates						
Non-exporters	5718	44.2%	18.4%	14.1%	0.0%	200.0%
Exporters	7048	54.5%	18.9%	15.3%	0.0%	966.2%
Wages by employee						
Non-exporters	5867	45.4%	45	42	0	600
Exporters	7049	54.5%	49	45	0	374
Capital per worker						
Non-exporters	5870	45.4%	444	61	0	209050
Exporters	7050	54.6%	237	60	0	182437
Number of employees (FTE)						
Non-exporters	5872	45.4%	1049	291	0	90445
Exporters	7049	54.5%	962	319	0	70408
Sales						
Non-exporters	5742	44.4%	231743	75119	0	11047545
Exporters	7050	54.6%	391436	102317	1207	28015962
Profits						
Non-exporters	4837	37.4%	10863	2468	– 537983	4516342
Exporters	5857	45.3%	16882	3095	– 628042	3085297
Value added						
Non-exporters	5671	43.9%	64093	19432	– 206256	8669474
Exporters	6954	53.8%	82412	22405	– 336835	5949330

Table 4.3 Summary statistics, by MNEs and non-MNEs, in thousand euros

	Observations (greater than zero)	Proportion of total	Mean	Median	Minimum value	Maximum value
Labour productivity (value-added)						
Non-MNEs	3046	23.6%	121	63	– 159	21550
MNEs	9715	75.2%	123	69	– 3475	50776
Labour productivity (sales)						
Non-MNEs	3064	23.7%	807	253	0	92110
MNEs	9858	76.3%	928	278	0	157554
TFP						
Non-MNEs	3065	23.7%	19	14	0	910
MNEs	9858	76.3%	26	19	0	1646
FDI flows						
Non-MNEs	27	0.2%	7	0	– 929	9744
MNEs	4070	31.5%	8159	0	– 14087588	20934024
FDI stocks						
Non-MNEs	84	0.7%	-2	0	– 3083	97
MNEs	6764	52.3%	114728	1522	– 17194898	55942376
Investment in fixed assets						
Non-MNEs	2916	22.6%	16264	3628	– 1632863	9697000
MNEs	9427	72.9%	25375	3675	– 7990746	6101393
Fixed assets (stock, mid-period)						
Non-MNEs	3064	23.7%	81758	20167	0	10101000
MNEs	9856	76.3%	112906	19752	0	10924706
Depreciation rates						
Non-MNEs	3020	23.4%	15.4%	12.5%	0.0%	158.7%
MNEs	9746	75.4%	19.6%	15.6%	0.0%	966.2%
Wages by employee						
Non-MNEs	3062	23.7%	42	41	0	303
MNEs	9854	76.3%	49	45	0	600
Capital per worker						
Non-MNEs	3064	23.7%	526	80	0	209050
MNEs	9856	76.3%	270	57	0	182437
Number of employees (FTE)						
Non-MNEs	3065	23.7%	548	237	1	30931
MNEs	9856	76.3%	1140	331	0	90445
Sales						
Non-MNEs	3028	23.4%	181962	66028	0	14737557
MNEs	9764	75.6%	361426	98804	0	28015962
Profits						
Non-MNEs	2676	20.7%	10072	2323	– 537983	3779430
MNEs	8038	62.2%	15413	3017	– 628042	4516342
Value added						
Non-MNEs	3007	23.3%	41677	15823	– 140097	5879196
MNEs	9618	74.4%	84163	23254	– 336835	8669474

When we plot histograms for the main variables, and divide the population between exporters and non-exporters, and between MNEs and local firms, the differences mentioned above do not seem to be significantly large. In the Appendix we show the histograms for value-added and labour productivity using value-added for exporters and MNEs, but the results are similar for the rest of the productivity indicators.

To obtain more information the export productivity premia, we run Kolmogorov-Smirnov tests on the equality of distributions between exporters and non-exporters for our three productivity indicators. For all three productivity measures, we reject that exporters have smaller productivity values than non-exporters, but we cannot reject that the productivity distributions are equal. Thus, it is not clear by looking only at the distributions that there is an export productivity premia and we need to check this using panel regressions.

