Appendices for *Econometrica* website and not for publication

A   Constructing the French Firm Level Data

Up to 1992 all shipments of goods entering or leaving France were declared to French customs either by their owners or by authorized customs commissioners. These declarations constitute the basis of all French trade statistics. Each shipment generates a record. Each record contains the firm identifier, the SIREN, the country of origin (for imports) or destination (for exports), a product identifier (a 6-digit classification), and a date. All records are aggregated first at the monthly level. In the analysis files accessible to researchers, these records are further aggregated by year and by 3-digit product (NAP 100 classification, the equivalent of the 3-digit SIC code). Therefore, each observation is identified by a SIREN, a NAP code, a country code, an import or export code, and a year. In our analysis, we restrict attention to exporting firms in the manufacturing sector in year 1986 and in year 1992. Hence, we aggregate across manufacturing products exported. We can thus measure each firm's amount of total exports in years 1986 and 1992 by country of destination. Transactions are recorded in French Francs and reflect the amount received by the firm (i.e., including discounts, rebates, etc.). Even though our file is exhaustive, i.e., all exported goods are present, direct aggregation of all movements may differ from published trade statistics, the second being based on list prices and thus excluding rebates.

We match this file with the Base d'Analyse Longitudinale, Système Unifié de Statistiques d'Entreprises (BAL-SUSE) database, which provides firm-level information. The BAL-SUSE database is constructed from the mandatory reports of French firms to the fiscal administra-

tion. These reports are then transmitted to INSEE where the data are validated. It includes all firms subject to the "Bénéfices Industriels et Commerciaux" regime, a fiscal regime mandatory for all manufacturing firms with a turnover above 3,000,000FF in 1990 (1,000,000FF in the service sector). In 1990, these firms comprised more than 60% of the total number of firms in France while their turnover comprised more than 94% of total turnover of firms in France. Hence, the BAL-SUSE is representative of French enterprises in all sectors except the public sector.

From this source, we gather balance sheet information (total sales, total labor costs, total wage-bill, sales, value-added, total employment). Matching the Customs database and the BAL-SUSE database leaves us 229,900 firms in manufacturing (excluding construction, mining and oil industries) in 1986 with valid information on sales and exports. All values are translated into U.S. dollars at the 1986 exchange rate.

B   Comparison with Danish and Uruguayan Firms

Here we present a number of regression of the form

$$\ln \overline{X}_{ni} = \beta_0 + \beta_1 \ln X_n + \beta_2 \ln y_n + \varepsilon_{ni}$$

where $\overline{X}_{ni}$ is the mean sales in n among exporters from i, X_n is n's market size, and y_n is GDP per capita in n (taken from the World Bank's *World Development Indicators*) We have data for i = France in 1986 and 1992, for i = Denmark (from Pedersen, 2009) in 1993, and for i = Uruguay (compiled by Raul Sampognaro) averaged across 1992 and 1993. We weren't able to get y_n for every country so the regressions with that variable included have fewer

observations.

| | France | | | | Denmark | | Uruguay | |
	1986		1992		1993		1992/93	
Intercept, $\widehat{\beta}_0$	-5.1	-4.3	-4.6	-4.0	-4.7	-5.3	-5.3	-4.6
	(0.23)	(0.34)	(0.22)	(0.38)	(0.66)	(0.94)	(0.66)	(1.09)
Market size, $\widehat{\beta}_1$	0.35	0.42	0.34	0.41	0.32	0.30	0.32	0.36
	(0.02)	(0.04)	(0.02)	(0.04)	(0.05)	(0.07)	(0.06)	(0.10)
GDP per capita, $\widehat{\beta}_2$		-0.17		-0.14		0.09		-0.12
		(0.05)		(0.07)		(0.12)		(0.19)
Observations	113	100	98	94	39	37	63	61
Adjusted R^2	0.68	0.71	0.72	0.74	0.33	0.33	0.24	0.23

The estimated elasticity with respect to market size, $\widehat{\beta}_1$, is similar across the 3 sources, while the effect of y_n is not robust (significantly different from zero only for France).

