Firm-level heterogeneity

- Technology and costs: revisiting Tech. gap theories
- What do firms know? What do they produce?
- Measuring heterogeneity

Marco Grazzi

University of Bologna, visiting ND on the Indiana University BCSP

Notre Dame University, April 14^{th} 2015

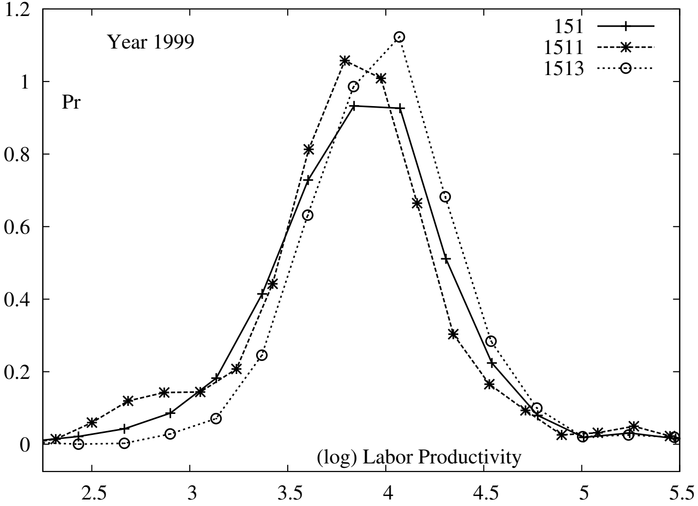
New Perspectives on firm level heterogeneity

- Recent works in Int. Econ focus on the role of firms and acknowledges substantial heterogeneity among firms
- Export status is yet another dimension under which firms differ and contributes to observed industry level heterogeneity
- Theoretical explanations (Melitz, 2003) capture the central role of produtivity differences in determining reallocation of market shares, but are rather silent about the determinants of such (persistent) productivity differences

In this scenario we investigate:

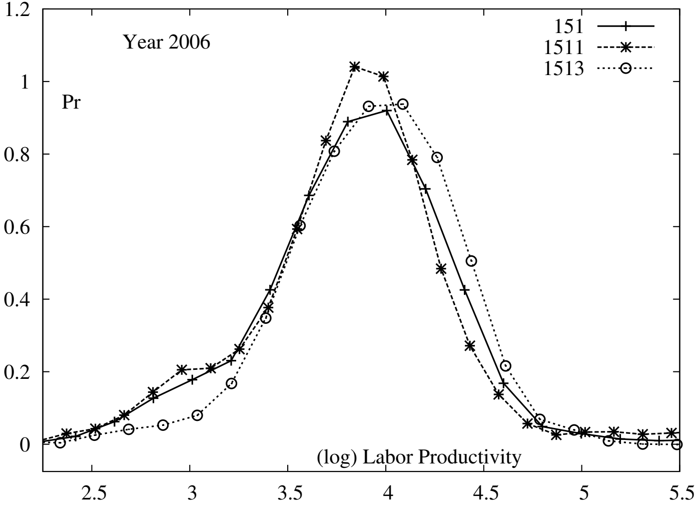
- The role of technological and cost competition in explaining who is trading (ext. margin) and export volumes (int. margin)
- What is the role of the underlying technological competences ($\sim$patents) in shaping firms' diversification patterns ($\sim$products)?
- How to measure firm-level heterogeneity (multi-dim)

Heterog. performances Meat Products (1999)



exp(3) ≈ 20 th. euro; exp(4.5) ≈ 90 th. euro

Heterog. in performances is persistent (year 2006)



$\exp(3) \approx 20$ th. euro; $\exp(4.5) \approx 90$ th. euro

Dataset

Sources and Variables (Italy)

- Firm level data. Source: SBS-like dataset (ISTAT).
    - Firms $\geq$ 20 empl (Micro.3)
    - Balance Sheet and Income Statement data
- Exports data. Source: customs trade data (ISTAT).
    - Universe of exporters
    - Firm-product-country, per year
- Patent data. Source: USPTO & EPO (Bureau van Dijk)
    - Universe of patent holders

Other countries available

- France (other works)
- China and India (only firm-level exports, no product-country disaggregation)
- Brazil (in progress)

The international competitiveness: a macro perspective

- Tech. gap theories look at sector-country trade
- Tech. gap: trade flows are primarily driven by sector-specific abs. adv., in turn stemming from widespread tech. asymmetries between countries
- Theoretical framework

$$X_{ij} = f(T_{ij}, C_{ij})$$

- Evidence at country and sector-country level
  - Technology - as proxied by patents - is relevant in explaining export shares
  - The role of costs is much less clearcut

The international competitiveness: firm level analysis

- Wide and persistent intra-industry heterogeneity in:
  - ex-ante choice of input mix (Dosi and Grazzi, 2006).
  - ex-post performance (Bartelsman and Doms, 2000).
  - partecipation on the export market (Bernard et al., 2012) and innovation activity (Basile, 2001; Caldera, 2010).
- Central role of firms and firm heterog. In recent trade literature (Melitz, 2003) all is driven by efficiency parameter $\Rightarrow$ What is the distinct role of technological and cost competition in explaing trade?
- Integrating the classical technology gap approach to international trade (Soete, 1981; Fagerberg, 1988) and the more recent literature on firms in international trade (Melitz, 2003; Bernard et al., 2007).

Table: Country- and sector-level studies (see complete table online)

Authors	Years-Cntr-Sec	Methodology	Main results
Soete (1981, 1987)	1963-77 - 20 - 40	cross-sectional estim. of 4 equations in 1977	Patents (+)
Fagerberg (1988)	1961-83 - 15 - all econ	2SLS estimation of a six equations model	R&D-Patents (+), Investments (+), Costs ()
Dosi et al. (1990)	1963-77 - 20 - 40	cross-sectional analysis	Investments (+), Patents (+), Costs ()
Greenhalgh (1990)	1954-81 - 1, UK - 31	error correction model	#Innovations (+), Prices ()
Amendola et al. (1993)	1967-87 - 16 - all manuf	autoregressive-distributed lag model	Patents (+), Investments (+), Costs ()
Magnier and Toujas-Bernate (1994)	1975-87 - 5 - 20	error correction model	R&D (+), Investments (+), Prices (-)
Amable and Verspagen (1995)	1970-91 - 5 - 18	error correction model	Patents (+), Investments (+), Costs (-)
Landesmann and Pfaffermayr (1997)	1973-87 - 7 - 2	almost ideal demand system	R&D (+), Costs (-)
Wakelin (1998b)	1988 - 9 - 22	OLS estimation of pooled & sect. data	R&D (+), Patents (+), Investments (), Costs(-)
Carlin et al. (2001)	1970-92 - 14 - 12	distributed lag model	Patents (), R&D (), Investments (+), Costs (-)

Table: Firm-level studies (see complete table online)

Authors	Country	Data source	Structure	Firms
Wakelin (1998a)	UK	SPRU innov. survey	cross-section	320
Sterlacchini (1999)	Italy	field study	cross-section	143
Basile (2001)	Italy	Mediocredito surveys	panel	6000
Roper and Love (2002)	Germ. & UK	product development survey(PDS)	cross-section	1087(UK) 1190(Germ.)
Barrios et al. (2003)	Spain	ESEE survey	panel	around 2000
Beise-Zee and Rammer (2006)	Germ.	CIS	cross-section	4786
Lachenmaier and Wöß mann (2006)	Germ.	IFO innovation survey	cross-section	981
Aw et al. (2007)	Taiwan	Statistical Bureau's census and R&D survey	panel	$\sim$518 $\sim$1311
Castellani and Zanfei (2007)	Italy	CIS2 and ELIOS	cross-section	785
Harris and Li (2009)	UK	CIS3 and Annual Respondents Database	cross-section	3303

Table: Observations by manuf. sectors, COE + Micro3, year 2000

	(I)	(II)	(III)	(IV)	(V)
All manufacturing	30,599	100.00	100.00	75.87	100.00
Food, beverages, tobacco	2,049	6.70	7.75	74.33	4.80
Textiles, wearing, leather	5,379	17.58	13.70	72.91	13.94
Wood	776	2.54	1.49	66.88	0.67
Paper & printing	1,709	5.59	5.06	69.28	2.56
Coke & petroleum	108	0.35	0.90	41.67	2.61
Chemicals	1,174	3.84	6.67	91.99	10.11
Rubber & plastics	1,863	6.09	5.15	86.74	4.68
Other non-metallic	1,697	5.55	5.09	64.76	3.34
Basic metals	866	2.83	4.57	82.56	4.99
Fabricated metal	4,668	15.26	9.66	63.52	5.27
Machinery	4,433	14.49	15.22	87.95	20.70
Computing & electrical	2,681	8.76	10.41	74.67	9.93
Transport equipment	1,023	3.34	9.57	77.61	11.07
Other manufacturing	2,173	7.10	4.74	85.18	5.33

Note. (I) Number of firms; (II) percentage share of firms within each sector; (III) shares of employment; (IV) percentage of exporting firms within each sector; (V) shares of export volumes.