4.2 Data organized by firm

We also present data at the firm-level. We estimate averages for each firm by year, when the firm was active in more than one year, otherwise it is the unique year observation for that firm. From this classification we also obtained the main data for starters, stoppers and switcher for exporting and MNEs.

Table 4.4 Summary statistics organized by firm, values are firm averages, in thousand euros

	Observations (greater than zero)	Proportion of total	Mean	Median	Minimum value	Maximum value
Labour productivity (value-added)	2393	18.5%	140	67	-988	37549
Labour productivity (sales)	2439	18.9%	917	271	1	97087
TFP	2440	18.9%	24	18	0	965
Investment in fixed assets	2116	16.4%	18591	3,366	– 1140549	3170500
FDI flows	667	5.2%	4254	0	– 2863014	4199508
FDI stocks	1263	9.8%	59085	171	– 259025	43669436
Fixed assets (stock, mid-period)	2438	18.9%	88989	17,904	0	9267152
Depreciation rates	2293	17.7%	19.0%	16.0%	0.0%	200.0%
Exports	1191	9.2%	52578	2,755	0	8807231
Wages by employee	2434	18.8%	47	4,396	0	281.5
Capital by employee	2438	18.9%	441	63	0	195743
Number of employees (FTE)	2439	18.9%	830	274	0	70408
Sales	2315	17.9%	262667	75,031	0	19108176
Profits	1802	13.9%	12453	2,262	– 302088	4516342
Value added	2263	17.5%	62472	18,084	– 90821	8669474

5 Consistency between SFGO and ABR

In Prog6.do we merged SFGO with ABR to obtain key info from ABR. The resulting database is: SFGO_Final. ABR is organized by plants using the variable BE_ID and has also information on FIRM_ID. Since ABR begins in 2000, there are many observations in SFGO without ABR data.

The structural change variable is plant (BE_ID) specific, so we cannot merge it with SFGO. This is also true for the sector identifier, since different plants may belong to different sectors and this may differ from the firm's sector. Thus, it is not possible to make a comparison of sectors between databases, without some kind of sector grouping in the ABR database.

We found small discrepancies between both datasets. Only 102 observations are in SFGO but not in ABR, this is around 1% of data after 2000. Therefore, using the ABR we can link the information con the SFGO with the plant-level dataset from the CBS Production Statistics.

With respect to the MNE identifier, there are discrepancies but they are due to the adjustments we made to the MNE identifier. The original MNE indicator is not equal to ABR in only 3.5% (446 observations), but for our corrected MNE subsidiaries ABR is different in 7.7% (995 observations) and Dutch MNEs in 15.6% (2019 observations)

6 Export productivity regressions

Following the empirical literature on firm heterogeneity we regress our three productivity measures against the dummies X_ID and MNE to obtain the productivity premia for exporters and MNEs, respectively. We also use export intensity (i.e. the rate of exports to total sales for each firm) as an explanatory variable. In particular, the estimating equation is:

$$P_i = \alpha + \beta X_i + \gamma \mathbf{C} + \varepsilon_i \qquad (6.1)$$

where P_i is the productivity indicator of firm i, which can be labour productivity defined as total sales per employee, or value added per employee or TFP. X is either the exporter or MNE identifier or the export intensity ratio. Finally, **C** is the vector of control variables, which include a combination of sector (2-digit sector code) and size (by number of employees),[5] and year. The productivity premia is defined as: $(\exp(\beta)-1)*100$.

We use both pooled and fixed-effect panel regressions, and our results are robust to the use of weights (e.g. the expansion-factor variable) and the elimination of outliers, which are defined as the 1st and 100th percentile of the distribution when observations are ordered using value-added by worker.

Table 6.1 presents our regressions when we use weights and eliminate the outliers. The coefficients are highly significant in all regressions. For labour productivity using sales and for TFP we find positive and significant export premia, which are lower than the MNE premia. For labour productivity using value-added the premia is 26% for exporters, and 20% for MNEs. This last result is surprising, and could be related to the fact that a large proportion of foreign MNE in The Netherlands only serve the local market.

