

INSTITUTE
OF ECONOMICS



Scuola Superiore
Sant'Anna

LEM | Laboratory of Economics and Management

Institute of Economics
Scuola Superiore Sant'Anna

Piazza Martiri della Libertà, 33 - 56127 Pisa, Italy
ph. +39 050 88.33.43
institute.economics@sssup.it

LEM

WORKING PAPER SERIES

Varieties of deindustrialization and patterns of diversification: why microchips are not potato chips

Giovanni Dosi ^a
Federico Riccio ^a
Maria Enrica Virgillito ^a

^a Institute of Economics and EMbeDS, Scuola Superiore Sant'Anna, Pisa, Italy.

2020/11

May 2020

ISSN(ONLINE) 2284-0400

Varieties of deindustrialization and patterns of diversification: why microchips are not potato chips*

Giovanni Dosi^{**1}, Federico Riccio¹, and Maria Enrica Virgillito¹

¹Institute of Economics and EMbeDS Department
Scuola Superiore Sant'Anna[†]

Abstract. Contrarily to the notion of a natural tendency of deindustrialization, this paper, documenting the existence of a variety of patterns of deindustrialization, performs a cross-country, long-term analysis. Looking at industrial sectors and their technological characteristics, categorised on the ground of the Pavitt (1984) taxonomy, we do find a markedly uneven process of deindustrialization with Science Based and Specialised Suppliers not presenting any inverted U-shaped pattern, neither in employment nor in value added. The heterogeneity holds both for the four Pavitt aggregates and under further disaggregation at industry level. We then study whether the uneven sectoral composition might have exerted an impact on the timing of deindustrialization. Overall, our analysis brings support to the notion that “microchips” are not “potato chips” in their influence on the patterns of long-term economic development of different countries. Moreover, during the phase of globalisation the probability for low-income countries to be stuck to produce “potato chips” has increased and that of transition toward the production of “microchips” has been reducing.

Keywords: Deindustrialization · Structural Change · Diversification · Technological Change

JEL classification: J20, L60, O14, O25, O30.

[†] Piazza Martiri della Libertà 33, I-56127, Pisa (Italy). E-mail addresses: gdosi<at>santannapisa.it, f.riccio<at>santannapisa.it, m.virgillito<at>santannapisa.it

* The authors wish to thank participants to the 31st EAEPE conference in Warsaw and the 8th WIPE in Reus for their useful comments on the manuscript. The authors acknowledge support from European Union’s Horizon 2020 research and innovation programme under grant agreement No. 822781 GROWINPRO – Growth Welfare Innovation Productivity. This manuscript version is made available under the CC-BY-NC-ND 4.0 license.

** Corresponding author

1 Introduction

A process of generalised deindustrialization represents a widespread feature of the current phase of capitalist development. From the peak of 40% share of the overall workforce employed in the manufacturing sector in the sixties, nowadays the overall manufacturing employment share in advanced economies ranges between 10%-25%.

Deindustrialization beyond a certain income threshold is often thought of a “natural tendency” of capitalism. In such a perspective, deindustrialization should occur evenly across countries and industrial sectors. However, the historical evidence is at odds with this view. Countries deeply differ both in their process of industrialization and eventually deindustrialization, and with that, in the consequences that deindustrialization bears thereafter upon the pattern and “quality” of development. Consider two “archetypical” alternative examples like Brazil and South Korea. Both countries experienced a rapid process of catching-up with the former starting indeed in the fifties from much more higher levels of GDP per capita. However their development path followed completely different trajectories, with Brazil falling behind since the eighties and South Korea nowadays included among the most prosperous economies. And now they differ even in the response to major exogenous shocks. South Korea has recently demonstrated a superb capability in managing the spreading of the Covid-19 pandemic. The pandemic is teaching us that, other things being equal, more industrialised countries are better equipped to manage them.

Contrarily to the belief on any “natural tendency” in deindustrialization, this paper performs a cross-country, long-term analysis, documenting the existence of a variety of patterns of deindustrialization. Looking at industrial sectors and their technological characteristics, categorised on the ground of the Pavitt (1984) taxonomy, we do find a markedly uneven process of deindustrialization with *Science Based* and *Specialised Suppliers* sectors not presenting any inverted U-shaped pattern, neither in employment nor in value added. The heterogeneity holds both for the four aggregates of the Pavitt taxonomy and under further disaggregation at industry level.

We then study whether the uneven sectoral composition might have exerted an impact on the timing of deindustrialization. We find, first, that *Scale Intensive* and *Specialised Suppliers* industries have reduced their employment shares more dramatically than *Supplier Dominated* ones. Countries stuck into the latter seem to have missed major opportunities of catching-up. Second, after performing a cluster analysis, we do find that the higher the degree of diversification in industrial composition the higher the level of GDP per capita. In this respect, four different clusters in terms of composition of manufacturing employment shares do emerge, distinct in terms of their degrees of diversification. Finally, we do find evidence that the post-1990 time period is characterised by increasing probabilities of falling behind even for industrialised rich countries. That is, globalization, far from fostering growth opportunities for developing countries, has contributed in general to freeze technological upgrading and opportunities of

catching-up from traditional toward innovative manufacturing sectors. With China, of course, standing out as a major exception.

Overall, our analysis brings support to the notion that “microchips” are not equivalent to “potato chips”:¹ the industrial composition of manufacturing highly influences the patterns of long term economic development of the countries. During the phase of globalization the probability for low-income countries to produce “potato chips” has increased while the transition probability toward the production of “microchips” has been reducing.

The paper is organised as follows: Section 2 discusses the main theoretical background, Section 3 presents the data and performs the analysis of deindustrialization patterns disaggregating by the Pavitt taxonomy. Section 4 analyses the possible processes of premature deindustrialization by technological classes while in Section 5 we perform a cluster analysis to detect the underlying patterns of country diversification. Finally, our conclusions and policy implications are sketched in Section 6.

2 Deindustrialization, sectors and quality of specialization

The very first question one needs to address is whether it is still relevant to discuss about the role of the manufacturing sector for the process of economic development. Back to the structuralist perspective but even earlier to List (1841), what a country produces, does matter. According to the “Hirschman-Prebisch” approach, the manufacturing sector is the *engine of growth* for two specific reasons namely, first, the rate of productivity growth occurring in the manufacturing sector is comparatively high, second, the gains from productivity growth are transferred into wages with a higher elasticity (Hirschman, 1958; Prebisch, 1959; Szirmai, 2012).

The benefits from manufacturing manifest since the first Industrial Revolution. The latter, resulting from the interaction between technical change and division of labour inside factories, has represented a turning point for the overall process of economic development and since then, the manufacturing sector has represented the locus of capabilities accumulation, learning by doing and dynamic increasing returns (Arrow, 1962; Kaldor, 1967; Abramovitz, 1986; Cimoli and Dosi, 1995). Indeed, the development of a manufacturing sector has happened to be a necessary condition for catching-up (Szirmai and Foster-McGregor, 2017), with countries gradually moving from low-technology (e.g. textile) toward high-technology sectors (e.g. ICT). The manufacturing sector is an engine of growth also because it generates positive spillover effects for the economy as whole. These spillovers show up both in terms of wage effects, by means of Kaldorian virtuous circles (Kaldor, 1967), but also in terms of employment multipliers, meaning jobs indirectly created by the manufacturing sector in other sectors of the economy – ranging from knowledge intensive services, to societal services as education and health, to social consumption services as media industries –.

¹ The statement “It doesn’t make any difference whether a country makes potato chips or computer chips!” is attributed to J. Boskin, chairman of President H. W. Bush’s Council of Economic Advisers, although he denies it. The expression “microchips, not potato chips” became popular after the article by Thurow (1994).

However, being as such the engine of productivity growth, the overall employment share of the manufacturing sector has experienced a strong contraction, with an increasing fraction of the workforce nowadays employed in the service sector. This is indeed the key symptom of the process of deindustrialization. If manufacturing is a key stage for the process of catching-up (Rodrik, 2013), and wages are generally higher than in the service sector as a whole, when do countries deindustrialize and what happens thereafter?

Historically, take-off and economic growth are associated with a movement of the labour force from agriculture to manufacturing, and finally to the service sector (Kuznets and Murphy, 1966; Chenery and Syrquin, 1975; Landesmann and Pichelmann, 1999; Nuvolari and Russo, 2019). This transition occurred led by rates of technological innovation and diffusion, uneven across sectors, in time and in space. The prevalence of the labour-creative or destructive effect accompanying the process of structural change mainly depends on whether output growth (demand) is higher/lower than productivity growth (stemming from technical change). Demand growth and productivity growth are linked first by the price elasticity channel: productivity differentials, reducing consumer prices, lead to a stimulus to demand in sectors experiencing high productivity growth (Clark, 1957; Baumol, 1967; Pasinetti, 1983). On the other hand, the structure of Engel curves – i.e., the income elasticity of demand – plays a role in influencing the dynamics of outputs and indirectly of inputs, primarily labour. For these reasons, the employment absorbing sectors (toward which the labour force moved) have been generally characterised by high labour intensity and high-income elasticity of demand, at least in the initial phase of sectoral development (Freeman, Clark, and Soete, 1982). Demand growth ensures structural change toward those new sectors associated with a sufficient creation of new jobs. At the same time, the effect of labour productivity growth in influencing the direction of structural change is more ambiguous. The transition from agriculture to manufacturing was a shift from a lower productivity toward a higher productivity sector. What we observe nowadays is the shift to both high productivity sub-sectors in manufacturing (e.g., ICT and biorobotics) and in services (e.g. software), and to low productivity ones (health and education).

Contrary to any simple argument in favour of a natural tendency to deindustrialization (Ramaswamy and Rowthorn, 1997), the complex relationship between demand patterns and productivity dynamics strongly hints at the possible country specificity in the dynamics of structural change. Besides the relationship between product and process innovation in balancing labour demand patterns (Dosi, Piva, Virgillito, and Vivarelli, 2019), one has to bring into the picture the role of globalization. In that respect, Rodrik (2016) renewed the deindustrialization debate highlighting the tendencies for developing economies to experience a premature shift in both manufacturing employment and value added shares in the age of globalization. By splitting the sample in two periods, before and after the fall of the Berlin Wall, there is evidence of an acceleration process of deindustrialization after 1990. In turn, the changing importance of globalization and international trade in explaining deindustrialization clearly downplays the natural tendency argument. One interpretation of “premature deindustrialization” rests

on the declining prevalence of manufacturing products in the whole basket of consumption. However, Haraguchi, Cheng, and Smeets (2017) explain the patterns of deindustrialization of the West as a process of relocation of manufacturing activities in the East. Therefore, more than a story of exhaustion of opportunities for the manufacturing sector, premature deindustrialization results from the process of international relocation of industrial activity in few catching-up countries primarily China (Yu, Dosi, Lei, and Nuvolari, 2015). All in all, the interactions between productivity, demand and relocation patterns cannot be detected at the aggregate levels, given the complex product-sectoral dynamics behind.

The scant literature on the industry-level patterns of deindustrialization (Haraguchi, Verspagen, and Amann, 2019; Kunst, 2019) points at heterogeneous dynamics in output, productivity, employment and occupations across manufacturing sub-sectors. However, such heterogeneity is far from random. On the contrary, it is rooted in the historical patterns of industrialization of each country. Vindicating a general evolutionary/structuralist interpretation of why income levels and growth rates differ (Dosi, Pavitt, and Soete, 1990), producing the metaphorical equivalent of potato chips, even within manufacturing, is quite different from producing microchips. This is so because the learning opportunities, the process of accumulation of capabilities and the ensuing scope of increasing returns are dramatically different across sectors.

Thus, the patterns of sectoral specialization/diversification dramatically impact upon the overall process of economic growth (Fagerberg, 1987; Lee and Malerba, 2017). This is revealed by strikingly simple zero-parameter statistics (Tacchella, Cristelli, Caldarelli, Gabrielli, and Pietronero, 2013; Hausmann, Hwang, and Rodrik, 2007) as a proxy for the complexity of the products in which each country is specialised. Even more so, such an inter-country diversity ought to emerge by opening up the black box of technological attributes of the different sectors. In that respect, the Pavitt (1984) taxonomy reveals to be a quite useful lens of analysis to classify sectors according to their technological content, position in the supply chain, and overall quality of specialization. Indeed, to our knowledge, what follows is the first empirical analysis looking at the deindustrialization process by means of such sectoral and technological perspective. The development of differentiated capabilities according to the prevailing sectoral compositions (Cimoli and Dosi, 1995), we will show, influences the subsequent country varieties of deindustrialization, or non-industrialization at all, their timing and consequences.

3 Varieties of deindustrialization: sectoral and technological heterogeneity

In this section we first present the data structure and then document the patterns of sectoral and technological heterogeneity in industrial structures and deindustrialization patterns. We consider both employment and nominal valued added shares.²

² Nominal value added does not distinguish between movement in quantities and prices. However, industry-country specific deflators are still not available.

3.1 Overview of the data and preliminary evidence

We employ the INDSTAT2 data set (UNIDO, 2018) which comes from national industrial surveys on manufacturing industries. It excludes the informal sector and firms with less than five employees, spanning a wide ensemble of countries (173) from 1963 to 2015. It includes data for 23 manufacturing industries according to the ISIC Rev.3.1 classification (see Appendix A for a detailed description of the data coverage).³

Real GDP per capita (at constant 2011 dollars), population and total employment come from the Penn World Table 9.0 version (Feenstra et al., 2015) while total value added in current U.S. dollars comes from the Manufacturing Value Added data set (UNIDO MVA). The final sample coverage results from the combination of these sources.

By means of a non-parametric descriptive analysis, we start investigating the relevance of a disaggregated sectoral analysis of deindustrialization. We regress our measures of deindustrialization on country log-(per capita) income level, by using the locally weighted scatter-plot smoothing (*lowess*),⁴ thus running a local polynomial fit of y (employment and value added industry-level shares) on x (income level) for n localised subsets of the data. A weighting function attributes higher weights to observations closer to the estimation point.

