

NBER-ORBIS Firm Level Data Initiative

Pierre-Olivier Gourinchas and Şebnem Kalemli-Özcan

NBER IFM Data Sources, July 2019

PRELIMINARY

- 1 Description of the ORBIS database: benefits and pitfalls
- 2 Some recent papers using the data
- 3 The NBER-ORBIS licence initiative

- ORBIS database provided by Bureau van Dijk (BvD) (2005–), harmonized worldwide (310 million companies in 100+ countries).
- Collected from official business registers, annual reports, newswires, and webpages (160 info providers).

Main features of the data:

- 1 Balance sheets and profit and loss account (advantage over Census).
- 2 Good coverage relative to Census iff one merges across different vintages of data or use ORBIS historical database.
- 3 Coverage is based on country regulations in terms of small private firms filing to business registers
- 4 Covers many of small and private firms (advantage over Compustat/Worldscope). Listed firms are 1 percent of the sample.
- 5 Detailed industry classification (4-digit)
- 6 Mimics official size distribution for most countries with good coverage

Coverage Relative to Eurostat (Manufacturing Wage Bill) (from Kalemli-Ozcan et al. 2015)

	Spain	Italy	Portugal	Germany	France	Norway	Romania
1999	0.69	0.59					
2000	0.71	0.63			0.70	0.54	
2001	0.73	0.62			0.72	0.68	
2002	0.75	0.69			0.75	0.67	
2003	0.74	0.68			0.73	0.73	
2004	0.75	0.71			0.71	0.66	0.85
2005	0.74	0.72			0.71	0.67	0.85
2006	0.74	0.73	0.91	0.34	0.72	0.71	0.75
2007	0.74	0.73	0.94	0.34	0.73	0.73	0.83
2008	0.72	0.84	0.97	0.28	N/A	0.65	0.81
2009	0.72	0.81	0.96	0.28	0.71	0.85	0.92
2010	0.73	0.83	0.96	0.30	0.73	0.82	0.90
2011	0.74	0.86	0.97	0.28	0.75	0.82	0.72
2012	0.71	0.85	0.96	0.25	0.73	0.87	0.74

Share of Total Wage Bill by Size Class (2006)

[▶ other variables](#)[▶ compnet comparison](#)

	Spain	Italy	Port	Germany	France	Norway	Romania
--	-------	-------	------	---------	--------	--------	---------

ORBIS-AMADEUS

1-19 emp	0.19	0.11	0.18	0.01	0.08	0.14	0.06
20-249 emp	0.47	0.53	0.50	0.33	0.30	0.43	0.32
250+ emp	0.34	0.36	0.32	0.67	0.61	0.43	0.61

Eurostat (SBS)

1-19 emp	0.20	0.22	0.21	0.07	0.14	0.15	0.07
20-249 emp	0.43	0.44	0.49	0.26	0.31	0.41	0.33
250+ emp	0.37	0.34	0.30	0.67	0.55	0.44	0.61

Other Countries, Average Coverage 2000-2016,
Manufacturing: **Historical Product of ORBIS**, Most
Restrictive—TFP sample

AT	0.34	IL	0.32
BE	0.65	JP	0.48
CH	0.12	KR	0.63
CL	0.40	LU	0.32
CO	0.66	LV	0.57
CZ	0.46	MX	0.13
DK	0.71	NL	0.49
EE	0.65	PL	0.26
FI	0.56	SK	0.32
FR	0.53	SI	0.56
GB	0.64	SE	1.06
GR	0.42	US	0.27
HU	0.49	BG	0.57
IS	0.46	IE	0.90

Coverage Relative to Eurostat Gross Output: Aggregate Economy (from Kalemli-Ozcan et al. 2015)