The Micro Evidence: Selection into export

$$P(EXP_{it}=1)=\Phi(\beta_1 WAGE_{it-1}+\beta_2 PROD_{it-1}+\beta_3 INV_{it-1} + \beta_4 PAT_{it-1}+\beta_5 EMP_{it-1}) \quad (1)$$

i for firms

$EXP_{it}=1$, if a firm exports

EMP, number of employees

PAT, dummy for patenting firm

The Micro Evidence: Export volumes

$$\begin{aligned} EXP_{it} = & \beta_1 WAGE_{it-1} + \beta_2 PROD_{it-1} + \beta_3 INV_{it-1} \\ & + \beta_4 PAT_{it-1} + \beta_5 EMP_{it-1} + \epsilon_{it} \end{aligned} \tag{2}$$

Main results

- **Patents and investments** do matter. More relevance at the *intensive* (volums) than at the *extensive* (selection) margin.
- **Wages**: capture more differential skills (even controlling for labour productivity). Overall, not a hindrance to export strategy.
- **Product innovation** is more relevant than process innovation in determining firms export success (CIS data).

Selection: results

	WAGE	PROD	INV	PAT(D)	Obs.	firms
ALL MANUFACTURING	0.034***	0.119***	0.011***	0.115***	181524	39761
Food, beverages, tobacco	−0.007	0.132***	0.009**	0.144***	14136	2941
Textiles, wearing, leather	−0.052***	0.253***	−0.017***	0.053	32356	8030
Wood	0.044	0.204***	0.010	0.206***	4854	1028
Paper & printing	−0.274***	0.131***	0.023***	0.122*	10635	2268
Chemicals	0.038*	0.014	0.004	0.025	9261	1714
Basic metals	0.105***	0.063***	0.012***	0.163***	7108	1236
Machinery	0.054***	0.070***	0.009***	0.066***	24312	5010
Computing & electrical	0.095***	0.150***	0.041***	0.114***	15294	3624
Transport equipment	0.169***	0.051***	0.012***	0.140***	5725	1244

Note. Probit estimation. Marginal effects computed at means with robust standard errors clustered at the firm level in parentheses. (D) for discrete change of dummy variable from 0 to 1. Coefficient of EMP omitted. Year dummies included. *** $p < 0.01$, ** $p < 0.05$, * $p < 0.10$.

Export volumes: results

	WAGE	PROD	INV	PAT	Obs.	firms
ALL MANUFACTURING	0.032	0.824***	0.082***	0.562***	138241	31255
Food, beverages, tobacco	0.367***	0.852***	0.157***	1.073***	9931	2310
Textiles, wearing, leather	−0.094	1.117***	−0.069***	0.799***	23326	5778
Wood	0.317	0.246	0.017	1.762***	3226	743
Paper & printing	−1.188***	0.903***	0.223***	1.389***	7249	1719
Chemicals	−0.179	0.713***	0.278***	0.277**	8153	1578
Basic metals	−0.577***	0.989***	0.051*	0.210	5743	1064
Machinery	0.105	0.858***	0.029**	0.479***	21544	4531
Computing & electrical	−0.026	0.236***	0.149***	0.722***	12056	2796
Transport equipment	0.198	0.874***	0.131***	0.987***	4680	1041

Note. Pooled OLS estimation with robust standard errors clustered at the firm level in parentheses. Coefficient of EMP omitted.
*** $p < 0.01$, ** $p < 0.05$, * $p < 0.10$.

Main results: Robustness checks

These results are robust to a number of controls and robustness checks:

- Heckman-type selection (selection variable: lagged exp status)
- workforce composition
- GMM estimation of dynamic specification with short-run and long-run effects:
    - Lab cost display some negative and sign effect in the short run; however it vanishes in the long run
    - Technology variables display quite a different pattern: Inv (both short and long); patents (long)

Short-run *vs.* long-run

- We adapt the empirical framework of Amendola et al. (1993) to firm-level data and consider an autoregressive distributed lag model

$$\begin{aligned} EXP_{it} = & \sum_{l=1}^{K} \eta_l EXP_{it-l} + \sum_{l=1}^{L} \alpha_l WAGE_{it-l} + \sum_{l=1}^{L} \beta_l PROD_{it-1} + \sum_{l=1}^{L} \gamma_l INV_{it-1} \\ & + \sum_{l=1}^{L} \delta_l PAT_{it-1} + \sum_{l=1}^{L} \phi_l EMP_{it-1} + d_t + \epsilon_{it} \end{aligned} \quad (3)$$

- In order to identify the short-run coefficients, we employ a "twostep system GMM" estimator, to control both for unobserved heterogeneity and for the potential endogeneity of cost and technology variables. (K=1; L=3)
- We use less distant lags (typically at $t-2$ and $t-3$) to instrument, in the first difference equation, both the lagged value of the dependent variable (EXP_{it-1}) and the variables that we take as endogeneous, that is wage, productivity, investment intensity, and patents.
- Long-run coefficients are calculated from the short-run:

$$x_{long-run} = \frac{\sum_{l=1}^{3} x_l}{1-\eta_1} \quad (4)$$

where $x \in \{\alpha, \beta, \gamma, \delta\}$.

CIS data: Innovation premia

- Community Innovation Survey data (Survey not census)
  - Covers *all* firms ≥ 250 and a *sample* in employment range 10-250
  - No firm level time series, other than firms ≥ 250 (which are 1100)
- Better proxy of innovation ouput
  - "The firm introduced a *product* or *process* (or both) innovation"
  - "The product innovation was relevant to the market or only to the firm"
- Useful for comparative analysis *across* European countries

$$Exp_i^D = \alpha INN_i + \beta sector_i + \epsilon_i$$

$$Exports_i = \alpha INN_i + \beta sector_i + \epsilon_i$$

where $INN = INDPT$ or $INPCS$ depending on the specification.

Innovation premia: results (%)

	CIS3 98-00		CIS4 02-04	
	(a)	(b)	(a)	(b)
Product innovation premia				
Exporting firms	14.8	10.9	13.2	9.4
Export volumes	116.2	54.7	115.1	51.3
Process innovation premia				
Exporting firms	10.0	6.4	11.7	8.3
Export volumes	80.2	22.9	84.2	25.0

Note. The table reports innovation premia, in percentage.
Columns (b) control for total employment.
All differences are significant at the 1% level.

Innovation effects: selection equations

$$
\begin{aligned}
EXP_i^D = & \alpha WAGE_i + \beta PROD_i + \gamma INPCS_i + \delta INPDT_i \\
& + \zeta BOTH_i + \phi EMP_i + \epsilon_i
\end{aligned}
$$

Only for firms that introduced a product innovation:

$$
EXP_i^D = \alpha WAGE_i + \beta PROD_i + \gamma NEWMKT_i + \epsilon_i
$$

Innovation effects: export market shares equations

$$\begin{aligned} Exports_i = & \alpha WAGE_i + \beta PROD_i + \gamma INPCS_i + \delta INPDT_i \\ & + \zeta BOTH_i + \phi EMP_i + \epsilon_i \end{aligned}$$

Only for firms that introduced a product innovation:

$$Exports_i = \alpha WAGE_i + \beta PROD_i + \gamma NEWMKT_i + \epsilon_i$$

Innovation effects: selection results

	CIS3		CIS4	
	(1)	(2)	(3)	(4)
WAGE	-0.046	-0.029	-0.005	0.028
	(0.028)	(0.020)	(0.031)	(0.026)
PROD	0.142^{***}	0.073^{***}	0.120^{***}	0.011
	(0.021)	(0.018)	(0.019)	(0.013)
INPDT	0.092^{***}		0.092^{***}	
	(0.011)		(0.011)	
INPCS	0.025^{*}		0.050^{***}	
	(0.014)		(0.012)	
BOTH	0.077^{***}		0.093^{***}	
	(0.011)		(0.011)	
NEWMKT		0.019		0.025^{**}
		(0.013)		(0.012)
N	4521	1852	3609	1185
pseudo R^2	0.183	0.185	0.172	0.159

Note. Marginal effects computed at means with robust standard error in parenthesis. Discrete change from 0 to 1 for dummy variables. Columns (1) and (2) are for CIS3 regression, while columns (3) and (4) are for CIS4 regression. Sector dummies included. *** $p < 0.01$, ** $p < 0.05$, * $p < 0.10$