Figure 1 shows the *lowess* fit by sectors for both employment (left-hand side) and value added (right-hand side), plotting the average value among the n sub-samples. Consider employment shares. Although all sectors display the typical inverted U-shape relation, there are significant differences both in the concavity, in the peak share and location. For example, Food & Beverage and Textile peak at lower income levels and rapidly fade away for higher ones. Basic and Fabricated Metals and Machinery sectors emerge at higher income levels and present a higher peak share. The former reaches 2% of total employment and it slowly declines, while the latter reaches the 3% of total employment and shows only mild signs of decline. Figures for value added are more heterogeneous. Textile and Food & Beverage, both at 4% share at the initial per capita income, have completely different patterns: while the Textile sector continuously declines, Food & Beverage first peaks at the 6% and then rapidly falls off. Conversely Machinery peaks at 4% of value added but at relatively high per capita income level.

Let us move to a parametric analysis, employing the specification proposed in Chenery and Taylor (1968) recently adopted in the deindustrialization debate (Rodrik, 2016), and try to detect whether the process of deindustrialization is affected by country income levels. Chenery and Taylor (1968), focusing on advanced countries, identify a decline in the growth of the employment share of manufacturing sectors already in the 1960s. Adding a quadratic GDP per capita term allows to detect the nonlinear negative relationship between manufacturing shares and income growth. This quadratic relationship is

³ Whenever gaps and reclassification of sectors did arise, we aggregated over some sectors.

⁴ Note that trying to detect deindustrialization aggregating different countries has some methodological caveats. It implicitly assumes that the different income levels represent various development stages of a single "representative" country.

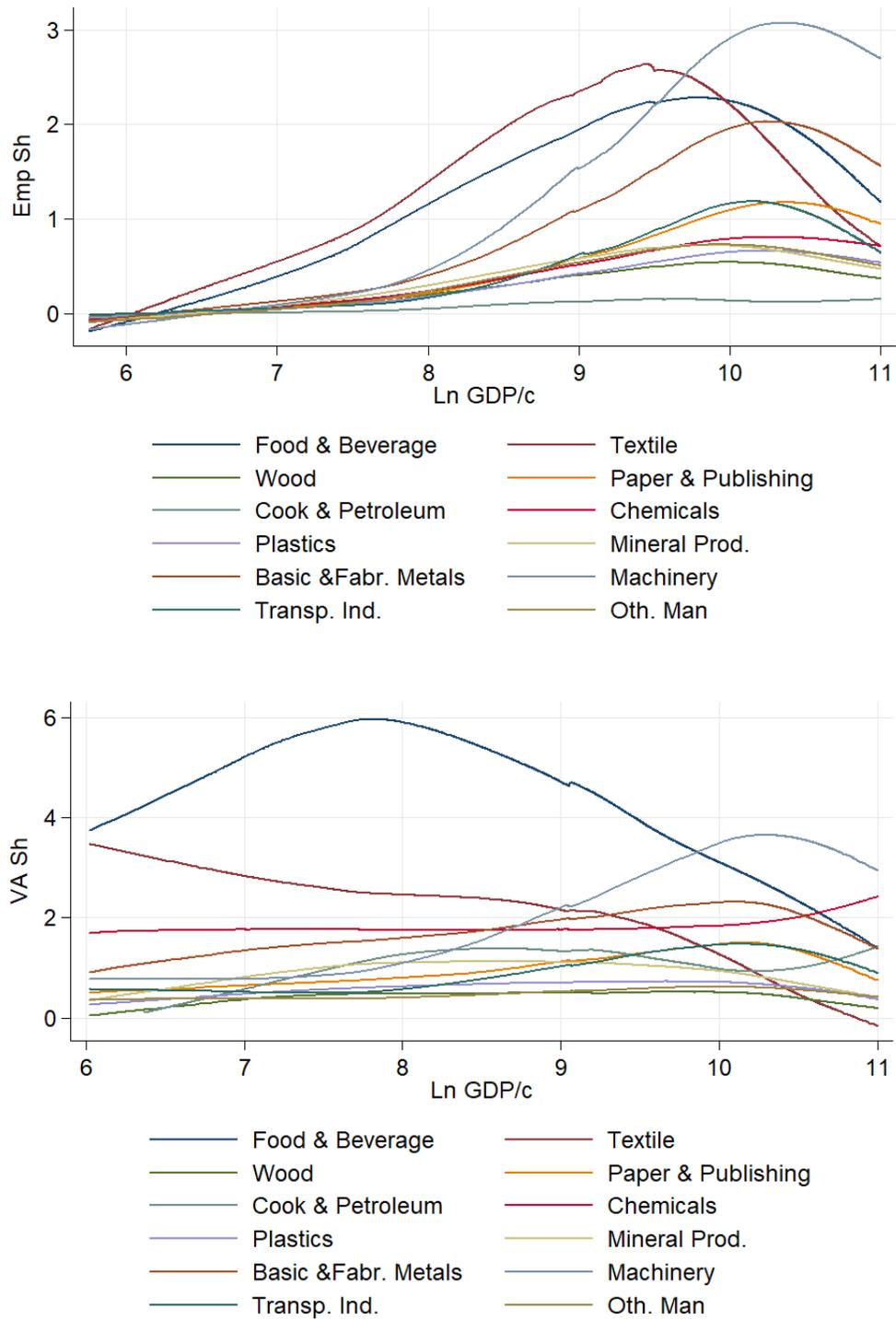


Fig. 1. Lowess fit of employment and value added shares by aggregated ISIC Rev.3 sectors.

meant to proxy changes in the consumption patterns, following the Engel curve. We employ therefore the following specification at the sectoral level by the ISIC Rev.3 classification:

$$Share_{jit} = \alpha_0 + \alpha_1 y_{i,t} + \alpha_2 y_{i,t}^2 + \alpha_3 Pop_{i,t} + \alpha_4 Pop_{i,t}^2 + \sum_{dec=1}^5 \gamma_{1,dec} D_{dec} + \sum_{i=1}^N \gamma_{2,i} C_i + \epsilon_{j,i,t} \quad (1)$$

where $Share_{jit}$ is the value added or employment share of each manufacturing sector j , of country i , at time t , $y_{i,t}$ stands for the logarithm of real GDP per capita of the country, $Pop_{i,t}$ for the logarithm of population, D_{dec} are time dummies for each decade and C_i are country fixed effects.⁵ We perform a further robustness test for the U-shape: as argued by Lind and Mehlum (2010), a positive α_1 and a negative α_2 are only necessary conditions for the emergence of a concave pattern. Thus, one needs to test the sign of the slope of the curve before and after the peak.

Table 1 shows the regression results of the panel estimation. We can identify at least three distinct sectoral varieties: some industries just present a declining path, meaning that, starting at the lowest GDP per capita levels covered by the data set, the sectors underwent a steady contraction process. The slope test confirms only the negative derivative. Some other sectors show the usually expected inverted U-shape. Finally, some sectors show a linear, or even convex, relationship between manufacturing shares and GDP per capita. While highly significant in all other manufacturing sectors, the GDP per capita term is not significant for the Machinery industry and for the residual category Furniture, Recycling & Other manufacturing activities. The sectors failing the U-shape test are chemical, machinery and the residual sector which all display a *positive* relationship. The population terms are not significant in the majority of the cases, and their exclusion does not affect the results.

Non U-shaped sectors in employment share present the same pattern as in value added. The only differences relate to the Food, Tobacco & Beverage, and to the Wood sectors showing a linear negative relationship with GDP per capita, failing the U-test.

Figure 2 plots the fitted values for the “typical country” using the estimation results of Table 1. We aggregate by averaging countries and decades fixed effects.

The heterogeneity in the patterns of deindustrialization is also reflected in different timings and shares at the peak. Figure 3 analyses the peak dynamics by industry. The red vertical line corresponds to the level of GDP per capita at the peak for the whole manufacturing sector. The horizontal line is the unweighted manufacturing share at the peak. The vertical line is the reference to detect the sectoral contribution to deindustrialization. Industries on the left-hand side of the red vertical line peaked at lower per capita income levels than the overall manufacturing industry, therefore positively contributed to deindustrialize, while those one on the right-hand side peaked at higher per capita income levels,

⁵ Haraguchi et al. (2017) introduce a cubic term of GDP/c however we did not find a significant effect for the cubic term in our sectoral regression.

Industry		y	y ²	Decade FE	Country FE	Mean Dep. Var	adj. R ²	obs	N countries	Slope at min.	Slope at max.	L-M p-value
Food, Beverage and Tobacco	emp	2,66*** (.863)	-0.144*** (.048)	Yes	Yes	1,78	0.088	4678	150	1,00***	-0.91***	0.000
	va	0.22 (2.872)	-0.06 (.148)	Yes	Yes	4,01	0.143	3858	139	-.45	-1,17	-
Textile and wearing	emp	15,11*** (4.25)	-0.868*** (.248)	Yes	Yes	1,97	0.264	4688	149	5,13***	-6,39***	0.000
	va	9.91** (4.121)	-0.57** (.226)	Yes	Yes	1,81	0.294	3872	139	3.15**	-4,00***	0.017
Wood Products	emp	1,07*** (.405)	-0.061 (.0241)	Yes	Yes	0.39	0.1	4608	148	0.37***	-0.44**	0.010
	va	0.59 (.565)	-0.04 (.029)	Yes	Yes	0.463	0.056	3772	137	0.143	-0.33**	0.260
Paper, Publishing and Printing	emp	2,52*** (.530)	-0.133*** (.029)	Yes	Yes	0.66	0.307	4667	149	0.98***	-0.79***	0.000
	va	3.31*** (.815)	-0.17*** (.044)	Yes	Yes	1,02	0.192	3828	138	1.30***	-0.83***	0.001
Coke and Petroleum Industry	emp	-0.03 (.176)	0.002 (.010)	Yes	Yes	0.12	0.063	3679	126	-0.01	0.03	0.460
	va	2.75 (1.977)	-0.17 (.111)	Yes	Yes	1,17	0.035	2904	111	.72	-1.42**	0.148
Chemicals	emp	0.97*** (.321)	-0.046** (.018)	Yes	Yes	0.54	0.268	4597	148	0.43***	-0.18	0.100
	va	-2.42 (-2.499)	0.16 (.147)	Yes	Yes	1,63	0.082	3735	139	-0.49	1,56*	0.262
Rubber and Plastic Industry	emp	1,51**** (.510)	-0.085*** (.031)	Yes	Yes	0.43	0.066	4371	142	0.53***	-0.61**	0.010
	va	1,78*** (.588)	-0.10*** (.030)	Yes	Yes	0.642	0.049	3507	129	0.64***	-0.55***	0.004
Oth. non-metallic products	emp	0.86*** (.307)	-0.041*** (.017)	Yes	Yes	0.55	0.345	4634	148	0.39***	-0.16	0.130
	va	2.45*** (.552)	-0.13*** (.030)	Yes	Yes	.924	0.123	3784	136	0.94***	-0.66***	0.000
Basic and Fabricated Metals	emp	3,40*** (.932)	-0.170*** (.053)	Yes	Yes	1,2	0.177	4680	150	1,44***	-0.82**	0.030
	va	5.44*** (1.229)	-0.28*** (.066)	Yes	Yes	1,80	0.156	3825	139	2.11***	-1.42***	0.000
Machinery, Scientific Eq. and ICT	emp	2,52* (-1.767)	-0.079 (.102)	Yes	Yes	1,72	0.193	4474	146	1,61***	0.57	-
	va	3.12 (2.285)	-0.08 (.124)	Yes	Yes	2,24	0.148	3585	134	2.16	1,14	-
Transp. Industry	emp	2,39*** (.735)	-0.116*** (.040)	Yes	Yes	0.68	0.195	4283	143	1,06***	-0.48*	0.050
	va	3.17*** (1.128)	-0.15*** (.057)	Yes	Yes	1,02	0.133	3468	128	1.39***	-.50*	0.084
Furniture, Recycling and Oth Man	emp	0.3 (.623)	-0.012 (.036)	Yes	Yes	0.52	0.068	4524	148	0.16	-0.010	0.490
	va	1.24* (.703)	-0.06 (.038)	Yes	Yes	0.540	0.021	3714	138	0.51**	-0.26	0.143

Table 1. Regression results for Eq. (1). Standard error in parentheses, clustered at the country level. *p<0.1, ** p<0.05, *** p<0.01. Industry level employment and value added data are taken from UNIDO-INDSTAT2 (see Appendix A), and data on GDP per capita and total employment from the Penn World Table 9.0. Mean stands for the mean value of the dependent variable. Decade dummies are for 1960s, 1970s, 1980s, 1990s, 2000s, 2010s. Population terms are included but not reported. y stands for GDP per capita. L.M. p-values is the outcome of the Lind and Mehlum U-test.