In Table 6.2 we show our results when export intensity is used as the main independent variable. We use both a linear and a quadratic function. However, only the linear function is significant and thus, productivity is positively related with export intensity. This is more distinct for the case of labour productivity based on value-added, where the coefficient size is much larger than for the other two productivity indicators.

[5] We use this combination to reduce the problem of having a size variable (number of employees and size codes) both as a denominator of the dependent variable (when defined as labour productivity), and as an independent variable.

Table 6.1 **Export and MNE premia, weighted pooled regressions excluding outliers for different productivity definitions**

	Labour productivity (sales)			Labour productivity (value-added)			TFP		
Exporter	0.08	0.06		0.23	0.20		0.06	0.02	
	(0.01)***	(0.01)***		(0.02)***	(0.02)***		(0.01)***	(0.01)**	
MNE		0.09	0.10		0.13	0.18		0.16	0.16
		(0.01)***	(0.01)***		(0.02)***	(0.02)***		(0.01)***	(0.01)***
Export premia (%)	8.45	6.52		26.07	22.73		5,78	2.36	
MNE premia (%)		9.00	10.78		13.73	19.88		17.06	17.76
Observations	12,409	12,409	12,409	12,447	12,447	12,447	12,406	12,406	12,406
R-squared	0.39	0.39	0.39	0.48	0.48	0.48	0.62	0.62	0.62

Constant term and control variables (2-digit-sector*size and year) not reported.
Standard errors in brackets: * significant at 10%; ** significant at 5%; *** significant at 1%
Source: SFGO and own estimations.

Table 6.2 **Export intensity, weighted pooled regressions excluding outliers for different productivity definitions**

	Labour productivity (sales)		Labour productivity (value-added)		TFP	
Export intensity	0.12	0.06	0.52	0.57	0.11	– 0.08
	(0.02)***	(0.06)	(0.032)***	(0.09)***	(0.02)***	(0.07)
Export intensity squared		0.07		– 0.06		0.25
		(0.08)		(0.12)		(0.09)***
Observations	12,409	12,409	12,447	12,447	12,406	12,406
R-squared	0.39	0.39	0.49	0.49	0.62	0.62

Constant term and control variables (2-digit-sector*size and year) not reported.
Standard errors in brackets: * significant at 10%; ** significant at 5%; *** significant at 1%
Source: SFGO and own estimations.

In Table 6.3 we disaggregate the data by firm size, when defined by the number of employees for labour productivity based on value-added. We find that for all size classes both the exporter and MNE productivity premia are positive and significant. In accordance to our previous finding, the MNE premia is usually smaller than the export premia. For the other two productivity indicators we also find positive and significant productivity premia, but in these cases the MNE premia is higher.

Table 6.3 **Export and MNE premia by size class, pooled weighted regressions excluding outliers for labour productivity (value-added)**

	Size 1: 0 to 49 employees		Size 2: 50 to 249 employees		Size 3: 250 to 499 employees		Size 4: 500 or more employees	
Exporter	0.73		0.24		0.14		0.18	
	(0.12)***		(0.03)***		(0.03)***		(0.02)***	
MNE		0.43		0.24		0.15		0.08
		(0.15)***		(0.03)***		(0.03)***		(0.03)***
Export premia (%)	106.99		27.64		15.5		20.04	
MNE premia (%)		53.75		27.61		15.98		7.83
Total observations	781	781	4,338	4,338	3,115	3,115	4,213	4,213
R-squared	0.42	0.39	0.36	0.36	0.32	0.32	0.49	0.49

Constant term and control variables (2-digit-sector*size and year) not reported.
Standard errors in brackets: * significant at 10%; ** significant at 5%; *** significant at 1%
Source: SFGO and own estimations.

In Table 6.4 we divide the data by year and compare the export and MNE premia. Both premia are positive and significant for each year, while the export premia is higher for most of the years.