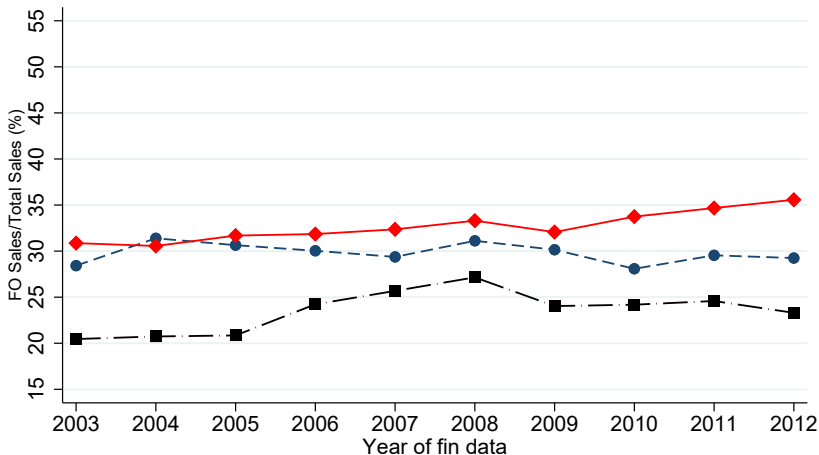
Year	AT	BE	BG	CZ	DE	DK	EE	ES	FI	FR	GB	GR	HR	HU
1999	0.21	0.60		0.39	0.29	0.04	0.76	0.63	0.35	0.54	0.46	0.95		0.79
2000	0.31	0.65		0.46	0.46	0.23	0.85	0.64	0.40	0.57	0.45	0.93		0.84
2001	0.47	0.63		0.56	0.49	0.29	0.87	0.67	0.44	0.65	0.46			0.71
2002	0.54		0.57	0.65	0.58	0.29	0.88	0.67	0.47	0.62	0.48	0.24		0.73
2003	0.49	0.63	0.60	0.67	0.56	0.29	0.93	0.68	0.50	0.67	0.48	0.47		0.66
2004	0.47	0.63	0.79	0.77	0.64	0.35	0.92	0.68	0.51	0.70	0.48	0.48		0.76
2005	0.47	0.63	0.83	0.79	0.81	0.34	0.94	0.69	0.53	0.69	0.51	0.45		0.80
2006	0.61	0.62	0.84	0.75	0.68	0.35	0.99	0.71	0.51	0.68	0.53	0.47		0.81
2007	0.68	0.62	0.91	0.79	0.76	0.35	0.95	0.68	0.55	0.70	0.48	0.47		0.79
2008	0.72	0.73	0.94	0.75	0.68	0.40	0.78	0.80	0.57	0.79	0.64	0.54	0.88	0.75
2009	0.69	0.69	0.92	0.79	0.66	0.38	0.89	0.80	0.53	0.76	0.78	0.46	0.84	0.87
2010	0.71	0.65	0.96	0.93	0.65	0.37	0.91	0.80	0.59	0.80	0.68	0.45	0.00	0.76
2011	0.73	0.74	0.88	0.85	0.61	0.42	0.93	0.82	0.58	0.81	0.83	0.45	0.82	0.72
2012	0.75	0.78	0.88	0.79	0.47	0.49	0.96	0.81	0.51	0.81	0.83	0.44	0.81	0.82
Average	0.56	0.66	0.83	0.71	0.60	0.33	0.90	0.72	0.50	0.70	0.58	0.52	0.67	0.77
Year	IE	IT	LT	LU	LV	NL	NO	PL	PT	RO	SE	SI	SK	
1999	0.13	0.45	0.16		0.54	0.17	0.63	0.36	0.45	0.58	0.52	0.29	0.26	
2000	0.32	0.50	0.18		0.51	0.21	0.63	0.40	0.49	0.53	0.56	0.66	0.33	
2001	0.36	0.50	0.19		0.58	0.20	0.77	0.47	0.48	0.68	0.60	0.66	0.40	
2002	0.32	0.54	0.39		0.62	0.22	0.79	0.44	0.59	0.70	0.63	0.77	0.51	
2003	0.35	0.53	0.53		0.63	0.21	0.65	0.47	0.60	0.75	0.65	0.75	0.75	
2004	0.36	0.57	0.52	0.63	0.84	0.22	0.67	0.57	0.63	0.83	0.67	0.74	0.77	
2005	0.34	0.57	0.54	0.59	0.87	0.23	0.59	0.54	0.80	0.82	0.68	0.74	0.82	
2006	0.44	0.58	0.59	0.48	0.87	0.25	0.67	0.61	0.82	0.78	0.71	0.74	0.79	
2007	0.44	0.60	0.78	0.43	0.86	0.25	0.71	0.65	0.83	0.82	0.70	0.73	0.78	
2008	0.49	0.72	0.62	0.44	0.88	0.28	0.59	0.55	0.82	0.86	0.73	0.80	0.97	
2009	0.46	0.68	0.60	0.28	0.80	0.31	0.78	0.66	0.83	0.92	0.84	0.80	0.93	
2010	0.59	0.64	0.69	0.34	0.79	0.35	0.79	0.61	0.86	0.92	0.88	0.95	0.92	
2011	0.54	0.71	0.57	0.80	0.82	0.30	0.72	0.56	0.81	0.92	0.82	0.84	0.95	
2012	0.40	0.61	0.48	0.72	0.78	0.28	0.67	0.59	0.93	0.91	0.76	0.96	0.91	
Average	0.40	0.59	0.49	0.52	0.74	0.25	0.69	0.54	0.71	0.79	0.70	0.74	0.72	

Firm Level Variables: Example

- Balance Sheet: Assets, Liabilities, Equity and Memo lines (number of employees, export revenue)
 - Assets: Tangible, Intangible, Cash...
 - Liabilities: Long Term Debt, Current Liabilities (Bank Loans, Creditors)...
 - Shareholder Funds: Capital/Equity
- Profit and Loss Account: Sales, EBIT, EBITDA and Memo lines (materials, cost of employees, R&D)
- **Panel data: changes in productivity and financial conditions over time.**

- ORBIS includes over 117 million ownership/subsidiary links
- Data shows full list of shareholders and subsidiaries in time series.
- Key feature of the ownership links:
 - 1 Nationality of the Owner
 - 2 Ultimate vs Direct
 - 3 Type and sector of Owner
 - Industry-FDI: Parent company industrial.
 - Financial-FDI: Parent company is a bank, financial company, insurance company, mutual fund other financial institution.