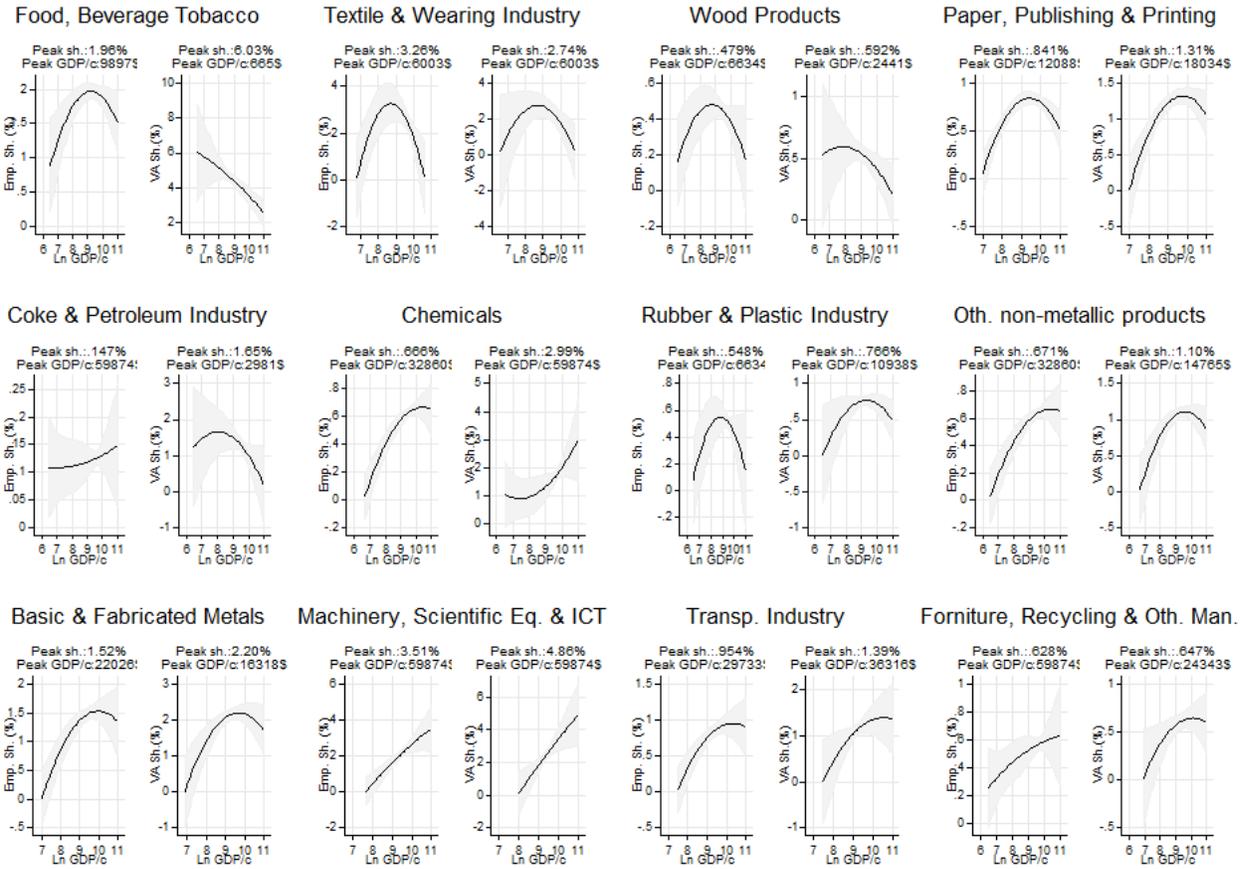


Fig. 2. Predicted employment and value added shares by industry classification from regression in Eq. (1). Country fixed effect is averaged and population size is set to the sample average. Shaded areas are 95% C.I..

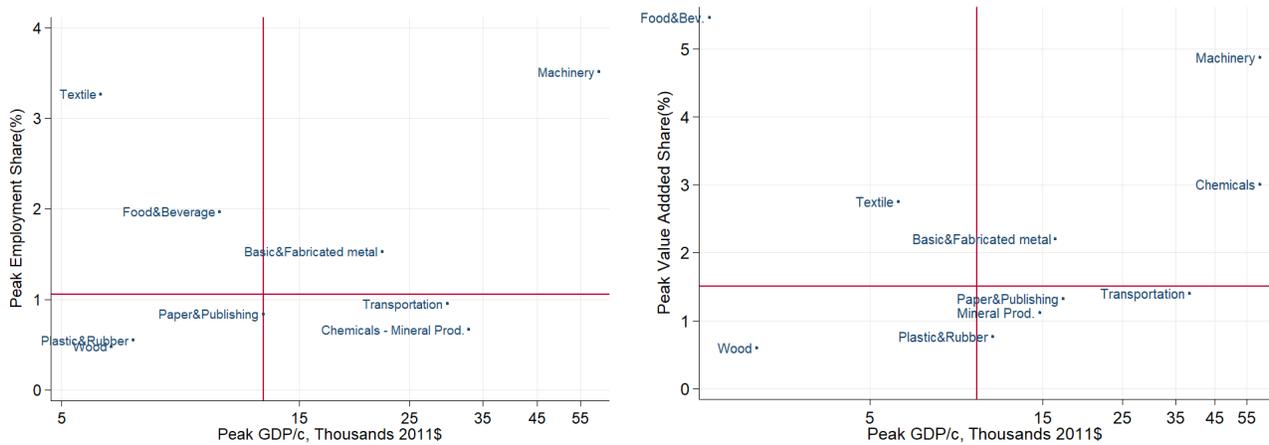


Fig. 3. Peak analysis by industrial sector. Predicted peaks and correspondent GDP per capita at the turning point for the typical country, derived from regression in Eq. (1). The red vertical line corresponds to the level of GDP per capita at the peak for the whole manufacturing sector. The red horizontal line is the unweighted manufacturing share at the peak.

therefore delaying deindustrialization. Sectors positioned above the horizontal line have bigger relative weights with respect to the others, and vice-versa. Conditional on per capita income levels, higher share “delaying” industries in terms of deindustrialization timing are located in the north-east quadrant such as Machinery and Metals Production, while lower share “delaying” sectors are in the south-east quadrant. Large “accelerating” sectors, as Textile and Food & Beverage are in the north-west quadrant while smaller ones are in the south-west quadrant (e.g. Wood, Plastic & Rubber). Note that industries at the far right (e.g. Machinery and Chemicals) are those displaying a convex relationship with GDP per capita and consequently, they do not reach a peak at all.

3.2 Technological patterns and learning regimes

Having documented the heterogeneity across sectors, let us ask whether such divergent patterns are driven by diversities in the underlying technological and learning regimes. We do that by resorting to the four classes within manufacturing identified by Pavitt (1984).⁶ The characteristics of the classes can be summarised as follows:⁷

- *Suppliers Dominated* industries (SD), wherein innovation is mainly driven by exogenous change in intermediate capital inputs and learning largely entails learning-by-using.
- *Scale Intensive* industries (SI), whose innovative abilities derive both from technological adoption of capital inputs but also by the capability to internally develop complex products and managing complex organizations. The learning is cumulative and its effects reinforced by economies of scale.
- *Specialised Suppliers* industries (SS), which provide capital equipment, instruments, components to a wide range of “downstream” industries. Learning is based on innovative efforts via formal R&D expenses as well as on tacit knowledge about artefacts design and users requirements.
- *Science-based* industries (SB) are those one whose technological advances are strongly related to the ones in basic and applied research. They have often contacts with research laboratories and the learning rates are typically quite high.

Circumstantial supportive evidence on the underlying technological differences (inevitably mixed with appropriability conditions) of the four Pavitt classes is presented in Figure 4. It shows the time series of the median 5-years cumulative patent count per worker by Pavitt classes for high-income countries vs medium- and low-income ones. The different learning regimes behind the Pavitt taxonomy are highlighted by the Science Based sector largely dominating the others in terms of patent intensity. There appears also a remarkable heterogeneity in the trend evolution of the four aggregates. Although the time patterns look somewhat similar, the scale is rather different when comparing high- and medium/low-income countries, and *technology gaps* do not appear to be closing.

⁶ The list of sectors grouped by Pavitt classes is presented in Table A2 of the Appendix.

⁷ There is a caveat to consider. Pavitt himself acknowledged that the 2-digit disaggregation is not enough to disentangle the different innovation patterns going on. This limitation is even more restrictive if both advanced and developing countries are considered together for a protracted period, as the processes of technological learning may differ.

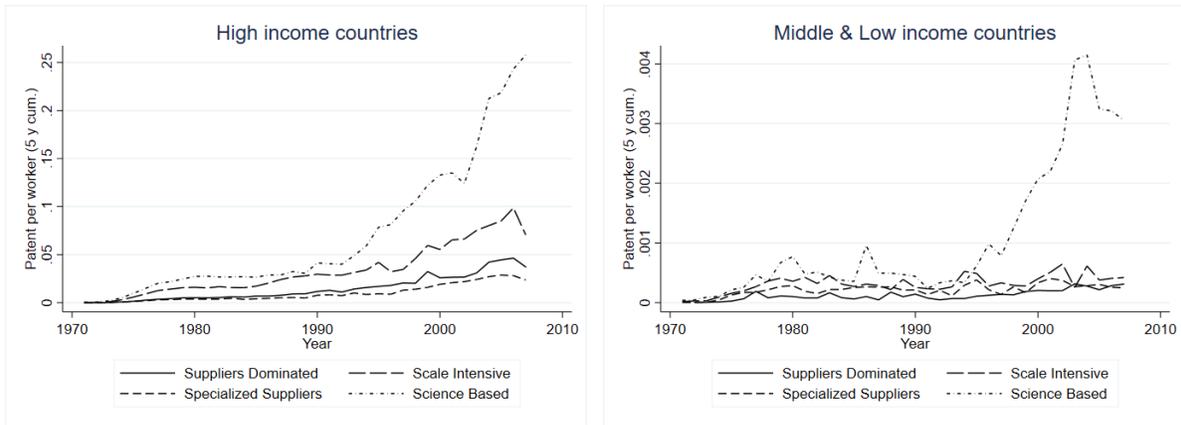


Fig. 4. Time series of the median 5-years cumulative patent count per worker by Pavitt classes for high-income countries and medium and low income countries. Patent data by IPC classes comes from the EPO-PATSTAT database and then converted into Rev. 3.1 ISIC class using the probabilistic algorithm proposed by Lybbert and Zolas (2014). Income groups are defined following the World Bank Classification. Countries are classified according to their per capita gross national income in 1989, using the World Bank Atlas Method (<http://datahelpdesk.worldbank.org/knowledgebase/articles/378832-the-world-bank-atlas-method-detailed-methodology>.)

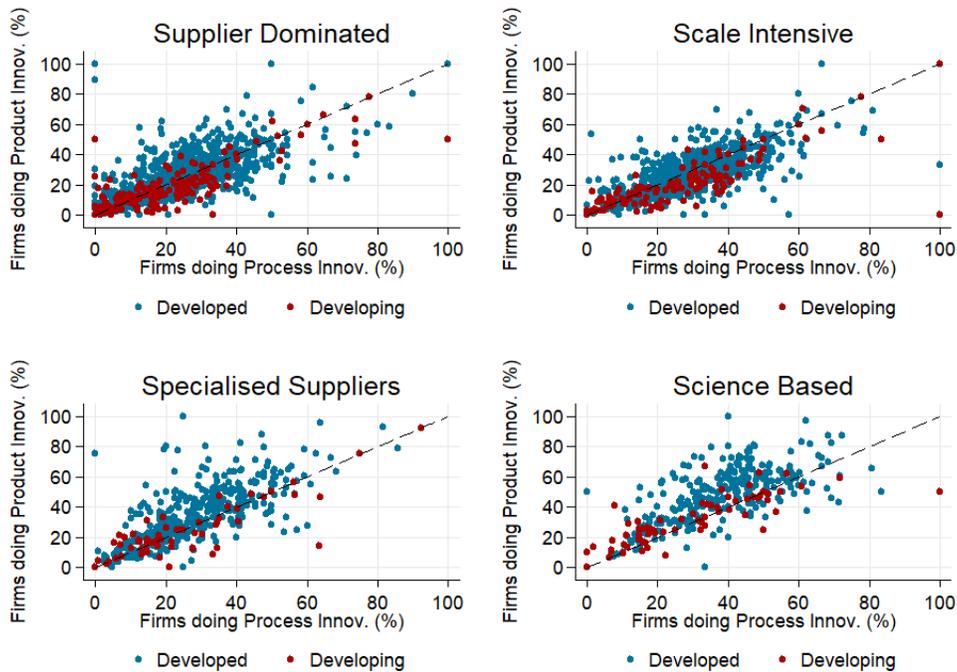


Fig. 5. Percentage of firms doing product vs the ones doing process innovation in each 2-digit sector, aggregated by Pavitt classes. Source: UNESCO Institute for Statistics, UIS online database (from 2005 to 2014). Ratios of firms doing product over process innovation - SD: 1.04; SI: 0.92; SS: 1.12; SB: 1.18 (developing countries) and SD: 1.09; SI: 1.02; SS: 1.30; SB: 1.32 (developed countries).

Another proxy for technological activities is captured by (self-reported) efforts toward product and process innovation by firms in each class, presented in Figure 5. Plausibly in the Pavitt taxonomy,

“downstream” industrial sectors as Supplier Dominated and Scale Intensive should present more complementarity between product and process innovation (or dominance of the latter in Supplier Dominated ones) as compared to the “upstream” sectors such as Science Based and Specialised Suppliers wherein innovation is largely based on the introduction of new products. The inter-sectoral comparison confirms the significant differences in the intensity of innovative activities, as well as the gap in such efforts between the Global North and the Global South. Finally, it also emerges an international division of generation of innovative activities which sees the prevalence in developing countries of patenting firms in the downstream aggregate wherein, as discussed above, learning intensities are lower.

	(1)		(2)		(3)		(4)	
	Supplier Dominated		Scale Intensive		Specialised Suppliers		Science Based	
	<i>Emp</i>	<i>Va</i>	<i>Emp</i>	<i>Va</i>	<i>Emp</i>	<i>Va</i>	<i>Emp</i>	<i>Va</i>
Ln. Gdp p/c	16.84*** (-4.506)	12.55** (-4.896)	8.64*** (-2.155)	17.17*** (-3.367)	3.22** (-1.319)	6.34*** (-1.826)	-1.78 (-1.211)	-6.77 (-4.402)
Sqr. Ln. Gdp p/c	-0.93*** (0.252)	-0.75*** (0.266)	-0.47*** (0.123)	-0.92*** (0.187)	-0.15** (0.074)	-0.32*** (0.098)	0.14* (0.077)	0.47* (0.265)
Ln. Pop	4.61 (-5.860)	-0.19 (-6.486)	1.22 (-2.908)	-0.15 (-3.762)	2.52** (-1.219)	2.44 (-1.833)	1.54 (0.991)	2.21 (-3.905)
Sqr Ln Pop	-0.12 (0.189)	0.00 (0.203)	0.00 (0.087)	0.06 (0.120)	-0.04 (0.037)	-0.05 (0.057)	-0.03 (0.029)	-0.01 (0.118)
Obs.	5510	3,897	5545	3,877	5075	3,572	5327	3,798
R-squared	0.140	0.314	0.321	0.242	0.190	0.186	0.168	0.202
N of countries	151	139	151	139	150	134	151	139
Country FE	YES	YES	YES	YES	YES	YES	YES	YES
Mean Sh.	4.43 %	7.59 %	2.18 %	5.05 %	1.26 %	1.83 %	.719 %	2.13 %
Slope at Min	6.104 ***	3.69**	3.278 ***	6.277***	1.45***	2.537***	-1.160	-1.206
Slope at Max	-6.28***	-5.679***	-2.91***	-5.243***	-.575	-1.483**	1.70***	4.68**
Lind-Mehlum p-value	.000	.020	0.000	.000	0.15	.011	0.318	.178

Table 2. Regression results for Eq. (1). Standard error in parentheses, clustered at the country level. *p<0.1, ** p<0.05, *** p<0.01. Industry level employment and value added data from UNIDO-INDSTAT2, and data on GDP per capita and total employment from the Penn World Table 9.0. Mean stands for the mean value of the dependent variable. Decade dummies are for 1960s, 1970s, 1980s, 1990s, 2000s, 2010s.