Table 6.4 **Export and MNE premia by year, pooled weighted regressions excluding outliers for labour productivity (value-added)**

	Export premia		MNE premia	
	(%)	Sign. level	(%)	Sign. level
1997	13.63	**	10.86	
1998	28.61	***	29.64	***
1999	18.29	***	20.49	***
2000	24.01	***	18.12	***
2001	25.30	***	18.85	***
2002	27.14	***	17.06	***
2003	25.48	***	22.11	***
2004	31.07	***	18.45	***
2005	23.40	***	17.37	**

Sign. level: * significant at 10%; ** significant at 5%; *** significant at 1%
Source: SFGO and own estimations.

Finally, in Table 6.5 we show the panel regressions that use firm-specific fixed-effects. As expected, the productivity premia are much lower using this specification. This is a result of unobservable firm characteristics being taken into account, which explain part of the superior performance of exporting firms.

Table 6.5 **Export and MNE premia, fixed-effects regressions excluding outliers for different productivity definitions**

	Labour productivity (sales)			Labour productivity (value-added			TFP		
Exporter	0.019	0.018		0.061	0.060		– 0.003	– 0.004	
	(0.012)	(0.012)		(0.015)**	(0.015)***		(0.013)	(0.013)	
MNE		0.043	0.045		0.041	0.045		0.044	0.044
		(0.024)*	(0.024)*		(0.019)**	(0.019)**		(0.024)*	(0.024)*
Export premia (%)	1.91	1.80		6.30	6.18		– 0.27	– 0.39	
MNE premia (%)		4.43	4.57		4.16	4.65		4.52	4.48
Number of firms	2,380	2,380	2,380	2,382	2,382	2,382	2,379	2,379	2,379
Total observations	12,409	12,409	12,409	12,447	12,447	12,447	12,406	12,406	12,406
R-squared	0.18	0.18	0.18	0.27	0.27	0.27	0.28	0.29	0.29

Constant term and control variables (2-digit-sector*size and year) not reported.

Standard errors in brackets: * significant at 10%; ** significant at 5%; *** significant at 1%

Source: SFGO and own estimations.

7 Regressions using the ISGEP methodology

To compare our results with those from ISGEP (2007), we use their standardized procedure, this is:

- Include only firms with 20 or more employees
- Only manufacturing firms (i.e. SECTOR_ID=2,3)
- Exclude outliers
- Include the following control variables: log of employees (ln_emp) and it's squared value (ln_emp2), log of real wages (ln_realwages) and dummy variables that combine year and sector at the 2-digit code.
- Use pooled data and firm-specific fixed effects.

Using this methodology we fail to find a consistent and robust productivity premia. For exporters we only find a premia for labour productivity using sales, for both the pooled and the fixed-effects regressions, and for TFP using pooled regressions. Moreover, the productivity premia using value-added is negative in the pooled regressions. For MNEs there is only a premia for labour productivity from sales and TFP for the pooled regressions. In Table 6.6 we report the regression results for the pooled regressions.

Table 7.1 **Pooled regressions with three different productivity definitions, following the ISGEP methodology.**

	Labour Productivity (value-added)	Labour Productivity (value-added)	Labour Productivity (sales)	Labour Productivity (sales)	TFP	TFP
Exporter	– 0.047		0.160		0.048	
	[0.017]***		[0.021]***		[0.024]*	
MNE		0.022		0.064		0.188
		[0.023]		[0.034]*		[0.031]***
ln_emp	– 0.415	– 0.433	– 0.728	– 0.690	0.005	-0.012
	[0.059]***	[0.059]***	[0.069]***	[0.069]***	[0.072]	[0.072]
ln_emp2	0.031	0.033	0.051	0.049	0.003	0.004
	[0.005]***	[0.005]***	[0.005]***	[0.005]***	[0.005]	[0.005]
ln_realwage	1,238	1,234	1,302	1,299	0.189	0.167
	[0.056]***	[0.055]***	[0.061]***	[0.061]***	[0.060]***	[0.060]***
year*sector_id	0.000	0.000	0.000	0.000	0.001	0.001
	[0.000]***	[0.000]***	[0.000]**	[0.000]***	[0.000]***	[0.000]***
Observations	4655	4655	4669	4669	4655	4655
R-squared	0.35	0.35	0.28	0.27	0.01	0.02

Constant not reported. Standard errors in brackets, where * is significant at 10%; ** significant at 5%; *** significant at 1%.