Comparison to Official Multinational Activity from OECD (from Kalemli-Ozcan et al. 2015)



- Total Economy: ORBIS Micro Data
- - ● - - Manufacturing: ORBIS Micro Data
- ◆— Manufacturing: OECD Aggregate Data

The download is not trivial from WEB based access. Several reasons for this:

Sources: WRDS (Wharton), BvD Web

The download is not trivial from WEB based access. Several reasons for this:

Sources: WRDS (Wharton), BvD Web

- AMADEUS provide most recent 10 years for the same company
- AMADEUS drops the firm if firm did not report anything in the last 6 years where ORBIS keeps the firm and if firm exits, turn status to inactive
- Reporting lag of 2 years so coverage of 2014, 2015 from 2016 download will be bad but will be good from 2018 download
- Capping: Online (WRDS, BvD) caps number of firms/observations that can be downloaded in certain period of time: researcher will not notice this since download will not crash but come as missing observations
- Update: For the same firm, same variable that is missing in 2012, can be there in 2018 since info is updated constantly.
- **To build a panel of maximum coverage over time one has to use Historical Product of ORBIS (Historical Database) that takes care of these issues and creates a panel incorporating sector and ID changes**

Table: ORBIS-AMADEUS TO WRDS, 2006, MANUFACTURING, SPAIN

	Several Vintages	Direct WRDS Download
Panel B: Wage Bill		
1-19	19.2%	0.6%
20-249	47.0%	44.4%
250+	33.8%	53.4%
Panel C: Output		
1-19	13.6%	1.2%
20-249	41.5%	42.6%
250+	44.8%	54.4%

IN: International Trade, International Macro-Finance, IO-Growth, Finance

- Elhanan Helpman, Marc J. Melitz, Stephen R. Yeaple, 2004. “Export Versus FDI with Heterogeneous Firms,” **American Economic Review.**
- Julian di Giovanni, Andrei A. Levchenko, 2012. “Country Size, International Trade, and Aggregate Fluctuations in Granular Economies,” **Journal of Political Economy.**
- Nicholas Bloom, Raffaella Sadun, John Van Reenen, 2012. “Americans Do IT Better: US Multinationals and the Productivity Miracle,” **American Economic Review.**
- Gita Gopinath, Sebnem Kalemli-Ozcan, Loukas Karabarbounis, Carolina Villegas-Sanchez, 2017. “Capital Allocation and Productivity in South Europe,” **The Quarterly Journal of Economics.**
- Javier Cravino, Andrei A. Levchenko, 2017. “Multinational Firms and International Business Cycle Transmission,” **The Quarterly Journal of Economics.**
- Elias Papaioannou and Greg Aminadav, 2019, forthcoming “Corporate Control around the World,” **Journal of Finance**

The Proposed NBER-ORBIS Data Initiative

- Initial assessment: Access to the data is limited
 - Few university subscribers
 - Main access modes not designed for economic research
- Benefits from access to raw and curated, centrally maintained data for economic research could be high.
- Historical product is very expensive (hundreds of \$/year).
Solution: Consortium model for a limited number of individual NBER researchers.
- In Sep. 2018, survey of NBER researchers in IFM, ME, EFG, IO, CF.
- Survey asked for willingness to contribute to data purchase.
- 59 researchers indicated a positive willingness to pay, ranging from less than \$4,500 to more than \$20,000.
- The potential is there.

- Expected # contributions: 20 to 25.
- Pricing: GoFundMe model! Expected to cover licensing and operational costs for the NBER.
- Sequence:
 - Step 1 Between now and end of August: solicit *firm* commitments from individual NBER researchers (research funds, grants...).
 - Step 2 Provide list of potential research projects to BvD.
 - Step 2 August: if enough commitments, NBER enters agreement with BvD for access to data.

- **Individual projects grouped under three ‘Umbrella Projects’, tentatively:**
 - Umbrella 1: Firms in the global economy
 - Umbrella 2: Firm financing and growth
 - Umbrella 3: International firm linkages
- **License agreement valid for one year.** Can be renewed, but access to data is maintained until completion for projects initiated during the year.
- **Access is only for the NBER researcher and his/her research assistants.**
- **Data hosted on NBER servers with secure access**

Advantages:

- Data set has an advantage over “Compustat type” data given the non listed firm coverage
- Data set also has an advantage over “Census type” data given the links between real and financial side of firm activity

The dataset is ideal:

- If we want to connect real and financial variables at the firm level, going outside the realm of listed firms and achieving national representation
 - Alternative: collect the data country by country via national central banks or statistical offices
- If we want to measure international flows of FDI and portfolio equity investment correctly in terms of direct and ultimate investors/risk bearers.