Let us now address the impact of the foregoing inter-technological and inter-country differences in the processes of industrialization (or failure of) upon the subsequent deindustrialization. We therefore aggregate employment and value added manufacturing shares by countries according to the four Pavitt classes averaging per each decade (see Figure 6). Regression results are presented in Table 3.2.

A first rather striking result is that not all classes show the typical inverse U-shape. Science Based sectors stand out as a major exception. The “canonic” shape, confirmed by the test, is corroborated only for Supplier Dominated and Scale Intensive sectors. Conversely, the test rejects the quadratic relationship between employment and value added shares and GDP per capita for both Science Based and Specialised Suppliers sectors.

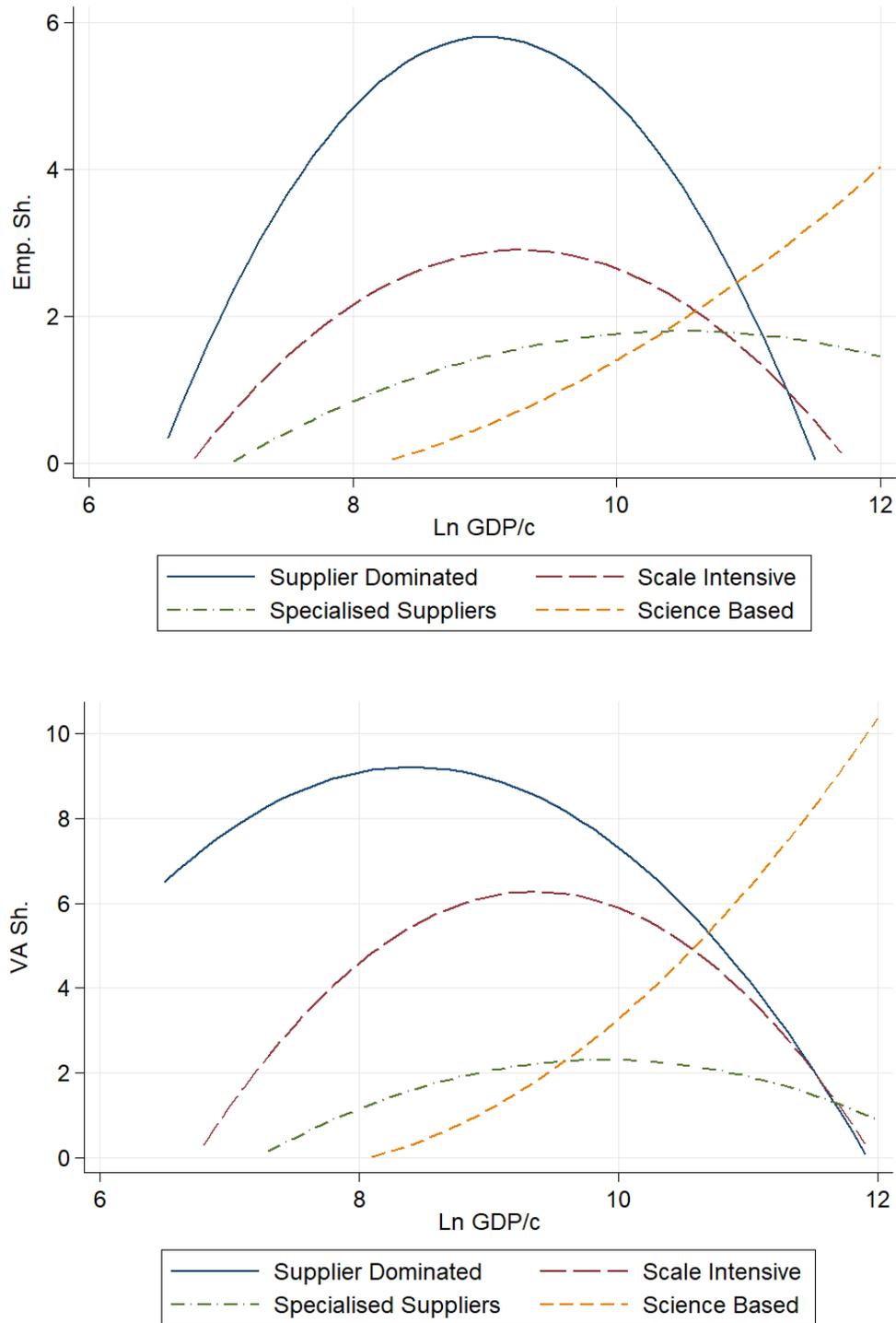


Fig. 6. Predicted employment and value added shares by Pavitt classes from regression in Eq. (1). Country fixed effects are averaged and population size is set to the sample average.

The overall increasing trend of Science Based is even more marked for value added shares, while at the opposite end of the spectrum the Supplier Dominated class shows an ever declining pattern. After reaching a peak of 9.2% in a very early phase of development (4500\$/c), the class continuously reduces its shares. Looking at Scale Intensive, the class expands up to a peak at 12000\$/c, while Specialised Suppliers keep growing till around 20000\$/c, presenting a far-less steeper slope.

Reading the evidence in terms of stages of economic development, Supplier Dominated industries represent the first stage of industrialization in line with a very robust historical stylised fact, with the textile/clothing sector largely contributing to the initial process of structural change, absorbing workers from agriculture. Overall, this class absorbs the largest fraction of employment share with a peak at 6%. However, Supplier Dominated is also the first class experiencing a decline, with deindustrialization starting at 8000\$/c. Scale Intensive sectors, with a peak at 3% of employment share, follow a similar pattern, although emerging at a later stage in the development process, well after Supplier Dominated took off. Any explicit mark of deindustrialization is hard to detect in Specialised Suppliers industries, as they reach the maximum expansion only in an advanced phase of development. Finally Science Based industries not only do not show signs of deindustrialization, but keep increasing the share of employed workers as per capita income grows. The sequence of expansion militates against the Rostow (1991) mechanical phases of economic growth.

Notably, by disaggregating into the four classes it clearly emerges the inverse relationship between the speed of deindustrialization and the relative employment share of each class. In fact, Supplier Dominated and Scale Intensive sectors absorb the largest fraction of the labour force and start to decline first. Indeed, a potential fallacy of composition might arise if one just look at the aggregate. The overall deindustrialization pattern derives therefore from the faster declining trend of the downstream aggregate vis-à-vis the constant/increasing trend of the upstream one.

In Figure 7 we present the changes in the composition of the overall manufacturing industry from 1971 to 2011⁸ both in terms of value added and employment. The variations represent reallocation of shares across Pavitt classes over the last forty years, for both developing and developed countries. In line with the previous analysis, the Supplier Dominated class records a massive reduction of its manufacturing shares, both in developed and even more so in developing countries. Conversely, the Scale Intensive class presents a more blurred pattern: while both employment and value added shares come down in developed countries, the latter largely increases in developing ones while employment does not vary. This is an indirect signal of the process of international relocation of scale intensive sectors to the South/East of the world with however small effects in terms of employment. Specialised Suppliers and Science Based activities grow both in terms of value added and of employment shares. Notably, the contribution of the Science Based class, particularly in developed countries, is positive in both em-

⁸ The time span has been chosen to cover most countries included in the analysis, assuring consistency between employment and value added. The precise time span does not affect the results.

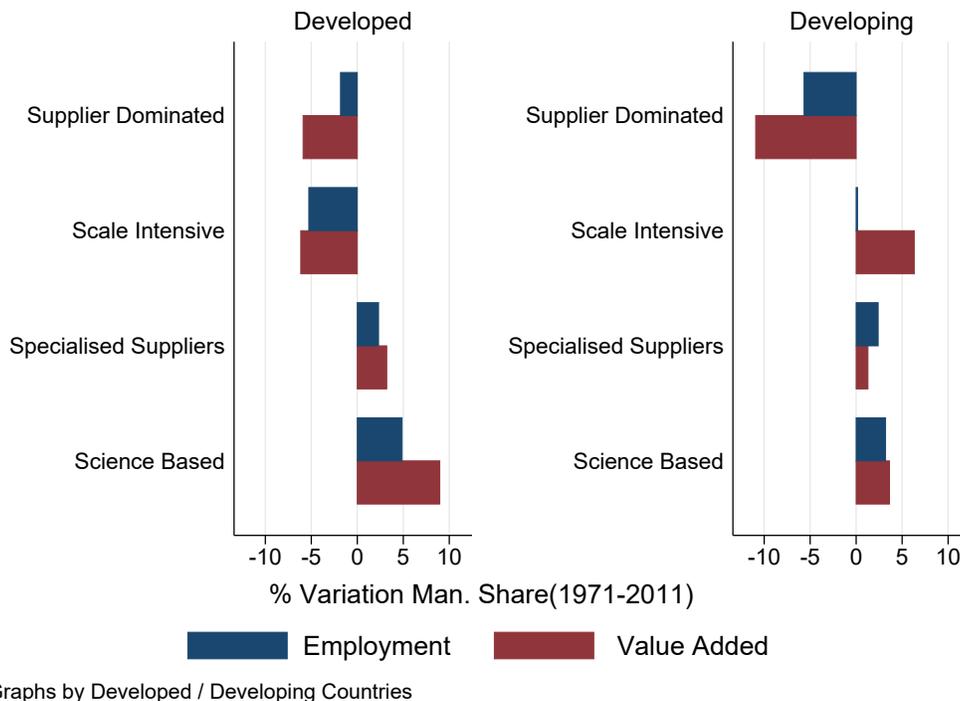


Fig. 7. Mean variation of employment and value added shares within the manufacturing sector in the period 1971-2011. Industry level value added data are taken from UNIDO-INDSTAT2.

ployment and value added. Overall, Specialised Suppliers and Science Based industries represent the employment-absorbing and value added-generating classes for both developing and developed countries (supposing a constant labour force participation in the labour market).

These results of course question any natural tendency hypothesis according to which deindustrialization is the simple result of income growth and technical progress. First of all, the very fact that value added and employment shares do not have a clear anti-correlated pattern, above some threshold, militates against the natural tendency story. Second, some highly innovative sectors, often belonging to the “upstream” aggregate, keep non reducing or even increasing employment shares as income grow.

4 Timing heterogeneity and premature deindustrialization

What about heterogeneity in the *timing of deindustrialization*? Let us consider it by disaggregating into the four Pavitt classes, and focusing on the sectoral contribution of each class to the aggregate premature deindustrialization process, splitting the analysis between pre- and post-1990 patterns.

In order to study such patterns, one needs to dissect the effect of time as such upon the shares dynamics. All technological classes experienced a downward time drift from 1960 to 2010. Supplier Dominated and Scale Intensive sectors display the most marked time-effect in both employment and value added. However there appears to be a remarkable heterogeneity across country-industry on each decade. The

downward shift is significantly different from zero only after the 1990s. Conversely, Science Based and Specialised Suppliers display a negligible temporal pattern accelerating in the last two decades.

Simple time regression does not allow to capture either heterogeneous shifts across per capita income or “regime shifts”. Following Rodrik (2016), we estimate Eq.(2), introducing a dummy variable for the ex-post Berlin Wall fall:

$$\begin{aligned}
 Share_{j,i,t} = & \alpha_0 + \alpha_1 y_{i,t} + \alpha_2 y_{i,t}^2 + \alpha_3 Pop_{i,t} + \alpha_4 Pop_{i,t}^2 + \\
 & + \alpha_5 y_{i,t} \# Period + \alpha_6 y_{i,t}^2 \# Period + \sum_{i=1}^N \gamma C_i + \epsilon_{j,i,t}
 \end{aligned} \tag{2}$$

Using a structural shift-type approach, the period dummy is interacted with GDP per capita. The interaction term allows to infer the impact of trade openness and increasing globalization for manufacturing shares, at different per capita income levels. Table 4 presents the regression results for Eq. (2). The interaction terms allow for a different patterns (U-shape or not) for the pre- and post-1990 periods beyond just a vertical shift, as in the case of decade dummies or year fixed effects. A visual analysis is presented in Figure 8 which shows the predicted relationship between GDP per capita and employment shares for the four Pavitt classes in the periods before and after 1990.

Interestingly, the classes most impacted by the interaction term are Scale Intensive and Specialised Suppliers sectors.⁹ Recall however that the two classes contribute differently to the overall deindustrialization process, with the former presenting a much steeper decline vis-à-vis the latter. The Supplier Dominated sector, which is the dominant deindustrializing class, on the contrary shows little signs of premature deindustrialization, with the peak share declining less than 0.5% in the post-1990 period. Finally, the Science Based class does not display any effect of the post-1990 dummy which also results to be non significant in the regression analysis. The two curves almost overlap. The patterns remain unchanged for value added although different in the scale.

In Figure 9 we present for each class the contribution to premature deindustrialization. Both in terms of value added and employment shares, Scale Intensive and Specialised Suppliers are those most affected by the globalization era (with a drop of more than 50% after 1990). Suppliers Dominated sectors are much less affected, while the only class recording a positive shift is the Science Based one.