Innovation effects: export shares results

	CIS3		CIS4	
	(1)	(2)	(3)	(4)
RWAGE	-0.638^{**}	-0.627^{*}	0.318	0.698^{*}
	(0.266)	(0.369)	(0.256)	(0.373)
RPROD	1.303^{***}	1.171^{***}	1.037^{***}	0.778^{***}
	(0.147)	(0.199)	(0.136)	(0.202)
INPDT	0.457^{***}		0.271^{***}	
	(0.096)		(0.126)	
INPCS	-0.019		0.075	
	(0.116)		(0.111)	
BOTH	0.291^{***}		0.342^{***}	
	(0.083)		(0.095)	
NEWMKT		0.047		-0.015
		(0.118)		(0.127)
N	3699	1680	3014	1110
R^2	0.425	0.478	0.407	0.510

Note. Robust standard error in parenthesis. Columns (1) and (2) are for CIS3 regression, while columns (3) and (4) are for CIS4 regression. Sector dummies included. *** $p < 0.01$, ** $p < 0.05$, * $p < 0.10$

Innovating firms and exogenous shocks

- Do innovating and non-innovating firms respond differently to exogenous shock? Along which margins do they adjust?
- Fluctuations in real exchange rates as measures of exogenous changes

$$RER_{ct} = ER_{ct}\frac{CPI_t}{CPI_{ct}}$$

- Extensive and intensive margins of firm's exports to a destination:

$$\ln X_{fpc} = \ln Quant_{fpc} + \ln UnitPrice_{fpc}$$

- The estimation equation:

$$\Delta \ln Y_{fpct} = c_1 + \delta_1 D_f^{BOTH} + \beta_1 \Delta \ln RER_{ct} + \gamma_1 \Delta \ln RER_{ct} * D_f^{BOTH} + d_j + \varepsilon_{ct}$$

Table: Exchange rates and firm's exports to product-country destinations

	Annual Differences					
	ln X_{fcpt}	ln X_{fcpt}	ln Q_{fcpt}	ln Q_{fcpt}	ln UV_{fcpt}	ln UV_{fcpt}
	(1)	(2)	(3)	(4)	(5)	(6)
D_{ft}^{BOTH}	0.005		0.007		-0.002	
	(0.009)		(0.009)		(0.005)	
$\Delta \ln RER_{ct}$	-0.327^{a}	-0.360^{a}	-0.308^{a}	-0.344^{a}	-0.018	-0.015
	(0.104)	(0.112)	(0.115)	(0.122)	(0.021)	(0.020)
$\Delta \ln RER_{ct} * D_{ft}^{BOTH}$	0.117^{b}	0.149^{b}	0.115^{c}	0.155^{b}	0.002	-0.005
	(0.053)	(0.062)	(0.064)	(0.075)	(0.029)	(0.030)
Country FE	Yes	Yes	Yes	Yes	Yes	Yes
Product FE	Yes	No	Yes	No	Yes	No
Firm-Product FE	No	Yes	No	Yes	No	Yes
N	329697	329697	329697	329697	329697	329697
adj. R^2	0.005	0.007	0.002	0.003	0.000	0.030

Note. Firm-product-country regressions with data on exports, quantity and unit value between 2000 & 2007. The dependent and independent variables are in annual differences. BOTH is a dummy for firms introducing both product & process innovations in CIS3 & CIS4. Robust std err clustered at country-year level in parenthesis. Year dummies incl. ^{a} $p < 0.01$, ^{b} $p < 0.05$, ^{c} $p < 0.10$

What do firms know? What do they produce? Some established evidence

- Importance of product and technological diversification for:
    - firm performance through growth opportunities, economies of scope, risk diversification (Hirsch and Lev, 1971; Montgomery, 1994; Bottazzi et al., 2001; Garcia-Vega, 2006)
    - accumulation of capabilities (Dosi, 1988; Pavitt, 1998)
- Relationship between technological knowledge and product portfolios:
    - Evidence on large firms: technological scope greater than product scope (Patel and Pavitt, 1997; Gambardella and Torrisi, 1998; Brusoni et al., 2001)
- Characteristics of diversification process:
    - Evidence of path-dependent and coherent processes (Teece et al., 1994; Bottazzi et al., 2001; Breschi et al., 2003; Bottazzi and Pirino, 2010)

What do firms know? What do they produce? Our contribution

Bringing together information on patents and products we analyze:

- Properties of the diversification breadth of both technological knowledge and product portfolios
- Scaling relation between the size of the firm and diversification of both technological knowledge and product scope
- Firms coherence in their technological and product diversification

Data sources

- Amadeus-Bureau van Dijk
  - Provides information on patent applications for more than 20,000 Italian firms (including the IPC classification code, the application date, and whether the patent has been granted or not)
- Archivio Statistico delle Imprese Attive (ASIA)
  - Census of all operating businesses: age, employment, total turnover, geographical location and main activity of the firm, 1998-2006
- Statistiche del Commercio Estero (COE) Custom data
  - Transactions level data: export values and quantity of the firm for HS6 product-country destination pairs
  - All cross-border transactions at the firm-product-country level, 2000-2007
- Resorting to Lybbert and Zolas (2014), we link IPC codes to 125 4-digits ISIC codes (Rev. 3), and HS6 codes to 145 4-digits ISIC codes (Rev. 3). Firms in our sample turn out to patent in 118 different technological fields and produce 138 different products.

Patents and firms, by period of application and patent office

Table:

Period	Total PATENTS	Total FIRMS	USPTO PATENTS	USPTO FIRMS	EPO PATENTS	EPO FIRMS
1949-1978	1,086	187	1,086	187	(Funded in 1977)	
1979-1995	8,055	1,426	3,929	863	4,126	1,168
1996-2006	**21,305**	**2,946**	**9,817**	**1,647**	**11,488**	**2,499**
2007-2014	9,340	1,948	4,871	1,006	4,469	1,550
1949-2014	39,786	4,411	19,703	2,586	20,083	3,709

Note. Number of USPTO and EPO granted patents owned by Italian firms. The period refers to the application date. Data from AMADEUS, ASIA, and COE.

Distribution of firms and patents by number of patents

Table:

	No. of patents										
	#1	#2	#3	#4	#5	#6	#7	#8	#9	#10+	Obs.
% FIRMS	41.68	19.14	9.37	6.45	4.41	3.29	1.66	1.60	1.15	11.24	2,946
% PATENTS	5.73	5.26	3.86	3.54	3.03	2.71	1.60	1.75	1.43	71.08	21,441 (*)

Source. Amadeus, ASIA, and COE, 1996-2006.
(*) There a few co-patent holders

Size, age, and product scope: patenting *vs.* non-patenting firms

$$X_i = \alpha + \beta D_{PAT_i} + \epsilon_i$$

Table:

	(1) $X_i = log(exports_i)$	(2) $X_i = log(age_i)$	(3) $X_i = log(age_i)$	(4) $X_i = log(\#products_i)$	(5) $X_i = log(\#products_i)$
D_{PAT_i}	3.092^{***}	0.320^{***}	0.189^{***}	1.003^{***}	0.371^{***}
	(0.047)	(0.016)	(0.017)	(0.015)	(0.012)
$log(exports_i)$			0.042^{***}		0.205^{***}
			(0.001)		(0.001)
N	139,360	139,360	139,360	139,360	139,360
adj. R^2	0.195	0.060	0.074	0.122	0.484
Sector dummies	Yes	Yes	Yes	Yes	Yes

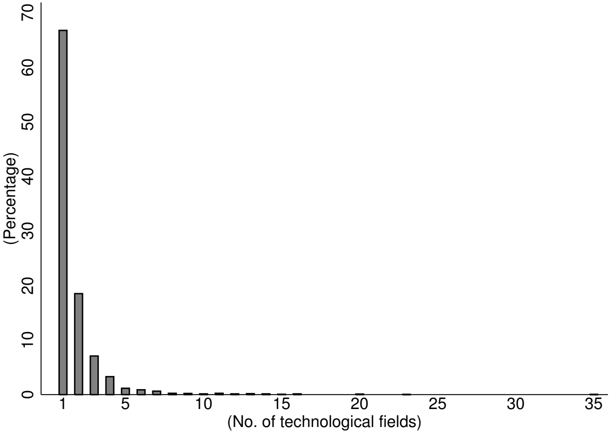
Note. Standard errors in parentheses.* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Patent and product price: patenting *vs.* non-patenting firms

- Investigate within each product-country pair the price (UnitValue) of patentees and non-patentees
- For each product-country pair we get two price distributions (1 for pat, 1 for non-pat)
- Employ Fligner-Policello (FP) test to study stochastic dominance
  - To compute the statistics we need obs > 20 both for pat and non-pat
- When the two distributions statistically differ, in (around) 80% of the product-country pair, patentees have higher price

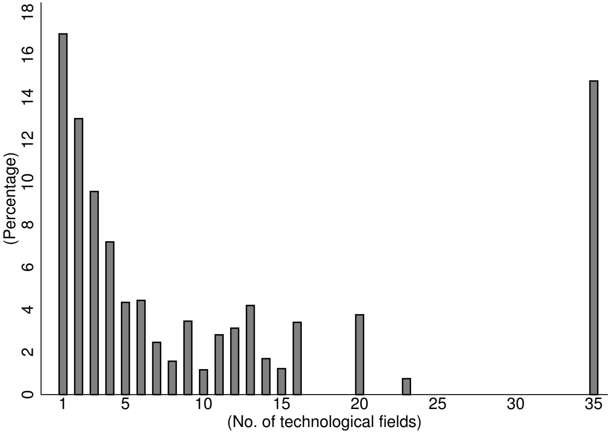
Distribution of firms by number of tech. fields