By pooling the four samples we can test for a common value of β_1 and we can explore whether there is a significant home-country effect (kicking in when the source and destination are the same, which we observe for France and Denmark, but not for Uruguay). We cannot reject the hypothesis of a common slope coefficient on market size if we allow for source-specific intercepts and a home-country effect. Imposing that restriction on the pooled regression yields:

	Estimate	Standard Error
Intercept, $\widehat{\beta}_0$	-4.30	(0.27)
Source-country effects:		
France (1986)	0	dropped
France (1992)	0.35	(0.09)
Denmark (1992)	0.09	(0.13)
Uruguay (1992-1993)	-0.40	(0.11)
Home-country effect	1.60	(0.38)
Market size, $\widehat{\beta}_1$	0.38	(0.03)
GDP per capita, $\widehat{\beta}_2$	-0.13	(0.04)
Observations	292	
Adjusted R^2	0.59	

Notice that the home effect is significantly positive and with a magnitude greatly exceeding other differences by source.

C   Monte Carlo Analysis

We perform a Monte Carlo evaluation of our estimation procedure, examining its performance in general and, more particularly, the sensitivity of our parameter estimates to sampling error in N_{nF}. We use the algorithm described in Section 4.2, with the estimated parameter values reported above along with the actual values of N_{nF}, to simulate 230,423 artificial French firms (the same as the total number in our actual sample).[50] We then proceed from scratch with the artificial dataset as we did with the actual French data, calculating the moments described in Section 4.3 and the number of firms selling in each market. We apply our estimation procedure, exactly as described, to these simulated data, using the new moments and counts of firms, to estimate Θ (using the same weighting matrix $\mathbf{W}$ as in the original estimation). We performed this exercise 25 times.

The table below reports the values used to create the simulated data (the "truth"), the same as our estimate above, the mean estimate across the 25 experiments, and the standard deviation:

	$\widetilde{\theta}$	λ	σ_α	σ_η	ρ
"truth"	2.46	0.91	1.69	0.34	-0.65
mean estimate	2.47	0.94	1.71	0.34	-0.65
standard deviation	0.04	0.19	0.02	0.00	0.04

The apparent bias in the estimates is very small although sampling error in N_{nF} appears to create uncertainty about the value of λ.

---

[50] Since we want to simulate the universe of firms rather than those exporting and selling in France, step 5 of the prodecure sets $\overline{u}(s) = \max_n \left\{ \overline{u}_n(s) \right\}$.

D   Robustness Checks

The table below shows how our parameter estimates differ when we employ different sets of moments in the estimation procedure. The first row repeats the baseline estimates described in the paper. The second row shows the results of dropping all moments that are based on the 95th percentiles (in the second and third set of moments listed in Section 4.3). The third row drops the 95th percentiles but also adds the 25th percentiles (to the second, third, and fourth set of moments). The fourth row employs the moments constructed as in the baseline but dropping the 56 countries below the median in terms of their popularity to French exporters (for the second, third, and fourth set of moments).

	$\widetilde{\theta}$	λ	σ_α	σ_η	ρ
Baseline moments	2.46	0.91	1.69	0.34	-0.65
Moments based on 50th and 75th percentile	2.57	0.75	1.69	0.34	-0.65
Moments based on 25th, 50th, and 75th percentile	2.46	0.56	1.47	0.36	-0.51
Only most popular destinations	2.57	0.63	1.68	0.34	-0.64

Note that the parameters are generally quite robust to these alternative moments, although λ (the parameter with the largest standard error in our baseline estimates) does vary considerably.