It is important to notice that Figure 9 displays average effects, in turn hiding major inter-country differences in primis the striking emergence of China as the *World Factory* and the dismal performance of Latin American countries. Overall, the revealed patterns of premature deindustrialization in fact highlight missing catching-up opportunities for a large share of developing countries which after 1990 faced much greater difficulties in industrializing. And such opportunities shrink particularly in those two classes nursing relatively complex learning capabilities – Scale Intensive and Specialised Suppliers –

⁹ Note that the regression analysis gives an average effect on the peak share, however it does not capture the entire area between the two graphs. The visual inspection allows to better grasp the overall impact.

	(1)		(2)		(3)		(4)	
	Supplier	Dominated	Scale	Intensive	Specialised	Suppliers	Science	Based
	Emp	Va	Emp	Va	Emp	Va	Emp	Va
Ln. Gdp/c	15.89*** (3.672)	13.04*** (4.461)	7.75*** (1.618)	15.33*** (3.034)	2.90** (1.150)	5.53*** (1.421)	-0.63 (1.003)	-4.95 (3.729)
Sqr. Ln. Gdp/c	-0.89*** (0.197)	-0.81*** (0.235)	-0.42*** (0.090)	-0.83*** (0.163)	-0.14** (0.065)	-0.28*** (0.071)	0.06 (0.061)	0.33 (0.217)
Ln. Pop	7.27 (5.650)	2.01 (6.658)	4.41** (2.200)	3.11 (3.943)	3.55*** (1.224)	4.74** (1.872)	1.38 (0.935)	1.78 (3.675)
Sqr Ln Pop	-0.24 (0.184)	-0.11 (0.206)	-0.14* (0.069)	-0.09 (0.127)	-0.10** (0.039)	-0.15** (0.060)	-0.04 (0.029)	-0.02 (0.115)
Ln GDP/c	0.49** (0.240)	0.16 (0.263)	0.46*** (0.089)	0.56*** (0.166)	0.21** (0.088)	0.29*** (0.101)	-0.06 (0.039)	-0.11 (0.139)
Sqr Ln GDP/c	-0.06** (0.026)	-0.03 (0.026)	-0.06*** (0.010)	-0.07*** (0.017)	-0.03*** (0.010)	-0.04*** (0.010)	0.01 (0.004)	0.01 (0.014)
Observations	5,510	3,897	5,545	3,877	5,075	3,572	5,327	3,798
R-squared	0.131	0.295	0.327	0.228	0.147	0.176	0.111	0.164
Number of id	151	139	151	139	150	134	151	139
Country FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Mean Sh.	4.43 %	7.59 %	2.18 %	5.05 %	1.26 %	1.83 %	.719 %	2.13 %

Table 3. Regression results for Eq. (2). Standard error in parentheses, clustered at the country level. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$. Industry level employment and value added data are taken from UNIDO-INDSTAT2, and data on GDP per capita and total employment from the Penn World Table 9.0.

which however do not often require interactions with “frontier science”. In a way, the post 1990 period, call it globalization phase, has favoured a polarization in terms of technological specialization, with leading advanced countries dominating Science Based industries and low income countries barely maintaining their shares in the Supplier Dominated ones.

5 Patterns of diversification

The process of economic development historically can be seen as an upgrading path over technological capabilities and learning regimes. If the textile sector has typically represented the opportunity to discipline and organise a previously unstructured labour force, steel and heavy metal industries often have been the chance to accumulate massive productive capacity and capital equipment. Finally, information and communication technologies, fine chemistry and drugs entailed a tighter connection between basic and applied research. Thus, the productive structure and the ensuing specialization and diversification patterns at any stage of development strongly influence the development process itself.

In order to capture the diversification profile of different countries across Pavitt classes, we perform a cluster analysis. The analysis allows to classify countries according to their industrial composition in terms of manufacturing employment shares. The cluster analysis allows to group and separate similar data objects. Among alternative cluster models, the K-means algorithm is an iterative method to partition the dataset into k non-overlapping subgroups. In our analysis, the clustered variable is the country

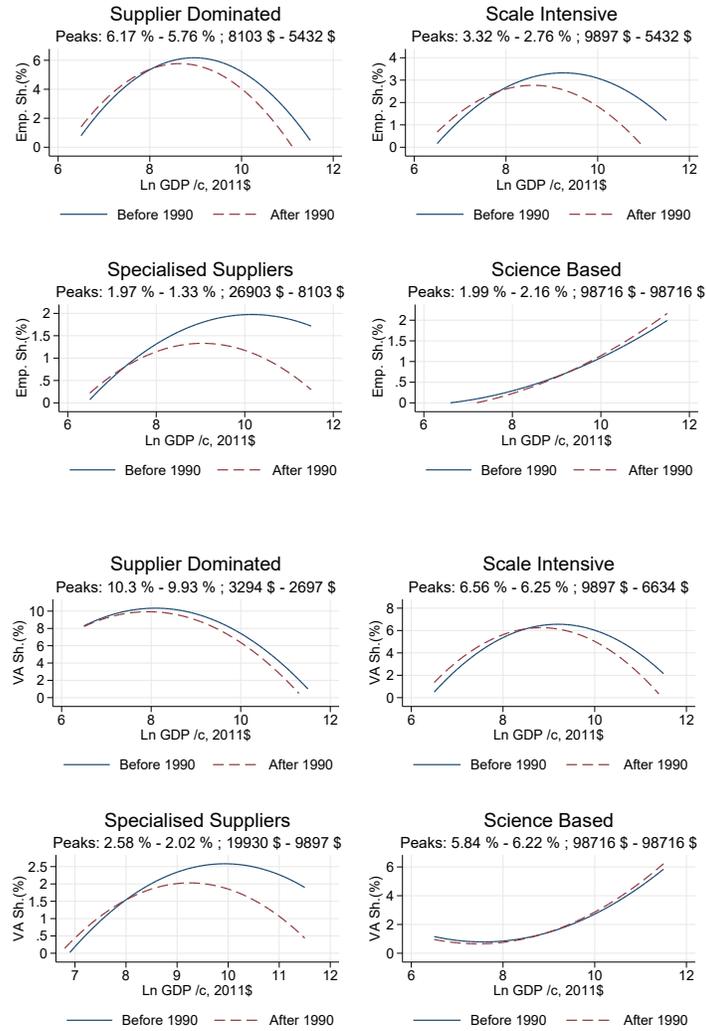


Fig. 8. Predicted employment and value added share by Pavitt classes from regression in Eq. (2). Average country fixed effect and average population size. Each subtitle indicates employment shares and the correspondent GDP per capita at the maximum in the two periods.

manufacturing employment share in each of the four Pavitt classes, averaged by decades, excluding petroleum-producing countries and those countries with less than one millions of inhabitants.

The K-means cluster analysis is an unsupervised learning algorithm used to partition the observations in groups such that each data point belongs to the cluster with “most similar” features.¹⁰ The iterative procedure starts with the extraction of k random observations that are used as centroids to group the data. Then, the algorithm minimizes the sum of the squared distance between the data points and the cluster’s centroid (i.e. the intra-cluster variance) and assigns each observation to the closest cluster. Finally, the centroids are updated using the new partitioned observations. The procedure is iterated

¹⁰ Appendix D contains a detailed description of the methodology used in this exercise.

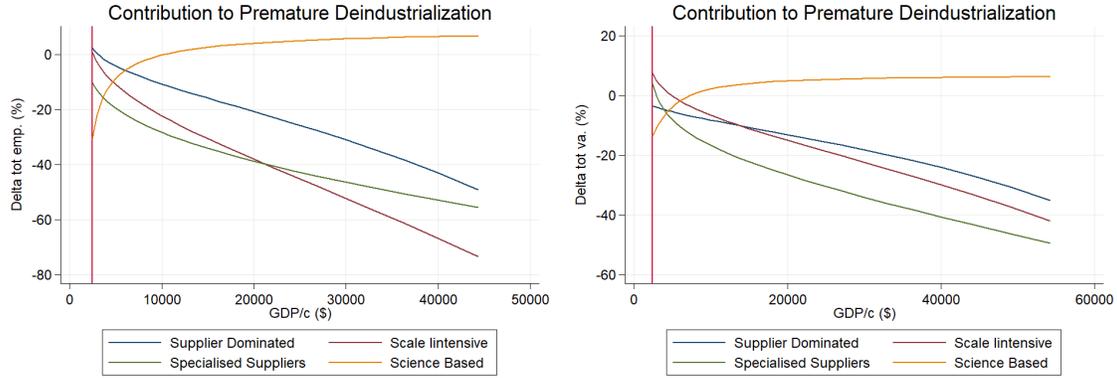


Fig. 9. Pavitt classes' contribution to premature deindustrialization. Those patterns derives from Eq. 2 in previous section and refer to "typical country" premature deindustrialization experience. The vertical red line stands for the level of GDP/c at the aggregate manufacturing peak.

until the partition stabilises and the centroids stop changing. To avoid local optima, we perform 1000 k-means clusters and we take the one with the smallest within cluster sum of squares (WSS). Based on the scree plots presented in Figure A3 in the Appendix, we select four clusters which are sufficient to explain more than eighty percent of the intra-cluster variance.

The results of the analysis are presented in Figure 10. Each box plot shows the distribution of employment shares across the four classes emerging in each cluster. We characterize the clusters in terms of degree of concentration. Starting from the most concentrated cluster, countries belonging to the bottom-right quadrant have an average industrial composition extremely concentrated in the Supplier Dominated class, absorbing 85% of the labour force. From the right to the left, the "Concentrated" cluster is populated by countries having a relatively lower fraction of employment shares in Supplier Dominated industries and a higher fraction in Scale Intensive ones. "Moderately diversified" countries present more equally distributed employment shares across the four classes, with Specialised Suppliers and Science Based recording around one tenth of the labour force. Finally, "Diversified" countries are those where the maximum share of Supplier Dominated is 40% while the minimum share of Science Based industries is 15%.

Figure 11 displays the time trajectories (up) and the evolution of the entropy index (bottom) of four selected countries belonging to the four clusters and starting with the same GDP per capita in 1970, namely Korea, Malaysia, New Zealand, Kenya. Note that the trajectories of per capita income diverge over time, at increasing rates. The bottom part of the figure shows the evolution of the corresponding entropy index. The latter, used to infer diversity in a given population, indicates the proportional abundance of the same Pavitt class. The four countries, with very similar entropy indices in 1970, again diverge in their diversification trajectories, with Korea and Malaysia increasing diversification within their industrial composition, while at the opposite extreme Kenya reducing it.

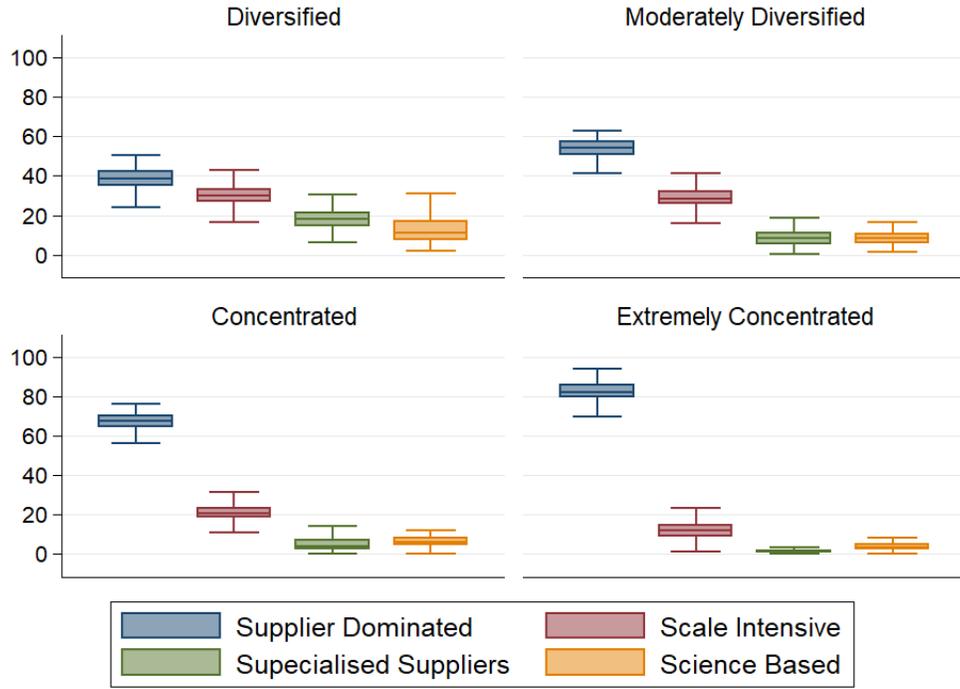


Fig. 10. Box plots of the manufacturing employment shares across Pavitt classes for each cluster.

Movements of the countries across clusters are expected as time goes by. If countries change their industrial composition they can either climbing the ladder or falling behind. Therefore, we study the one-step Markov transition probability for each country, that is the probability to move from cluster i to cluster j in the subsequent decade. The transition probability matrix (P) consists in the collection of one-step transition probabilities ($p_{i,j}$) for all combinations of states in the system (i, j):

$$p_{i,j} = Pr\{X_t = j | X_{t-1} = i\} \quad (3)$$

The final transition probability matrix (P) records the states (clusters) of the system at time t by column, while the states at time $t + 1$ by row. The main diagonal of the transition probability matrix reports the probabilities of remaining in the same cluster from one time step to the subsequent one. The upper triangular matrix represents the probabilities of moving backward to clusters with higher concentration level. The lower triangular matrix shows the probabilities of moving upward in the diversification ladder.