Figure:



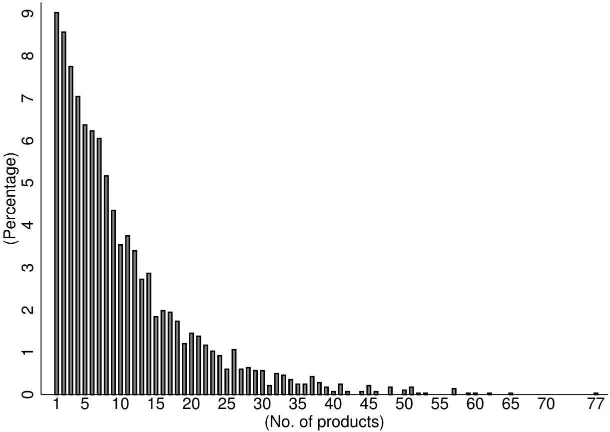
Distribution of firms' patents by number of tech. fields

Figure:



Distribution of firms by number of products

Figure:



Joint distribution of firms by # of tech. fields and products

#Tech	# Products										
fields	1	2	3	4	5	6	7	8	9	10+	Total
1	208	194	168	156	136	130	115	95	81	603	1,886
	(11.03)	(10.29)	(8.91)	(8.27)	(7.21)	(6.89)	(6.10)	(5.04)	(4.29)	(31.97)	(100.00)
2	35	38	38	31	32	27	40	31	23	228	523
	(6.69)	(7.27)	(7.27)	(5.93)	(6.12)	(5.16)	(7.65)	(5.93)	(4.40)	(43.59)	(100.00)
3	5	6	7	9	6	15	10	13	10	119	200
	(2.50)	(3.00)	(3.50)	(4.50)	(3.00)	(7.50)	(5.00)	(6.50)	(5.00)	(59.50)	(100.00)
4	4	4	3	1	3	1	4	3	6	64	93
	(4.30)	(4.30)	(3.23)	(1.08)	(3.23)	(1.08)	(4.30)	(3.23)	(6.45)	(68.82)	(100.00)
5	2	0	0	1	0	2	0	1	3	24	33
	(6.06)	(0.00)	(0.00)	(3.03)	(0.00)	(6.06)	(0.00)	(3.03)	(9.09)	(72.73)	(100.00)
6	0	0	0	0	0	0	1	1	0	23	25
	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(4.00)	(4.00)	(0.00)	(92.00)	(100.00)
7	0	0	0	1	0	0	0	1	0	16	18
	(0.00)	(0.00)	(0.00)	(5.56)	(0.00)	(0.00)	(0.00)	(5.56)	(0.00)	(88.89)	(100.00)
8	0	0	2	0	1	0	0	0	0	4	7
	(0.00)	(0.00)	(28.57)	(0.00)	(14.29)	(0.00)	(0.00)	(0.00)	(0.00)	(57.14)	(100.00)
9	0	0	0	0	0	0	1	0	0	5	6
	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(16.67)	(0.00)	(0.00)	(83.33)	(100.00)
10+	1	0	1	0	2	1	0	1	0	29	35
	(2.86)	(0.00)	(2.86)	(0.00)	(5.71)	(2.86)	(0.00)	(2.86)	(0.00)	(82.86)	(100.00)
Total	255	242	219	199	180	176	171	146	123	1,115	2,826
	(9.02)	(8.56)	(7.75)	(7.04)	(6.37)	(6.23)	(6.05)	(5.17)	(4.35)	(39.46)	(100.00)

Note. Absolute and percentage (in brackets) frequencies.

Product rank and firm knowledge

Table: Matching between technological fields and products

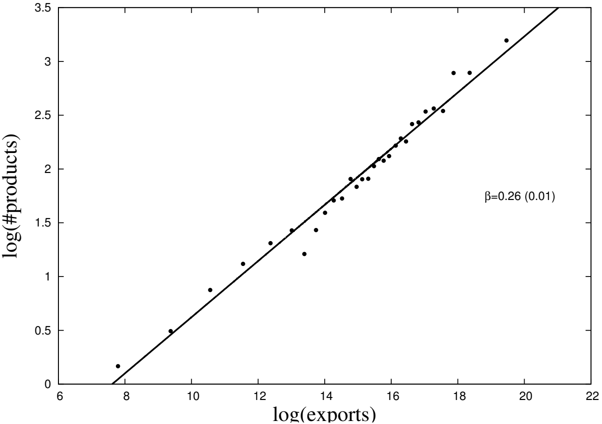
Product rank	# Products									
	1 (745)	2 (566)	3 (880)	4 (564)	5 (816)	6 (501)	7 (652)	8 (1,107)	9 (561)	10+ (13,941)
1	17.32	21.38	17.27	24.82	19.98	19.56	13.50	36.04	29.23	23.61
2		4.95	4.89	10.99	10.17	6.79	4.60	2.17	3.74	6.17
3			6.14	1.06	1.96	2.79	6.60	2.44	1.25	3.92
4				2.49	2.21	4.59	1.23	3.52	2.14	1.74
5					1.03	2.40	2.45	3.99	0.18	1.46
6						0.60	0.31	0.45	0.71	1.75
7							2.15	0.90	3.92	0.91
8								0.45	1.25	3.72
9									1.07	1.66
10+										0.81

Note. Each cell reports the percentage matching between technological fields and product categories across the relevant set of firm-products. In parentheses, the number of patents for the relevant set of firm-products.

- 1 Prod: (only) 17% of single-prod firms are active in a product in which they hold patents
- 10+ Prod: 23% of firms has as the most selling product one in which ...

Products-exports relationship

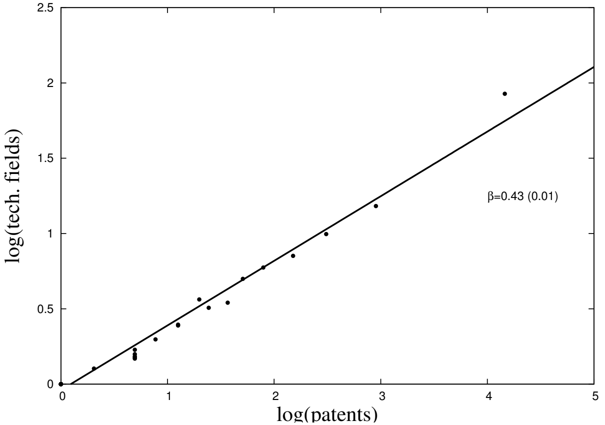
Figure: log(#products) against log(exports). Linear fit



- The increase in products and technologies we observe as firms get bigger is proportional to their growth rate.

Tech. fields-patents relationship

Figure: log(#tech. fields) against log(patents). Linear fit



- Scaling factor < 1: a large firm is less diversified than a collection of small (single product) firms which add up to the same size of the large one.

Measuring firm coherence: relatedness

We adopt a "survivor" measure of relatedness (Teece et al., 1994): if a selection mechanism is at work and if relatedness confers some advantage, then one captures that related activities will appear with higher frequency within the same firm.

Two main ingredients:

- A co-occurrences matrix C, whose generic cell J_{ij} is equal to the total number of firms active in both products:

$$J_{ij} = \sum_{k} C_{ik} C_{jk} \tag{5}$$

  with $C_{ik} = 1$ if firm k is active in product (or technological field) i, and 0 otherwise

- The p-value of the generic cell of the observed J_{ij}:

$$p_{ij}(J, H) = Prob[\tilde{J}_{ij} \leq J_{ij} | H] \tag{6}$$

  where $\tilde{J}_{ij}$ is the value of the relative cell under the null hypothesis H

Measuring firm coherence: null hypothesis

- Standard null hypothesis: the total number of firms active in a given sector (or product or patent class) is fixed and equal to the one observed in the actual data
  - the probability to obtain a given value of $\tilde{J}_{ij}$ is distributed according to a hypergeometric random variable
  - the implied distribution of firm scope converges to a binomial
- Alternative null hypothesis (Bottazzi and Pirino, 2010): both firms scope and the number of firms per activities are fixed and correspond to the observed ones
  - Deriving the implied distribution using Monte Carlo tecniques

Measuring firm coherence: formulas

- Weighted Average Relatedness measures the inverse of the average distance from a firm activity to *all* other activities:

$$WAR_k(H) = \frac{1}{n} \sum_i C_{ik} \left(\frac{\sum_{j \neq i} p_{ij}(H) w_{jk}}{\sum_{j \neq i} w_{jk}} \right) \tag{7}$$

where n is total number of products (technological fields) in which a firms is active and w_{jk} the weight of product (technological field) j with respect to firm k. We weight products with export share and technological fields with patent count