Given the consistent and robust productivity premia found earlier, the results using the ISGEP methodology are surprising. However, on a closer inspection we find that the main differences are related to the way in which the dummy variables are constructed and on the use of real wages as a control variable. In particular, the ISGEP methodology uses "grouped" variables, which assign a different number to different combinations of year and sector. Thus, this is not completely the same as the use of traditional dummy variables. However, the main difference is related to the way in which the control variables are combined. As explained before, we use a combination of sector and size, with years separated, while we did not use the log of employment variables. With this specification we wanted to avoid the inconvenience of having employment both in the denominator of the productivity indicator and as an independent variable. Another difference between our previous results and those using the ISGEP methodology is given by the use of real wages as a proxy for human capital. Finally, restricting our sample to manufacturing firms only does not produce different results.

8 Dynamic regressions

In this last section we show our results using dynamic regressions to test for two hypotheses: self-selection and learning-by-exporting. First, we use the ISGEP methodology that defines an export starter as a firm that exports in a given year, but did not export in the previous three years. All firms that are exporting throughout the whole period are excluded. Moreover, we use the same selection criteria as with the static regressions: include only manufacturing firms with 20 or more employees, exclude outliers and use the same set of control variables. Using this information, one can compare the productivity of the export starters with respect to the productivity of non-exporting firms three years prior to the starters actually exporting. These regressions will test the hypothesis of self-selection, where the econometric specification is the following:

$$P_{i\,t-3} = \alpha + \beta XS_{it} + \gamma \mathbf{C} + \varepsilon_i \qquad (8.1)$$

where $P_{i\,t-3}$ is the productivity indicator of firm i with a three-year lag. XSit is the dummy variable that identifies export starters at period t , and C is the vector of control variables.

For the second hypothesis, learning-by-exporting we use the same export starter identifier but now we compare the three-year productivity growth between non-exporting and export starters. Thus, the equation to be estimated is:

$$PG_{i\,t+3} = \alpha + \beta XS_{it} + \gamma \mathbf{C} + \varepsilon_i \qquad (8.2)$$

where $PG_{i\,t+3}$ is the growth of the productivity indicator of firm i in the following three years. However, using the ISGEP criteria of three-year gaps to define export starters, and limiting the sample to include only manufacturing firms yields only 11 observations to test the self-selection hypothesis, and none to test the learning-by-exporting hypothesis. Therefore, we do not report these results.

Instead, we widen the selection criteria and use a two-year gap to define an export starter, and include all the firms in the sample, regardless of size and sector. Using these criteria we obtain a sample between 2,125 and 574 observations, depending on the hypothesis tested and the control variables used.

Table 8.1 shows the regressions using equation 8.1 that test for self-selection. We find that export starters are more productive than non-exporters before they began to export. This confirms the self-selection hypothesis.

Table 8.1 Self-selection dynamic regressions for different productivity definitions

	Labour productivity (sales)	Labour productivity (value-added)	TFP
Export starter (2-year lag)	0.056	0.073	0.127
	[0.059]	[0.084]	[0.083]
Observations	991	1,001	991
R-squared	0.02	0.04	0.07

Constant term and control variables (2-digit-sector*size and year) not reported.

Standard errors in brackets: * significant at 10%; ** significant at 5%; *** significant at 1%.

Source: SFGO and own estimations.

9 Summary

We have processed and analyzed the information on the firm-level database of large Dutch firms (SFGO). Our results are consistent with the broad findings of other country-specific firm-level datasets (see for example, Wagner, 2007). Hence, we find that there are statistically significant export and MNE productivity premia, export concentration rates are high and there is empirical support to the self-selection hypothesis but no evidence for the learning-by-exporting hypothesis.