Figure 12 shows the one-step transition probability matrix before and after 1990. Throughout the period, the highest probability outcome is to remain in the same cluster (darker main diagonal), the second most probable outcome is climbing up with a minimum of 12% probability before 1990. In general, the transition probabilities go down moving away from the diagonal so that big jumps are quite improbable. The results somewhat change in the post 1990 period. Although the most probable outcome

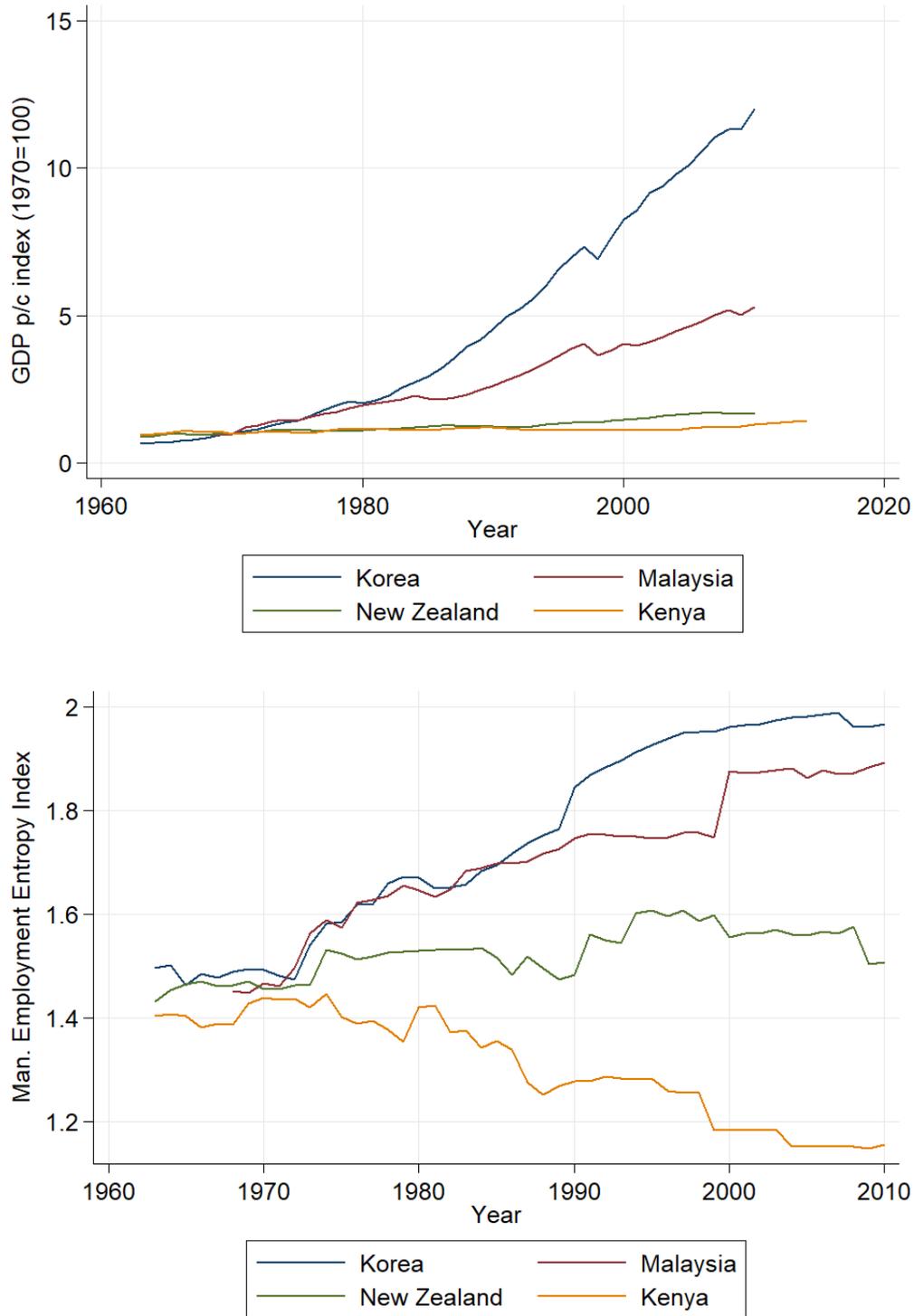


Fig. 11. Gdp per capita (top) and entropy index (bottom) of four selected countries. Base-year index 1970.

is still remaining in the same cluster, the probability of falling behind increases (darker upper triangular matrix).

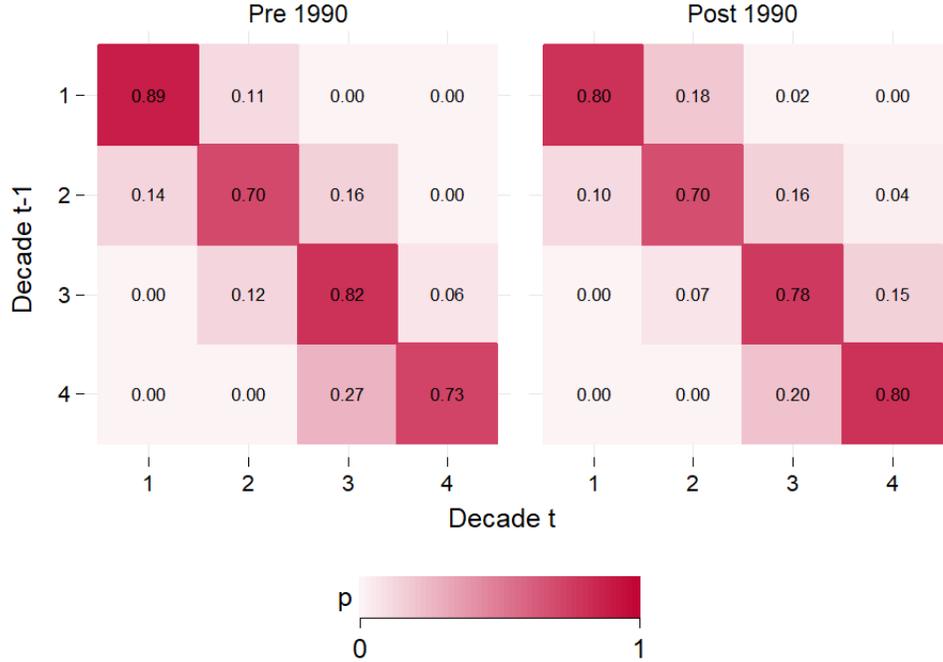


Fig. 12. Transition matrix with the probability of moving from one cluster to another in the subsequent period from 1970s to 2010s.

In order to link this evidence with the deindustrialization process, we re-estimate Eq. 1 for each cluster. Figure 13 shows the link between industrialization/deindustrialization and diversification conditional on GDP per capita for the average country-cluster. Diversified countries are those presenting the highest manufacturing share of total employment. At the opposite end, concentrated countries present remarkably lower peak shares and a premature deindustrialization process. The corresponding regression results are in Table 4.

6 Discussion and conclusions

The process of development is highly uneven. Cross-country evidence highlights a variety of industrialization/deindustrialization processes. The *institutional shock* associated with an accelerated globalization discussed by Rodrik (2016) has yielded an average acceleration in the timing and pace of deindustrialization. However, such average tendency hides ample heterogeneity driven by underlying diversities in technological capabilities and learning regimes. We try to account for them employing the Pavitt taxonomy of sectors/technologies. Indeed, dissecting the patterns at a finer sectoral level allows the identification of drivers otherwise concealed by fallacies of composition.

First, not all technological classes deindustrialize at the same pace. Although Supplier Dominated and Scale Intensive industrial sectors are the ones absorbing the highest fraction of the labour force in the

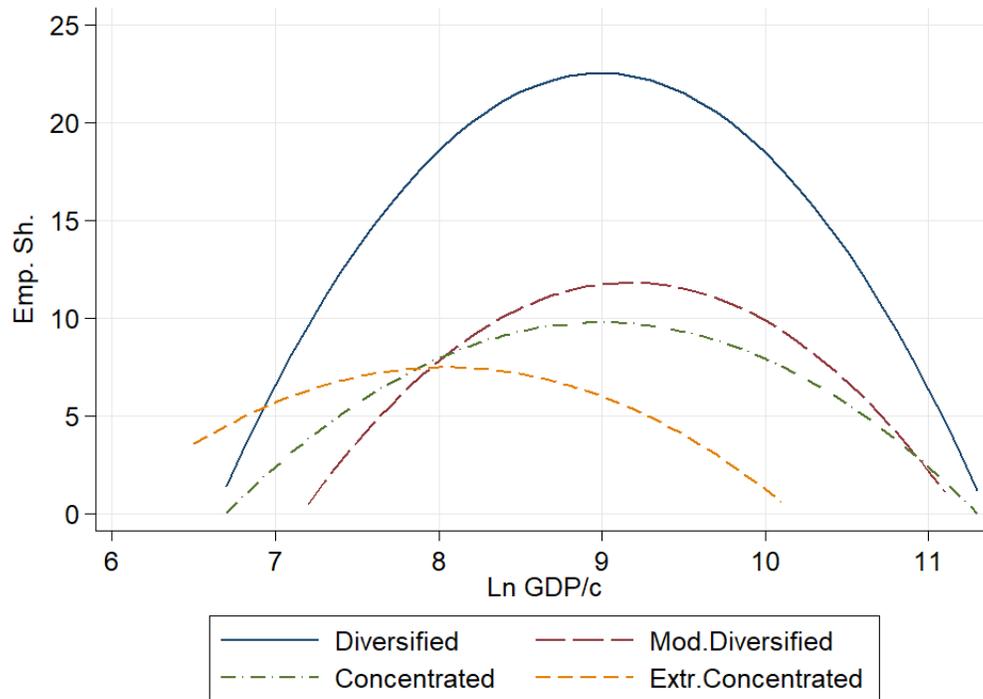


Fig. 13. "Typical country" patterns by clusters.

manufacturing, they are also the first losing shares in terms of both employment and value added. On the contrary Specialised Suppliers and Science Based ones are quite different: while the former mildly lose shares only at a late stage of the overall development process, the latter keeps increasing employment and value added shares.

These results question any natural tendency notion according to which deindustrialization is a natural process due to productivity improvements, hitting countries and sectors indifferently. Moreover, our results contradict, although indirectly, the negative employment-productivity relationships. Indeed, those sectors experiencing "anti-deindustrialization" patterns are the ones characterised by a higher degree of innovative contents, as proxied by patent intensity, but also by a higher ratio of product to process innovation. Overall, the most innovative manufacturing sectors are those experiencing increases in both employment and value added shares. However, their contribution to employment opportunities is relatively smaller vis-à-vis Supplier Dominated and Scale Intensive classes.

Second, varieties emerge in the timing and in the response to globalization. In fact, also premature deindustrialization is heterogeneous across technological classes, with those most hit being Scale Intensive and Specialised Suppliers. What emerges in the "globalization regime" is an interrupted process of industrialization for low-income countries which remain largely trapped in the (laggard) Supplier Dominated class. The reverse side of the coin is that Science Based employment and value added tend to be even more concentrated in advanced countries. Such an evidence seems also to suggest that the pur-

	(1)	(2)	(3)	(4)
	Diversified	Mod. Diversified	Concentrated	Extr. Concentrated
Ln GDP/c	72.08*** (10.357)	53.12*** (12.123)	33.13** (18.035)	26.34 (16.133)
Ln GDP/c Sqr.	-4.01*** (0.590)	-2.89*** (0.670)	-1.84* (0.993)	-1.64* (0.996)
Ln Pop.	56.56*** (17.839)	29.01 (18.964)	-18.55 (18.734)	-8.27 (20.553)
Ln Pop. Sqr.	-1.68*** (0.511)	-0.99* (0.559)	0.57 (0.585)	0.30 (0.623)
<i>Obs.</i>	1,240	1,154	1,202	632
<i>R-Sqr.</i>	0.542	0.251	0.102	0.116
<i>N. of country</i>	41	56	65	44
<i>Country FE</i>	YES	YES	YES	YES
<i>Mean Emp. Sh.</i>	16.7 %	9.62 %	8.18 %	5.70 %

Table 4. Regression results of Eq. 1 by clusters of appearance. Each country may varies cluster in each decade.

ported benefits deriving from foreign direct investments and related international knowledge spillovers promoted by globalization are somewhat missing, especially in the most innovative sectors. Overall, the post-1990 deindustrialization patterns suggest a tendency to an even more polarised and fragile international division of production. This should come as no surprise if technologies are subject to sector – and location specific – increasing returns. More globalization tends to reinforce pre-existing “comparative advantages” which might well entail traps reinforcing traps, and leaderships reinforcing leaderships (Dosi and Tranchero, 2018).

Third, we intersect technological diversification in terms of employment shares with country patterns of development. By identifying four clusters characterised by different diversification degrees, we show that countries presenting a more balanced and diversified industrial composition experience higher per capita income level. Additionally, by analysing transition probabilities we do find evidence of a stronger tendency toward falling behind threats rather than for catching-up opportunities after the acceleration of international trade and capital mobility.

Figure 14 gives a graphical representation of a successful process of diversification of climbing the ladder countries (up panel) and of a failing process of diversification of falling behind countries (bottom panel). As we have seen, dominant and failing countries are distinct both in terms of the share of each Pavitt class and in the timing and patterns of deindustrialization. By coupling the sectoral/technological composition of the manufacturing on the one hand, and the deindustrialization process on the other, we have been able to detect some structural features behind the development trajectories of the world economy.