Our analysis pools all manufacturing firms. To what extent do the patterns on which our moments are based vary across different types of firms? To start to answer this question we have classified our firms into consumption or investment goods producers, and replicated Figures 1, 3, and 4 for each group, available on request.[51] The patterns are very similar although some intriguing differences emerge. For producers of consumer goods, for example,

---

[51]We classify firms with NAP (nomenclature d'activités et de produits) codes 22-34 as investment good producers.

sales in France tend to be smaller, conditioning on number of destinations and popularity of destination. An interesting question is the extent to which our framework can account for these by treating the two types of firms as representing different regions of the distribution of the u's, α's and η's disproportionately, or whether accommodating these differences requires a richer parameterization. We leave this question for future research.

E Deriving the General Equilibrium Equations

To derive (42) we write country i's total absorption of manufactures is the sum of final demand and use as intermediates as:

$$X_i = \gamma(Y_i^A + D_i^A) + [(1 - \beta)(\sigma - 1)/\sigma]\, Y_i, \tag{48}$$

where Y_i^A is GDP and D_i^A the trade deficit. To relate Y_i^A to wages we write:

$$Y_i^A = W_i L_i + \Pi_i, \tag{49}$$

where Π_i are total net profits earned by country i's manufacturing producers from their sales at home and abroad.

Net profits earned in destination n both by domestic firms and by exporters selling there, which we denote Π_n^D, are gross profits X_n/σ less total entry costs incurred there, $\overline{E}_n$. Using (21) for $\overline{E}_n$:

$$\Pi_n^D = \frac{(\sigma - 1)}{\sigma \theta} X_n.$$

Producers from country i earn a share π_{ni} of these profits. Hence:

$$\Pi_i = \sum_{n=1}^{N} \pi_{ni} \Pi_n^D = \frac{(\sigma - 1)}{\sigma \theta} Y_i, \tag{50}$$

where the second equality comes from applying the conditions (41) for equilibrium in the market for manufactures.

Substituting (49) into (48) and using the fact that gross manufacturing production Y_i is gross manufacturing absorption X_i less the manufacturing trade deficit D_i:

$$Y_i + D_i = \gamma \left[W_i L_i + \frac{(\sigma - 1)}{\sigma \theta} Y_i + D_i^A \right] + \frac{(1 - \beta)(\sigma - 1)}{\sigma} Y_i.$$

Solving for Y_i yields expression (42).

F   The Data for Counterfactuals

For each country n, data on GDP Y_n^A and the trade deficit in goods and services D_n^A are from the United Nations Statistics Division (2007).[52] We took total absorption of manufactures X_n from our earlier work, EKK (2004). Bilateral trade in manufactures is from Feenstra, Lipsey, and Bowen (1997). Starting with the file WBEA86.ASC, we aggregate across all manufacturing industries. Given these trade flows $\pi_{ni} X_n$ we calculate the share of exporter i in n's purchases π_{ni} and manufacturing trade deficits D_n. The home shares π_{ii} are residuals.

The shares of manufactures in final output γ_n are calibrated to achieve consistency between our observations for the aggregate economy and the manufacturing sector.[53] In particular:

$$\gamma_n = \frac{X_n - (1 - \beta)(1 - 1/\sigma)(X_n - D_n)}{Y_n^A + D_n^A}.$$

---

[52] A couple of observations were missing from the data available on line. To separate GDP between East and West Germany, we went to the 1992 hardcopy. For the USSR and Czechoslovakia, we set the trade deficit in goods and services equal to the trade deficit in manufactures.

[53] The value of γ_n lies in the interval $(0.15, 0.55)$ for 100 or our 113 countries. The average is 0.36.

We exploit (49), (50), and (42) in order to derive baseline labor Y_i^L from data on GDP and deficits:

$$Y_i^L = \frac{[1 + (\sigma - 1)(\beta - \gamma_n/\theta)] Y_i^A - [(\sigma - 1)\gamma_n/\theta] D_i^A + [(\sigma - 1)/\theta] D_i}{1 + (\sigma - 1)\beta}.$$

All data are for 1986, translated into U.S. dollars at the 1986 exchange rate. See DEK (2008) for further details.