References

Delgado, M. A., J.C. Fariñas and S. Ruano, 2002, Firm productivity and export markets: A non-parametric approach, Journal of International Economics 57: 397-422.

International Study Group on Exports and Productivity, 2007, Exports and Productivity˜ Comparable Evidence for 14 Countries, ZEW Discussion Paper 07-069.

Mayer, T. and G.I. Ottaviano, 2007, The Happy Few: The Internationalisation of European Firms, vol. III of Bruegel Blueprint Series, Bruegel.

Statistics Netherlands, 2009, Internalisation Monitor 2009, The Hague.
Wagner, J., 2007, Exports and productivity: A survey of the evidence from firm-level data, World Economy, vol. 30, no. 1, pp. 60–82.

10 Appendix

Description of the successive STATA do-file that handle the SFGO files:

- Prog1: Mergers yearly data into Master_file, labels variable names (in English) and drops variables that are not used.
- Prog2: Processes the balance-sheet and loss and profit statements. Defines new variables (e.g. assets, investment, exports, sales, value added).
- Prog3: Estimates capital stocks and depreciation rates.
- Prog4: Creates summary information by firm (average and mode for main variables).
- Prog5: Analyzes the main variables obtained from the balance sheet (e.g. labour productivity, value added, exports) using summary statistics and histograms. It also distinguishes between: non-exporters, exporters and MNEs. Finally, it runs Kolmogorov-Smirnov tests.
- Prog6: Merges data with ABR and checks consistency of both. This program produces the file SFGO_Final.dta, which incorporates all the change and estimations from the previous do-files.
- Prog7: Runs panel regressions using the standardized procedure from ISGEP.
- Prog8: Runs the dynamic regressions using the ISGEP procedure.
- Prog9: Runs panel regressions using our own econometric specification.
- Prog10: Runs the dynamic regressions using our own procedure.
- ProgALL: Runs all the programs in a successive sequence.

Figure A.1 **Histogram for log of labour productivity (value-added), by exporter and non-exporter**

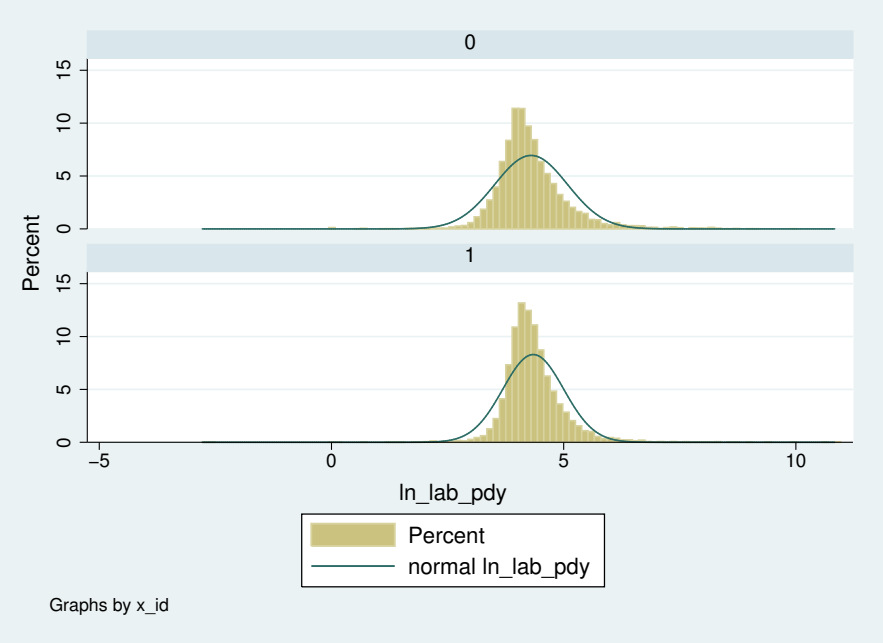

Figure A.2 **Histogram for log of labour productivity (value-added) by exporter and MNE identifier**