History is not a destiny. China is the most striking example of policies successfully working against any historical omen. More generally, our analysis lends support to the promotion of a wide array of industrial and innovation policies able to guide countries toward sustained innovative growth paths

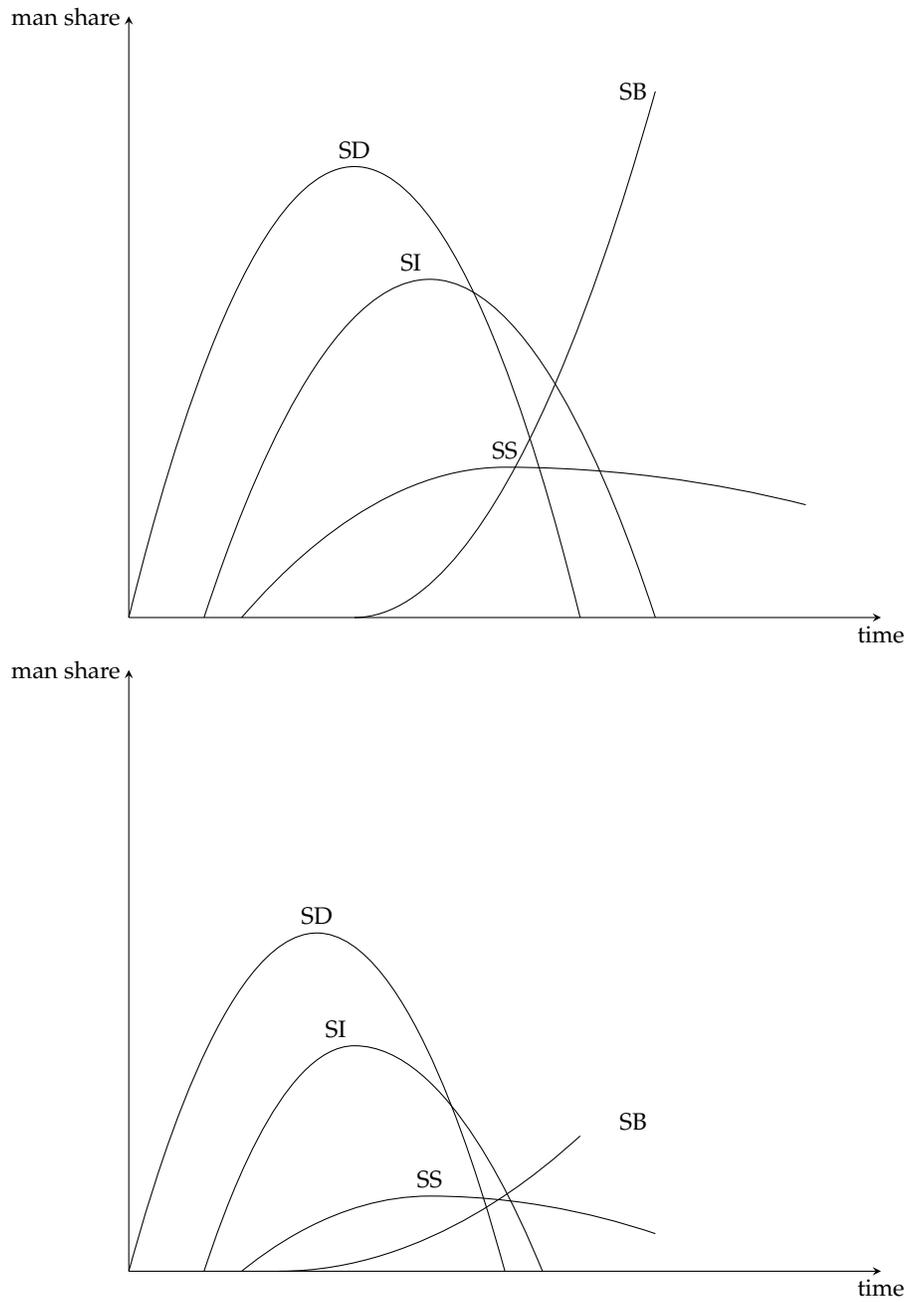


Fig. 14. Diversification patterns of climbing the ladder countries (up) and falling behind countries (down). SD: Supplier Dominated, SI: Scale Intensive, SS: Specialised Suppliers, SB: Science Based.

(Cimoli, Dosi, and Stiglitz, 2009). Otherwise, developing countries might well stick to the low and medium income traps. Nonetheless, climbing the ladder in the current phase of complex product specialization, barriers to entry and high degree of appropriability conditions is even more difficult than thirty years ago. Finally, participation in the global value chain is not a guarantee for economic growth. The stage of participation rather than the sheer access to global value chains ultimately determines prospects of growth. All in all, producing potato chips is not equivalent to producing microchips.

Industrial policies do not only pertain to developing countries. Industrial policies are urgently needed even in Europe and in the U.S. to invert an otherwise unavoidable trap of productivity slowdown, exploding inequalities and migration conflicts (Pianta, Lucchese, and Nascia, 2019).

Future research entails the analysis of the relationship between the patterns of deindustrialization and income inequalities. Indeed, deindustrialization and income distribution could not be more relevant than nowadays in a new economic phase marked by the Covid-19 pandemic, challenging the organization of global value chains and the entire division of labour and distribution of gains at the global level.

References

- Abramovitz, M. (1986). Catching up, forging ahead, and falling behind. *The Journal of Economic History* 46(2), 385–406.
- Arrow, K. J. (1962). The economic implications of learning by doing. *Review of Economic Studies* 29(3), 155–173.
- Baumol, W. (1967). Macroeconomics of unbalanced growth: the anatomy of urban crisis. *The American Economic Review* 57(3), 415–426.
- Bogliacino, F. and M. Pianta (2010). Innovation and employment: a reinvestigation using revised pavitt classes. *Research Policy* 39(6), 799–809.
- Chenery, H. and M. Syrquin (1975). *Patterns of development 1950-1970*. Oxford University Press.
- Chenery, H. and L. Taylor (1968). Development patterns among countries and over time. *Review of Economics and Statistics* 50, 391–416.
- Cimoli, M. and G. Dosi (1995). Technological paradigms, patterns of learning and development: an introductory roadmap. *Journal of Evolutionary economics* 5(3), 243–268.
- Cimoli, M., G. Dosi, and J. E. Stiglitz (2009). *Industrial Policy and Development: The Political Economy of Capabilities Accumulation*. Oxford University Press.
- Clark, C. (1957). *The conditions of economic progress*. London: Macmillan.
- Dosi, G., K. Pavitt, and L. Soete (1990). *The economics of technical change and international trade*. Pinter Publisher.
- Dosi, G., M. Piva, M. Virgillito, and M. Vivarelli (2019). Embodied and disembodied technological change: the sectoral patterns of job-creation and job-destruction. *IZA Discussion Paper, n. 12408*.
- Dosi, G. and M. Tranchero (2018). The role of comparative advantage and endowments in structural transformation. *LEM Working Paper Series, n. 33/2018*.
- Fagerberg, J. (1987). A technology gap approach to why growth rates differ. *Research policy* 16(2-4), 87–99.
- Feenstra, R. C., R. Inklaar, and M. P. Timmer (2015). The next generation of the penn world table. *American Economic Review* 105(10), 3150–3182.
- Freeman, C., J. Clark, and L. Soete (1982). *Unemployment and technical innovation*. London and New York: Pinter Publisher.
- Haraguchi, N., C. F. C. Cheng, and E. Smeets (2017). The Importance of Manufacturing in Economic Development: Has This Changed? *World Development* 93(C), 293–315.
- Haraguchi, N., B. Verspagen, and J. Amann (2019). The structural dynamics of deindustrialisation and the effect of accelerated globalisation on manufacturing. *mimeo*.
- Hastie, T., R. Tibshirani, and J. Friedman (2001). *The Elements of Statistical Learning*. Springer Series in Statistics. New York, NY, USA: Springer New York Inc.
- Hausmann, R., J. Hwang, and D. Rodrik (2007). What you export matters. *Journal of Economic Growth* 12(1), 1–25.
- Hirschman, A. O. (1958). *The Strategy of Economic Development*. Yale University Press.
- Kaldor, N. (1967). *Strategic Factors in Economic Development*. New York State School of Industrial and Labor Relations, Cornell University, Ithaca.
- Kunst, D. (2019). Premature Deindustrialization through The Lens of Occupations: Which Jobs, Why, and Where? *Tinbergen Institute Discussion Paper, TI 2019-033/V*.
- Kuznets, S. and J. T. Murphy (1966). *Modern economic growth: Rate, structure, and spread*, Volume 2. Yale University Press New Haven.
- Landesmann, M. and K. Pichelmann (1999). Employment patterns and the welfare services sector. In M. Landesmann and K. Pichelmann (Eds.), *Unemployment in Europe*, pp. 205–231. London: Macmillan.

- Lee, K. and F. Malerba (2017). Catch-up cycles and changes in industrial leadership: Windows of opportunity and responses of firms and countries in the evolution of sectoral systems. *Research Policy* 46(2), 338–351.
- Lind, J. T. and H. Mehlum (2010). With or without u? the appropriate test for a u-shaped relationship. *Oxford bulletin of economics and statistics* 72(1), 109–118.
- List, F. (1841). *Das nationale System der politischen Ökonomie*. Stuttgart/Tübingen.
- Lybbert, T. and N. Zolas (2014). Getting patents and economic data to speak to each other: An ‘algorithmic links with probabilities’ approach for joint analyses of patenting and economic activity. *Research Policy* 43(3), 530–542.
- Nuvolari, A. and E. Russo (2019). Technical progress and structural change: a long-term view. *MERIT Working Papers*, n. 22.
- Pasinetti, L. L. (1983). *Structural change and economic growth: a theoretical essay on the dynamics of the wealth of nations*. Cambridge, UK: Cambridge University Press.
- Pavitt, K. (1984). Sectoral patterns of technical change: Towards a taxonomy and a theory. *Research Policy* 13(6), 343–373.
- Pianta, M., M. Lucchese, and L. Nascia (2019). The policy space for a novel industrial policy in europe. *Industrial and Corporate Change*.
- Prebisch, R. (1959). Commercial policy in the underdeveloped countries. *The American Economic Review* 49(2), 251–273.
- Ramaswamy, R. and B. Rowthorn (1997). Deindustrialization: Causes and implications. IMF Working Papers 97/42, International Monetary Fund.
- Rodrik, D. (2013). Unconditional convergence in manufacturing. *The Quarterly Journal of Economics* 128(1), 165–204.
- Rodrik, D. (2016). Premature deindustrialization. *Journal of Economic Growth* 21(1), 1–33.
- Rostow, W. W. (1991). *The Stages of Economic Growth: A Non-Communist Manifesto* (3 ed.). Cambridge University Press.
- Szirmai, A. (2012). Industrialisation as an engine of growth in developing countries, 1950–2005. *Structural Change and Economic Dynamics* 23(4), 406–420.
- Szirmai, A. and N. Foster-McGregor (2017). Understanding the ability to sustain growth. *Gronigen Growth and Developing Centre GGDC working paper n. 173*.
- Tacchella, A., M. Cristelli, G. Caldarelli, A. Gabrielli, and L. Pietronero (2013). Economic complexity: conceptual grounding of a new metrics for global competitiveness. *Journal of Economic Dynamics and Control* 37(8), 1683–1691.
- Thurow, L. C. (1994). Microchips, not potato chips. *Foreign Affairs* 73(4), 189–192.
- UNIDO (2018). *INDSTAT 2, ISIC Revision 3*.
- Yu, X., G. Dosi, J. Lei, and A. Nuvolari (2015). Institutional change and productivity growth in China’s manufacturing: the microeconomics of knowledge accumulation and “creative restructuring”. *Industrial and Corporate Change* 24(3), 565–602.