- Weighted average relatedness of neighbors measures the inverse of the average distance from a firm activity *to its neighbor activity*

$$WARN_k(H) = \frac{1}{n} \sum_i C_{ik} \left(\frac{\sum_{j \neq i} p_{ij}(H) m_{ij}^k w_{jk}}{\sum_{j \neq i} m_{ij}^k w_{jk}} \right) \tag{8}$$

where $m_{ij}^k = 1$ if the pair ij is in the maximum spanning tree of firm k, defined as the graph with $n - 1$ links such that the sum of the relatedness measures on each link is largest

Summary of results

Interpretation

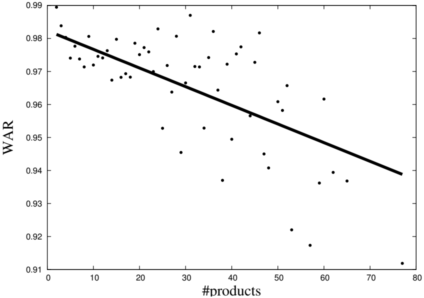
- WAR and WARN capture two different aspects of firm diversification.
- As firms diversify in new products and technological fields, WAR is expected, on average, to increase.
- However, if the competence-driven branching process is a reasonable account of firm diversification, the $WARN$ measure should not be affected by diversif.

Results

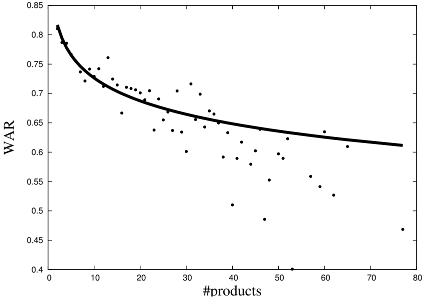
- WAR for products: As expected, as firms increase their product scope, the coherence across all its activities decrease.
- WARN for products: as firms introduce new products the coherence between neighboring activities slightly increase for relatively low levels of diversification, and stay constant for sufficiently diversified firms.
- The analysis of relatedness measure on patents suffer for lower degree of technological diversification as opposed to product

Product WAR as function of #products

Figure:



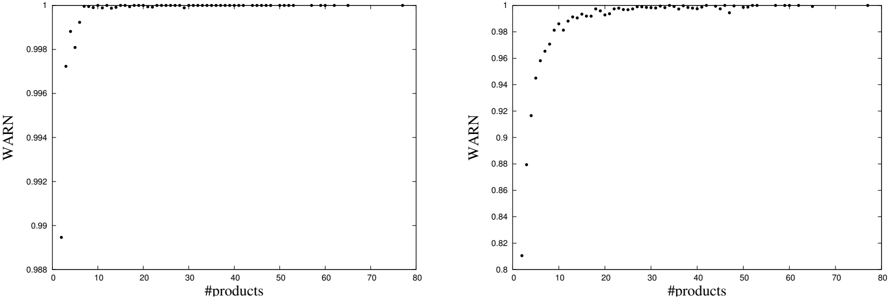
(a) Standard null hypothesis. Linear fit



(b) Alternative null hypothesis. Log-linear fit

Product WARN as funtion of #products

Figure:

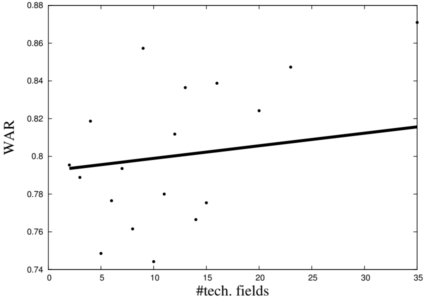


(a) Standard null hypothesis

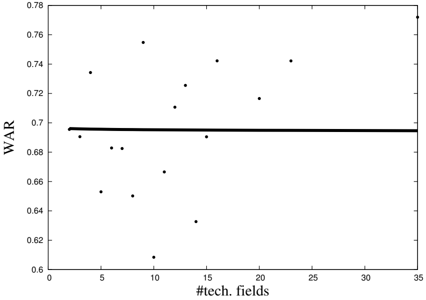
(b) Alternative null hypothesis

Technological WAR as funtion of #tech. fields

Figure:



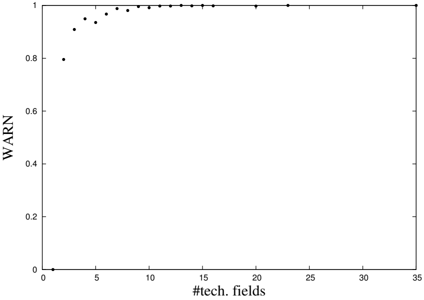
(a) Standard null hypothesis. Linear fit



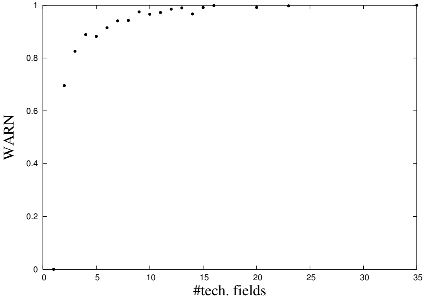
(b) Alternative null hypothesis. Log-linear fit

Technological WARN as funtion of #tech. fields

Figure:



(a) Standard null hypothesis



(b) Alternative null hypothesis

Stylized facts on heterogeneity

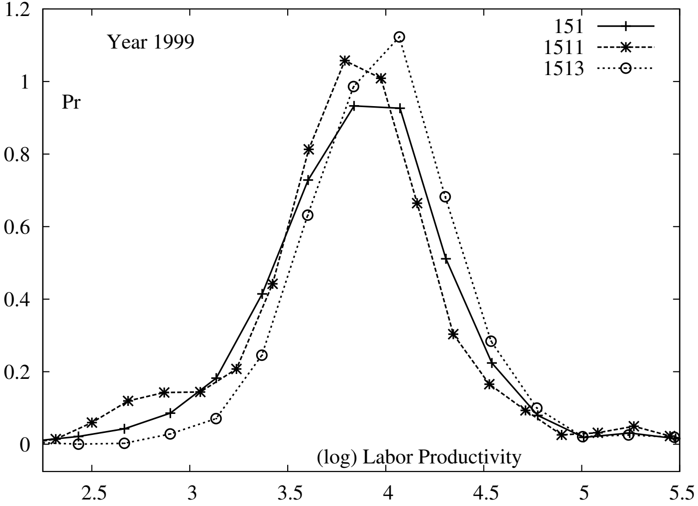
Robust evidence across many industries and countries (USA, Canada, UK, France, Italy, Netherlands, etc) consistently finds:

- wide asymmetries in productivity across firms
- equally wide heterogeneity in relative input intensities
- highly skewed distribution of efficiency, innovativeness and profitability indicators;
- different export status within the same industry
- **high intertemporal persistence in the above properties**
- **high persistence of heterogeneity also when increasing the level of disaggregation**

Disaggregation does not solve the problem

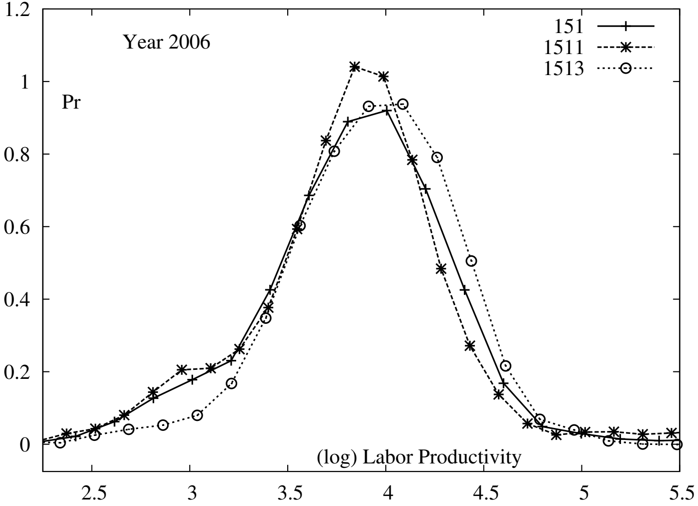
"*We [...] thought that one could reduce heterogeneity by going down from general mixtures as "total manufacturing" to something more coherent, such as "petroleum refining" or "the manufacture of cement." But something like Mandelbrot's fractal phenomenon seems to be at work here also: the observed variability-heterogeneity does not really decline as we cut our data finer and finer. There is a sense in which different bakeries are just as much different from each others as the steel industry is from the machinery industry.*" (Griliches and Mairesse, Production function: the search for identification, 1999)

Heterog. performances Meat Products (1999)