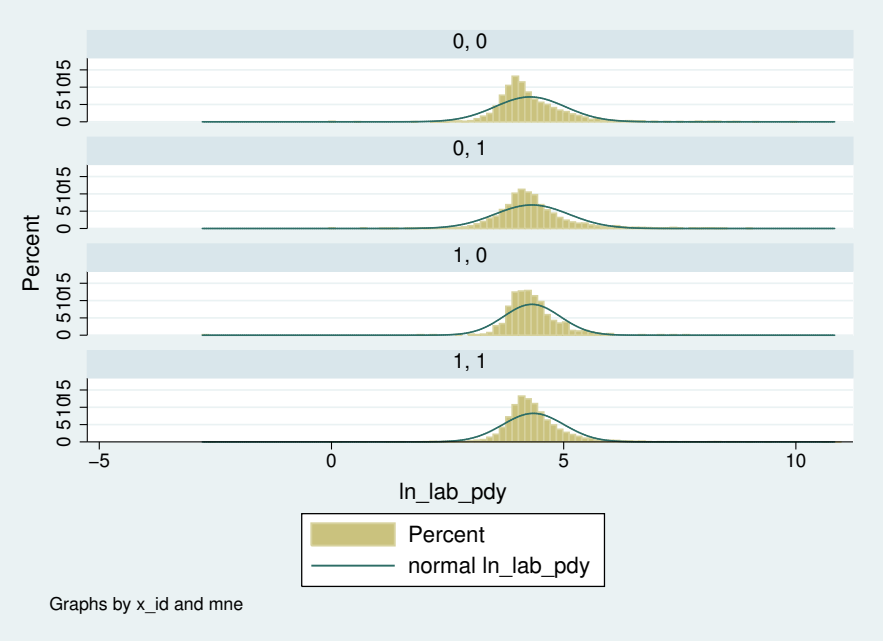

Figure A.3 **Histogram for logarithm of value added by exporter and non exporter**

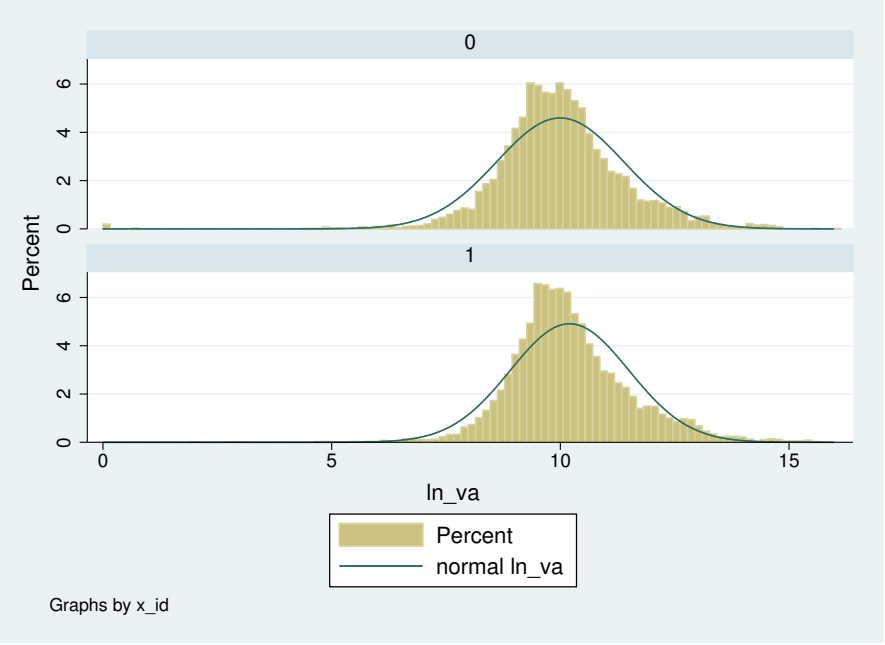

Figure A.4 **Histogram for logarithm of value added by exporter and MNE identifier**