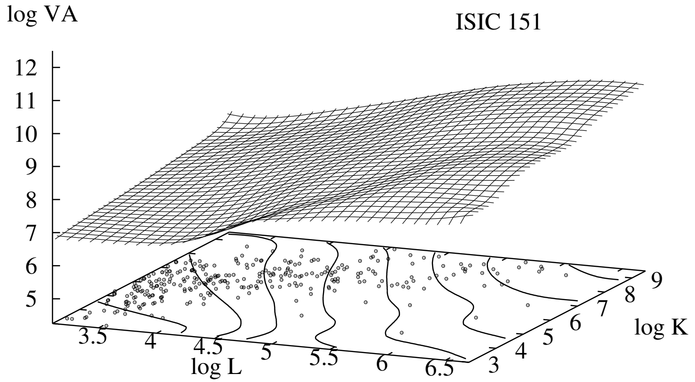
exp(3) ≈ 20 th. euro; exp(4.5) ≈ 90 th. euro

Heterog. in performances is persistent (year 2006)



exp(3) ≈ 20 th. euro; exp(4.5) ≈ 90 th. euro

Heterog. in adopted techniques



Strands of literature related to firm heterogeneity

- Intra-industry differences in productivity
  - Griliches and Mairesse (1999); Bartelsman and Doms (2000); Disney et al. (2003); Syverson (2011)
- Productivity and Selection
  - Baily et al. (1992); Baldwin and Rafiquzzaman (1995); Foster et al. (2008)
- Trade
  - Melitz (2003); Bernard et al. (2007); Melitz and Ottaviano (2008)

This is puzzling....

This evidence poses serious challenges to:

- Theories of competition and market selection

- Theoretical and/or empirical analyses which rely upon some notion of industries as aggregates of *similar/homogeneous* production units:
  - models based on industry production function
  - empirical exercises based on some notion of efficiency frontier
  - but also sectoral input-output coefficient à la Leontief are meaningless if computed as averages over such very dispersed and skewed distributions
  - indicators of technical change based on variations of such aggregates (isoquants or input-output coefficients) may be seriously misleading

Our attempt

- Can we give a representation of the production technology(ies) of an industry without denying heterogeneity, but fully taking it into account?
- ... and without imposing any hypothesis on functional forms or input substitutions?
- Can we produce empirical measures of the technological characteristics of an industry which explicitly take into account heterogeneity?
- we make an attempt going back and developing upon W. Hildenbrand "Short-run production functions based on microdata" *Econometrica*, 1981

Hildenbrand's analysis

- Represent firms in one sector as empirical input-output vectors of production at full capacity
- with some weak additional assumptions (divisibility) derives the empirical production possibility set for the industry (geometrically, a zonotope)
- and shows the following main properties of the derived efficiency frontier:
  - returns to scale are never constant
  - the elasticities of substitution are not constant

Our contribution

Building upon Hildenbrand (1981) we derive:

- indicators of industry heterogeneity
- rigorous measures of technical change at the industry level which do not assume any averaging out of heterogeneity
  - rate and direction of technical change
- Industry dynamics: how firm entry and exit affects heterogeneity and tech change
- We provide an application on Italian industrial census data
- Compare with existing measure of productivity

Production activities and Zonotopes

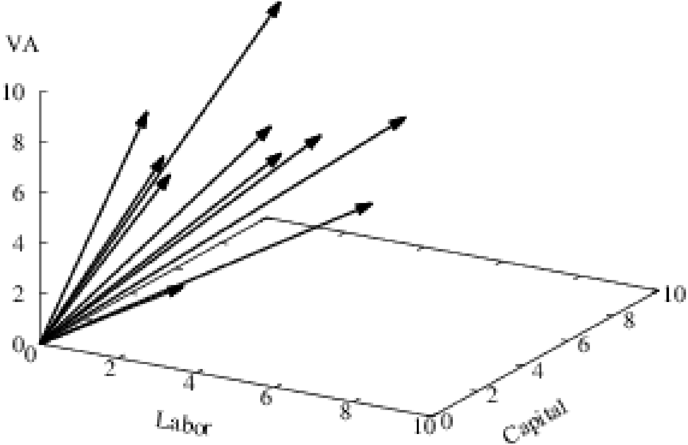
- The *ex post* technology of a production unit is a vector

$$a = (\alpha_1, \ldots, \alpha_l, \alpha_{l+1}) \in \mathbb{R}_+^{l+1},$$

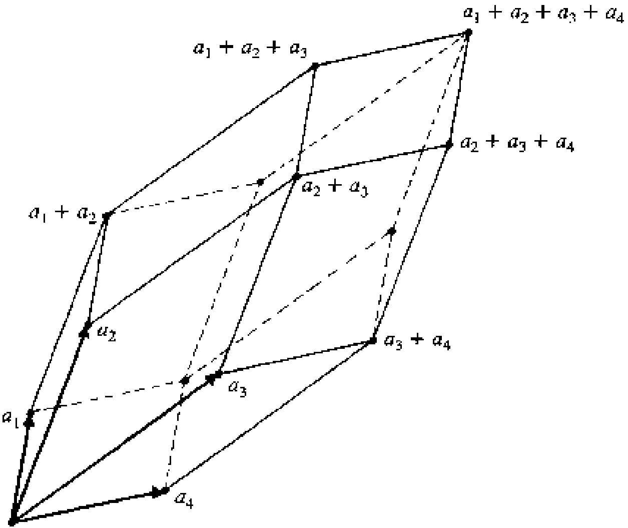
  i.e. a **production activity** a that produces, during the current period, α_{l+1} units of output by means of $(\alpha_1, \ldots, \alpha_l)$ units of input.
  - Holds also also for the multi-output case
- The **size** of the firm is the length of vector a, i.e. a multi-dimensional extension of the usual measure of firm size.
- The short run production possibilities of an industry with N units at a given time is a finite family of vectors $\{a_n\}_{1 \leq n \leq N}$ of production activities
- Hildenbrand defines the **short run total production set** associated to them as the Zonotope

$$Y = \{y \in \mathbb{R}_+^{l+1} \mid y = \sum_{n=1}^{N} \phi_n a_n, 0 \leq \phi_n \leq 1\}.$$

Production activities in a 3-dimensional space

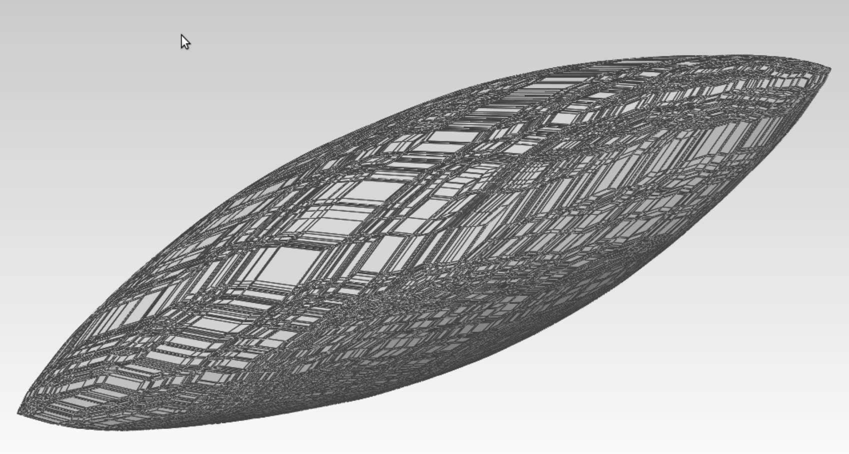


The Zonotope



$$Y = \sum_{n=1}^{4} [0, a_n] \subset R^3$$

Zonotope generated by 300 random vectors



Volume of Zonotopes and Gini index

- The **volume** of the zonotope Y in $\mathbb{R}^{l+1}$ is given by:

$$Vol(Y) = \sum_{1 \leq i_1 < \ldots < i_{l+1} \leq N} \mid \Delta_{i_1,\ldots,i_{l+1}} \mid$$

  where $\mid \Delta_{i_1,\ldots,i_{l+1}} \mid$ is the module of the determinant $\Delta_{i_1,\ldots,i_{l+1}}$.
- Interested in getting an **absolute** measure of the heterogeneity in techniques; independent both from the number of firms making up the sector and from the unit in which inputs and output are measured.
- This absolute measure is the **Gini volume** of the Zonotope (a generalization of the well known Gini index):

$$Vol(Y)_G = \frac{Vol(Y)}{Vol(P_Y)} \quad , \tag{9}$$

  where $Vol(P_Y)$ is the volume of the parallelotope P_Y of diagonal $d_Y = \sum_{n=1}^{N} a_n$, that is the maximal volume we can get when the industry production activity $\sum_{n=1}^{N} a_n$ is fixed.

Remark on complete heterogeneity

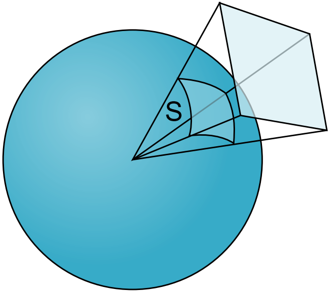
- Complete heterogeneity is not feasible
  - Note that alike the complete inequality case in the Gini index, i.e. the case in which the index is 1, also the complete heterogeneity case is not feasible in our framework, since in addition to firms with large values of inputs and zero output it would imply the existence of firms with zero inputs and non zero output. It has to be regarded as a *limit* similarly to the 0 volume in which all techniques are equal, i.e. the vectors $\{a_n\}_{1 \leq n \leq N}$ are proportional and hence lie on the same line.

Unitary production activities

- What is the role of size in industry heterogeneity?
- Compare volume of the original zonotope, Y, to that where all firms have the same size $\overline{Y}$
- Zonotope $\overline{Y}$ generated by the normalized vectors $\{\frac{a_n}{\|a_n\|}\}_{1\leq n\leq N}$, i.e. the unitary production activities.
- The Gini volume $Vol(\overline{Y})_G$ evaluates the heterogeneity of the industry in a setting in which all firms have the same size (norm is equal to one)
- The only source of heterogeneity is the difference in adopted techniques
  - Differences in firm size do not contribute to the volumes
- Intuitively, if the Gini volume $Vol(Y)_G$ is bigger than $Vol(\overline{Y})_G$ then big firms contribute to heterogeneity more than the small ones
  - and viceversa

Solid Angle

- In geometry, a solid angle (symbol: Ω) is the two-dimensional angle in three-dimensional space that an object subtends at a point.
- It is a measure of *how large* the object appears to an observer looking from that point.
- It can considered as the multi-dimensional analog of the support of the distribution of one variable
- An object's solid angle is equal to the area of the segment of a unit sphere that the object covers, as shown in figure 10.



External activities

- **External** production activities define the span of the solid angle
- Normalized production activities $\{\frac{a_n}{\|a_n\|}\}_{1\leq n\leq N}$ generate an arbitrary pyramid with apex in the origin.
- Note: in general, not all vectors a_i, $i = 1, \ldots, N$ will be edges of this pyramid.
  - It might happen that one vector is *inside* the pyramid generated by others

$\Rightarrow$ **external** vectors $\{e_i\}_{1\leq i\leq r}$ are edges of the pyramid.

- All the others will be called **internal**.
- Define the external Zonotope Y_e generated by vectors $\{e_i\}_{1\leq i\leq r}$.
- Pairwise comparison of $Vol(Y_e)_G$ and $Vol(Y)_G$ shows relative importance of the *density* of internal activities in affecting heterogeneity.

Angles and technical change

- Our measure of efficiency of the industry is the angle that the main diagonal, d_Y, of the zonotope forms with the space generated by all inputs
- This can be easily generalized to the case of multiple outputs
  - $\Rightarrow$ Appendix for the general case
- In a 2-inputs, 1-output setting, if $d_Y = (d_1, d_2, d_3)$, this is equivalent to study

$$\mathrm{tg}\theta_3 = \frac{d_3}{\|(d_1, d_2)\|} \tag{10}$$

- If the angle increases, then productivity increases

Direction of Technical change

- How relative inputs use varies over time
- Consider the angles that the input vector forms with the input axis
- In the two-inputs, one-output case

$$\mathrm{tg}\varphi_1 = \frac{d_2}{\|d_1\|} \tag{11}$$

- If input 1 is labor and input 2 is capital, an increase in φ_1 suggests that technical change is biased in the labor saving direction.

Normalized technical change

It is also interesting to measure the changes in the *normalized* angles, i.e. the ones related to the diagonal $d_{\overline{Y}}$.
In particular the comparison of the changes of two different angles is informative on the relative contribution of bigger and smaller firms to productivity changes and hence, on the possible existence of economies/diseconomies of scale.

Entry and exit of a firm: general case

How entry/exit of a firm contributes to heterogeneity and tech change.
If $Z \in \mathbb{R}^{l+1}$ is the Zonotope generated by vectors $\{a_n\}_{1 \leq n \leq N}$ and $b = (x_1, \ldots, x_{l+1}) \in \mathbb{R}^{l+1}$ is a new firm, the volume of the zonotope X generated by $\{a_n\}_{1 \leq n \leq N} \cup \{b\}$ is:

$$Vol(X) = Vol(Z) + V(x_1, \ldots, x_{l+1})$$

where $V(x_1, \ldots, x_{l+1}) = \sum_{1 \leq i_1 < \ldots < i_l \leq N} \mid \Delta_{i_1, \ldots, i_l} \mid$ and $\Delta_{i_1, \ldots, i_l}$ are the determinant of the matrix $A_{i_1, \ldots, i_l}$ whose rows are the vectors $\{b, a_{i_1}, \ldots, a_{i_l}\}$. The diagonal of X is $d_X = d_Z + b$ and the heterogeneity for the *new* industry is the real function on $\mathbb{R}^{l+1}$:

$$Vol(X)_G = \frac{Vol(Z) + V(x_1, \ldots, x_{l+1})}{Vol(P_X)} \quad .$$

To study the variation (i.e. gradient, hessian etc...) of $Vol(X)_G$ is equivalent to analyze the impact of a new firm on the industry.

Entry and exit: 3-dimensional case

As an example, in the 3-dimensional case we get:

$$V(x_1,x_2,x_3) = \sum_{1\le i<j\le N} \mid x_1(a_i^2 a_j^3 - a_i^3 a_j^2) - x_2(a_i^1 a_j^3 - a_i^3 a_j^1) + x_3(a_i^1 a_j^2 - a_i^2 a_j^1) \mid,$$

$d_X = \sum_{i,j,k=1}^{N}(a_i^1 + x_1)(a_j^2 + x_2)(a_k^3 + x_3)$ and

$$Vol(X)_G = \frac{Vol(Z) + V(x_1,x_2,x_3)}{d_X}$$

is a 3 variables function with $Vol(Z)$ and $\{a_n^1, a_n^2, a_n^3\}_{1\le n\le N}$ constants. If we set x_3, i.e. the output of the firm b, constant $Vol(X)_G$ becomes a function of two variables, that is $Vol(X)_G = Vol(X)_G(x_1,x_2)$, which can be easily studied from a differential point of view. So, for example, when this function increases then the new firm positively contributes to the industry heterogeneity and viceversa.

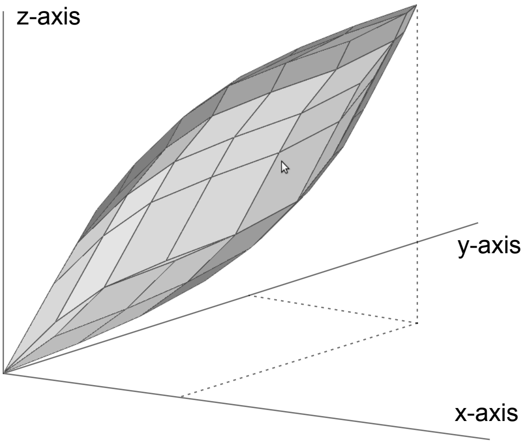
A toy illustration

Production schedules of 10 hypothetical firms composing an industry, 2-inputs, capital and labor, and one output.

	Year 1			Year 2			Year 3			Year 4		
Firm	L	K	VA	L	K	VA	L	K	VA	L	K	VA
1	7	4	9	7	4	9	**7**	**4**	**9**	**7**	**4**	**9**
2	1	4	5	1	4	5	**1**	**4**	**5**	**1**	**4**	**5**
3	**6**	**2**	**9**	**6**	**2**	**9**	**6**	**2**	**9**	**6**	**2**	**9**
4	**1.5**	**8**	**10**	**1.5**	**8**	**10**	**1.5**	**8**	**10**	**1.5**	**8**	**10**
5	5	2	8	5	2	8	5	2	8	5	2	8
6	**1**	**3**	**8**	**1**	**3**	**8**	**1**	**3**	**8**	**1**	**3**	**8**
7	**2**	**2**	**7**	**2**	**2**	**7**	**2**	**2**	**7**	**2**	**2**	**7**
8	3	5	7	3	5	7	3	5	7			
9	**2.5**	**2**	**2**	**2.5**	**2**	**2**						
10	**5**	**6**	**4.0**	4	4	6	**4**	**4**	**6**	**4.0**	**4**	**6**

Table: Production schedules in year 1 to 4, Number of employees (L), Capital (K) and Output (VA). External production activities in bold.

Zonotope generated by 10 vectors in Year 1



Solid angle in year 1 and 2.

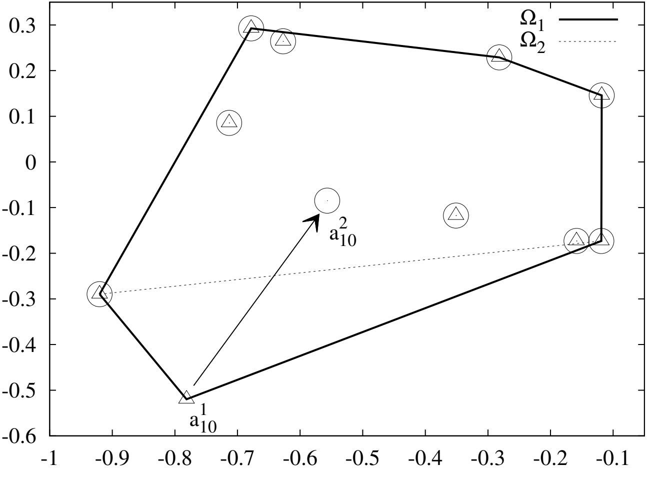


Figure: Planar section of the solid angle generated by all vectors in year 1 and 2. The section plane is the one perpendicular to the vector sum of generators in year 1. a_{10} moves inward from year 1 to year 2.

From Year 1 to Year 2

Firm 10	L	K	VA
Year 1	**5**	**6**	**4**
Year 2	4	4	6

- Firm 10 display unequivocal increase in productivity
- Vector of firm 10 rotates inward, it was external, it becomes internal
- All other firms unchanged

What happens to industry heterogeneity and tech. change ?

- A boundary vector shifts inward, production techniques are more similar $\Rightarrow$ heterogeneity reduces
- Zonotope's main diagonal gets steeper $\Rightarrow$ productivity increases
- tgφ_1 decreases; tech change is biased in the capital saving direction

Heterogeneity and Technical change in a toy example

	Year 1	Year 2	Year 3	Year 4
$Vol(Y^t)_G$	0.09271	0.07196	0.06518	0.06880
$Vol(\overline{Y}^t)_G$	0.09742	0.07905	0.06795	0.07244
$Vol(Y_e^t)_G$	0.12089	0.09627	0.07297	0.07297
Solid Angle	0.28195	0.22539	0.15471	0.15471
tgθ_3^t	1.3532	1.4538	1.51066	1.55133
tgφ_1^t	1.11765	1.09091	1.11475	1.05455
Malmquist Index		1.00460	1.02656	1.02859

Gini volume for the zonotopes Y^t; the zonotopes $\overline{Y}^t$ generated by the normalized production activities $\{\frac{a_j^t}{\|a_j^t\|}\}_{1\leq j\leq 10}$; the zonotopes Y_e^t generated by the external production activities; the solid angle; and the angles that account for the rate and direction of technical change.

From Year 2 to Year 3

- Firm 9, an external vector leaves the industry
- The outcome is smaller Gini volumes for all measures
- The solid angles also reduces
- As before, tgθ_3 increases. The exit of firm 9 further boost efficiency
- Tech change acts now in the labor saving direction

Accounting for entry and exit in the toy example

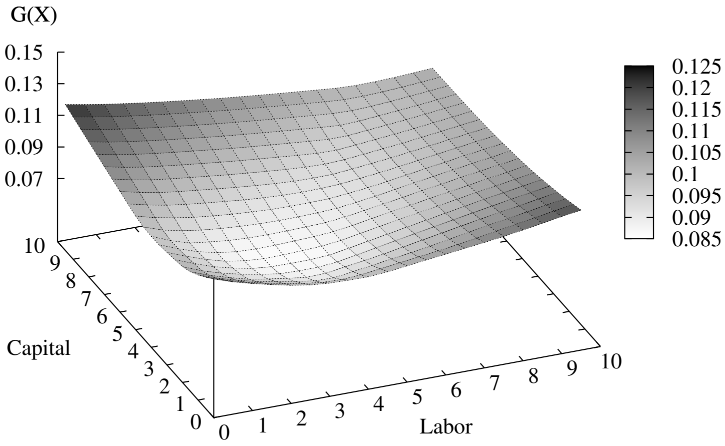


Figure: Variation of heterogeneity (z axis) when a firm of labor x, capital y and VA= 5

The Database Micro.3 1989-2006

- Micro.3 is the census of Italian firms bigger than 20 employees (change in data collection in 1998)
  - More than 40% of employment in the manuf. industry
  - More than 50% of value added in the manuf. industry
  - unbalanced panel of over 100,000 firms
- Integrated sources of data ⇒ Istat Census (SBS like), Financial Statements, CIS, trade, patents.
- Censorship of any individual information; data accessible at Istat facilities.

⇒ A Plus From 1998 availability of financial statements that is a legal requirements for *all* incorporated firms.

Measure of industry heterogeneity

NACE Code	$G(Y)$ '98	'02	'06	$G(\overline{Y})$ '98	'02	'06
1513	0.059	0.051	0.062	0.082	0.062	0.096
1721	0.075	0.068	0.103	0.075	0.078	0.124
1930	0.108	0.139	0.150	0.110	0.115	0.123
2121	0.108	0.043	0.062	0.081	0.064	0.081
2524	0.089	0.083	0.094	0.097	0.088	0.096
2661	0.079	0.088	0.099	0.100	0.094	0.110
2811	0.105	0.109	0.109	0.117	0.113	0.122
2852	0.088	0.102	0.110	0.100	0.103	0.111
2953	0.072	0.095	0.096	0.098	0.104	0.111
2954	0.078	0.074	0.093	0.086	0.130	0.113
3611	0.078	0.099	0.118	0.107	0.096	0.121

Table: Normalized volumes in 1998, 2002 and 2006 for selected 4 digit sectors.

Heterogeneity does not vanish over time

Industry level productivity change

NACE Code	(a) rates of growth of tg θ_3 1998-2002	2002-2006	(b) Malmquist TFP Index 1998-2002	2002-2006
1513	-11.9073	-11.4541	0.96509	0.85804
1721	10.5652	4.3723	0.99174	1.07459
1930	3.1152	25.2797	1.19082	1.12582
2121	-6.8362	-8.8206	1.02747	0.89696
2524	-15.2821	0.4118	0.96125	0.98312
2661	6.7277	-18.5953	0.74406	0.88080
2811	6.4256	-7.9102	1.02165	0.70803
2852	-12.0712	2.1536	0.90663	0.66255
2953	19.3637	-4.7927	1.01951	0.92981
2954	-0.3020	-21.2919	1.08091	1.34540
3611	-17.9141	0.0892	0.75043	1.11615

Table: (a) Angles of the zonotope's main diagonal, rates of growth; (b) Malmquist index TFP growth

Conclusions & further work

- Building on the seminal contribution by Hildenbrand (1981) we exploit the geometrical properties of the zonotope to study intra-industry heterogeneity and technical change
- We consider how entry and exit affects industry heterogeneity and productivity
- Comparison with existing measures (TFP estimates and productivity index)
- Further work / developments
  - Investigate industry dynamics (fine-grained sector) in presence of an exogenous shock (i.e. introduction of innovation)
  - What is the impact on the adopted techniques?
  - Is there convergence after the perturbation? How long does it take to revert to initial level of heterogeneity?
  - How fast the innovation does spread? How long does it take to observe productivity boosting effect of innovation?

Thank you!

Acknowledgment
Special thanks to Federico Ponchio that developed the software for computation and plotting:
`http://vcg.isti.cnr.it/~ponchio/zonohedron.php`

Appendix: Angle and Tech. Change

Let us consider a non-zero vector $v = (x_1, x_2, \ldots, x_{l+1}) \in \mathbb{R}^{l+1}$ and, for any $i \in 1, \ldots, l+1$, the projection map

$$\begin{aligned} pr_{-i} : \mathbb{R}^{l+1} &\longrightarrow \mathbb{R}^{l} \\ (x_1, \ldots, x_{l+1}) &\mapsto (x_1, \ldots, x_{i-1}, x_{i+1}, \ldots, x_{l+1}) \quad . \end{aligned}$$

Using the trigonometric formulation of the Pythagoras' theorem we get that if

- if ψ_i is the angle that v forms with the x_i axis;
- $\theta_i = \frac{\pi}{2} - \psi_i$ is its complement;
- $\|v_i\|$ is the norm of the projection vector $v_i = pr_{-i}(v)$

then the tangent of θ_i is:

$$\text{tg}\theta_i = \frac{x_i}{\|v_i\|}.$$

- We are interested in the angle θ_{l+1} that the diagonal, i.e. the vector d_Y, forms with the space generated by all inputs.
- Easily generalizable to the case of multiple outputs.

Back to Angle and Tech Change

Appendix: Malmquist Index

- Use TFP estimates to study aggregate (country, industry, etc) change in productivity
- Industry production at time 1 and 2 are described respectively by

$$Q_{1,1} = f_1(L_1, K_1) \qquad Q_{2,2} = f_2(L_2, K_2) \tag{12}$$

$Q_{1,1}$ denotes technology of time 1 and input quantities of time 1.

- The Malmquist index is the geometrical mean of $Q_{1,1}/Q_{1,2}$ and $Q_{2,1}/Q_{2,2}$

Back to toy example table

Back to Micro.3 table