Appendix

A Data coverage

Country	ISO3	Emp		Va		Country	ISO3	Emp		Va	
		%	first-last	%	first-last			%	first-last	%	first-last
Albania	ALB	6.33	1988-2014	7.94	1985-2014	Kyrgyzstan	KGZ	4.90	1990-2014	16.2	1997-2014
Algeria	DZA	7.14	1967-2014	8.75	1970-2014	Lao P.D.R.	LAO	1.36	1999-1999	9.90	1999-2014
Angola	AGO	1.14	1970-2014			Latvia	LVA	16.3	1990-2014	14.7	1993-2014
Argentina	ARG	11.9	1963-2002	20.7	1984-2002	Lebanon	LBN	7.54	1998-2007	8.24	1998-2014
Armenia	ARM	10.4	1990-2014	10.9	2004-2014	Lesotho	LSO	4.27	1982-2009	7.99	1980-1985
Aruba	ABW	5.09	1993-1999			Liberia	LBR	.258	1980-1986		
Australia	AUS	16.1	1963-2014	14.0	1970-2014	Libya	LBY			1.98	1970-1980
Austria	AUT	17.5	1963-2014	21.9	1970-2014	Lithuania	LTU	17.7	1992-2014	19.0	2000-2014
Azerbaijan	AZE	4.25	1990-2014	5.95	2001-2014	Luxembourg	LUX	14.1	1985-2014	11.2	1985-2014
Bahamas	BHS	2.31	1977-2014	2.72	1977-1998	Madagascar	MDG	3.26	1967-2006	6.18	1970-2006
Bahrain	BHR	12.7	1992-2014	14.2	2001-2014	Malawi	MWI	1.10	1964-2012	16.3	1970-2012
Bangladesh	BGD	2.59	1967-2011	14.0	1970-2011	Malaysia	MYS	11.7	1968-2014	23.1	1970-2014
Barbados	BRB	9.75	1970-1997	12.8	1970-1997	Maldives	MDV	1.32	2013-2014	2.25	2013-2014
Belarus	BLR	20.9	2005-2014	26.4	2005-2014	Malta	MLT	17.9	1963-2014	20.0	1970-2014
Belgium	BEL	20.1	1963-2014	21.3	1970-2014	Mauritius	MUS	20.6	1980-2014	20.2	1970-2014
Belize	BLZ	12.6	1991-1992	14.8	1989-1992	Mexico	MEX	4.36	1984-2014	18.8	1984-2014
Benin	BEN	.430	1980-1981	8.54	1974-1981	Mongolia	MNG	4.34	1990-2014	9.43	1990-2014
Bermuda	BMU	2.85	1990-2007	1.58	1996-2014	Montenegro	MNE	6.71	2010-2014		
Bolivia	BOL	1.51	1970-2012	14.4	1970-2014	Morocco	MAR	5.38	1976-2014	19.8	1976-2014
Bosnia Herzegov.	BIH	18.3	1990-2014	13.5	1990-2014	Mozambique	MOZ	1.50	1967-2000	29.6	1970-1973
Botswana	BWA	4.44	1981-2014	6.37	1981-2014	Myanmar	MMR	.170	1989-2013	19.0	1989-2013
Brazil	BRA	7.09	1963-2014	16.1	1990-2014	Namibia	NAM	4.97	1994-1994	12.5	1994-2014
Brunei Darussalam	BRN	2.78	2010-2010			Nepal	NPL	1.66	1986-2011	7.29	1986-2011
Bulgaria	BGR	21.5	1970-2014	14.9	1991-2014	Netherlands	NLD	13.5	1963-2014	16.5	1970-2014
Burkina Faso	BFA	.299	1974-1998	12.8	1974-1983	New Zealand	NZL	15.5	1963-2014	19.0	1970-2014
Burundi	BDI	.193	1980-2013	11.4	1971-2014	Nicaragua	NIC	4.43	1980-1985	19.0	1970-1985
Cambodia	KHM	2.56	1985-2000	10.6	1993-2000	Niger	NER	.044	1999-2014	6.38	1990-2014
Cameroon	CMR	1.21	1970-2008	13.9	1970-2002	Nigeria	NGA	.879	1963-1996	5.72	1970-1996
Canada	CAN	14.6	1963-2014	16.5	1970-2014	Norway	NOR	15.1	1963-2014	12.4	1970-2014
Central African Rep.	CAF	.419	1980-1993	17.7	1973-1993	Oman	OMN	4.41	1993-2014	8.07	1993-2014
Chile	CHL	7.39	1963-2014	16.6	1970-2014	Pakistan	PAK	2.26	1963-2006	10.9	1970-2006
China	CHN	7.51	1977-2014	31.1	1980-2014	Panama	PAN	4.56	1969-2014		
Hong Kong SAR	HKG	19.9	1973-2014	10.0	1973-2014	Papua New Guinea	PNG			3.23	1970-2001
Macao SAR	MAC	16.4	1991-2014	9.43	1978-2014	Paraguay	PRY	3.84	1970-2010	17.1	1970-2010
Taiwan Province	TWN	27.3	1973-2014			Peru	PER	6.20	1979-2014	16.6	1982-2014
Colombia	COL	4.18	1963-2014	16.3	1970-2014	Philippines	PHL	3.84	1963-2014	24.8	1970-2014
Congo	COG	1.11	1981-1988	10.1	1970-2009	Poland	POL	19.4	1970-2014	24.6	1970-2014
Cook Islands	COK			3.61	1998-2007	Portugal	PRT	16.2	1963-2014	17.1	1970-2014
Costa Rica	CRI	9.67	1963-2014	20.1	1970-2014	Qatar	QAT	8.10	1986-2014	9.69	1986-2014
Croatia	HRV	18.3	1990-2014	18.2	1990-2014	Rep. of Korea	KOR	12.7	1963-2014	26.6	1970-2014
Cuba	CUB			11.3	1976-1989	Rep. of Moldova	MDA	8.60	1990-2014	16.4	1993-2014
Cyprus	CYP	14.5	1963-2014	10.3	1970-2014	Romania	ROU	17.4	1963-2014		
Czechia	CZE	24.4	1990-2014	25.3	1995-2014	Russian Federation	RUS	15.5	1993-2014	17.7	1993-2014
Czechoslovakia	CSK			33.1	1970-1990	Rwanda	RWA	.626	1999-1999	7.96	1999-1999
Cote d'Ivoire	CIV	1.71	1966-1997	15.1	1970-1997	Saint Lucia	LCA			5.88	1991-1997
Denmark	DNK	15.9	1963-2014	16.4	1970-2014	Saudi Arabia	SAU	6.42	1976-2014	10.2	1976-2014
Dominican Rep.	DOM	10.2	1963-1997	32.5	1970-1984	Senegal	SEN	1.21	1974-2014	21.8	1974-2014
Ecuador	ECU	3.41	1963-2014	19.6	1970-2014	Singapore	SGP	18.3	1963-2014	23.5	1970-2014
Egypt	EGY	6.77	1964-2014	16.9	1970-2014	Slovakia	SVK	20.5	1991-2014	22.2	1993-2014
El Salvador	SLV	3.30	1975-1998	17.2	1970-1998	Slovenia	SVN	25.1	1990-2014	24.7	1990-2014
Eritrea	ERI			8.00	1992-2014	Somalia	SOM			5.22	1970-1986

Country	ISO3	Emp		Va		Country	ISO3	Emp		Va	
		%	first-last	%	first-last			%	first-last	%	first-last
Estonia	EST	19.7	1993-2014	17.3	1992-2014	South Africa	ZAF	12.4	1963-2014	19.3	1970-2014
Ethiopia	ETH	.434	1965-2014	5.79	1990-2014	Spain	ESP	13.9	1963-2014	19.2	1970-2014
Fiji	FJI	7.93	1980-2014	12.5	1970-2014	Sri Lanka	LKA	6.32	1966-2014	20.9	1970-2014
Finland	FIN	18.5	1963-2014	23.5	1970-2014	Sudan	SDN	1.48	1972-2001	8.87	1970-2001
France	FRA	18.2	1963-2014	16.8	1970-2014	Suriname	SUR	6.80	1974-2014	16.9	1987-2014
Gabon	GAB	4.00	1980-1995	6.19	1970-1995	Swaziland	SWZ	10.0	1980-2011	22.8	1970-2014
Gambia	GMB	1.04	1980-2004	7.72	1975-2004	Sweden	SWE	18.2	1963-2014	21.5	1970-2014
Georgia	GEO	4.33	1998-2014	12.7	2000-2014	Switzerland	CHE	17.7	1986-2014	19.8	1986-2014
Germany	DEU	18.6	1963-2014	22.5	1998-2014	Syrian Arab Rep.	SYR	7.20	1963-2010	7.96	1970-2005
Ghana	GHA	1.49	1963-2003	23.4	1970-2003	Tajikistan	TJK	5.19	1990-2014		
Greece	GRC	7.46	1963-2014	13.1	1970-2014	Thailand	THA	6.65	1968-2013	26.0	1970-2013
Guatemala	GTM	6.19	1968-2006	16.6	1971-1988	The Rep. of Macedonia	MKD	19.6	1990-2014	11.9	1990-2014
Haiti	HTI	1.05	1969-1997			Tonga	TON			9.51	1981-1981
Honduras	HND	5.48	1971-1995	17.3	1970-1996	Trinidad and Tobago	TTO	9.41	1966-2003	14.9	1970-2006
Hungary	HUN	21.6	1970-2014	23.1	1970-2014	Tunisia	TUN	8.67	1963-2014	16.0	1970-2014
Iceland	ISL	18.5	1967-2006	15.4	1970-2014	Turkey	TUR	7.01	1963-2014	22.7	1970-2014
India	IND	2.43	1963-2014	17.7	1970-2014	Uganda	UGA	.641	1963-2000	8.79	1971-2000
Indonesia	IDN	3.35	1970-2014	19.1	1970-2014	Ukraine	UKR	13.5	1992-2014	16.0	2001-2014
Iran	IRN	5.04	1963-2014	12.6	1970-2014	Utd. Arab Emirates	ARE	6.79	1970-2010	8.23	1977-2014
Iraq	IRQ	5.34	1970-2013	5.19	1970-2013	Utd. Kingdom	GBR	19.2	1963-2014	16.1	1970-2014
Ireland	IRL	15.3	1963-2014	21.7	1970-2014	Utd. Rep. of Tanzania	TZA	.889	1965-2014	9.13	1970-2014
Israel	ISR	15.5	1963-2014	18.7	1970-2014	Utd. States of America	USA	15.2	1963-2014	16.7	1970-2014
Italy	ITA	15.3	1967-2014	21.2	1970-2014	Uruguay	URY	10.9	1968-2014	18.9	1970-2014
Jamaica	JAM	6.46	1963-2006	14.3	1970-1996	Venezuela	VEN	7.60	1963-1998	24.7	1970-1998
Japan	JPN	15.9	1963-2014	25.4	1970-2014	Viet Nam	VNM	7.12	1998-2014	16.9	1998-2014
Jordan	JOR	8.94	1963-2014	17.8	1970-2014	Yemen	YEM	2.69	1989-2014	6.70	1998-2014
Kazakhstan	KAZ	5.81	1998-2014	11.6	2009-2014	Yugoslavia	YUG			32.3	1970-1989
Kenya	KEN	2.10	1963-2014	15.1	1970-2014	Zambia	ZMB	2.76	1963-1994	14.8	1970-2014
Kuwait	KWT	7.87	1970-2014	7.25	1970-2014	Zimbabwe	ZWE	5.17	1980-1996	11.7	2009-2014

Table A1: Data coverage for employment and value added. The table reports average total manufacturing shares and the first and last observations in the final dataset. Sources: UNIDO INDSTAT2 and Penn World Tables 9.1 for employment and MANVA dataset for value added. Note that sectoral coverage may be different.

B Revised Pavitt Taxonomy

Pavitt Class	Industry	Code
Supplier Dominated	Manufacture of food products and beverages	15
	Manufacture of tobacco products	16
	Manufacture of textiles	17
	Manufacture of wearing apparel; dressing and dyeing of fur	18
	Leather and leather products	19
	Manufacture of wood and of products of wood and cork	20
	Manufacture of fabricated metal product	28
	Manufacture of furniture; manufacturing n.e.c.	36
	Recycling	37
Scale Intensive	Manufacture of paper and paper products	21
	Publishing, printing and reproduction of recorded media	22
	Manufacture of coke, refined petroleum products and nuclear fuel	23
	Manufacture of rubber and plastics products	25
	Manufacture of other non-metallic mineral products	26
	Manufacture of basic metals	27
Specialized Suppliers	Manufacture of motor vehicles, trailers and semi-trailers	34
	Manufacture of machinery and equipment n.e.c.	29
	Manufacture of electrical machinery and apparatus n.e.c.	31
	Manufacture of other transport equipment	35
Science Based	Manufacture of chemicals and chemical products	24
	Manufacture of office, accounting and computing machinery	30
	Manufacture of radio, television and communication equipment and apparatus	32
	Manufacture of medical, precision and optical instruments, watches and clocks	33

Table A2. The Revised Pavitt taxonomy for manufacturing industries (Pavitt, 1984; Bogliacino and Pianta, 2010). Industry names and codes refer to ISIC Rev. 3.1.

C Robustness analysis: OECD Technological Classification

As a robustness check for the Pavitt taxonomy we employ a different technological classification which follows the Statistical Division of the OECD. The classification is based on the technological intensity of manufacturing industries proxied by industry level R&D intensity.¹¹ Our results in terms of variety of deindustrialization across classes and heterogeneous timing patterns in premature deindustrialization are robust even adopting this alternative classification.

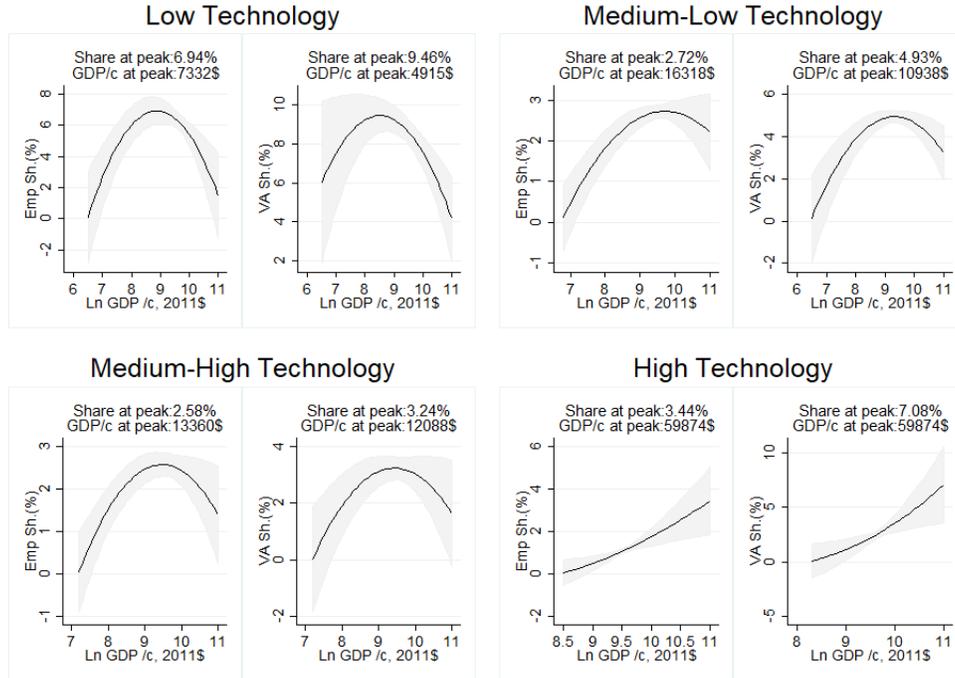


Fig. A1. Predicted employment and value added shares by OECD Technological Classes from regression in Eq. (1). Country fixed effect is averaged and population size is set to the sample average.

¹¹ The classification and some technical notes are available at: <https://www.oecd.org/sti/ind/48350231.pdf>

	Low		Medium-Low		Medium-High		High	
	<i>emp</i>	<i>va</i>	<i>emp</i>	<i>va</i>	<i>emp</i>	<i>va</i>	<i>emp</i>	<i>va</i>
Ln GDP /c	21.54*** (4.952)	14.53*** (4.980)	5.96*** (1.486)	11.29*** (2.198)	9.41*** (2.382)	12.23*** (3.488)	-3.11* (1.794)	-8.07 (4.915)
Ln GDP /c sqr.	-1.21*** (0.284)	-0.85*** (0.271)	-0.31*** (0.086)	-0.61*** (0.121)	-0.50*** (0.134)	-0.65*** (0.194)	0.23** (0.113)	0.55* (0.297)
Ln Pop.Sqr.	-6.00 (4.865)	0.49 (1.175)	0.69 (1.780)	0.90* (0.520)	2.11 (2.222)	0.89 (0.794)	2.40 (1.827)	2.01* (1.122)
Ln Pop.Sqr.	0.22 (0.156)	-0.06 (0.203)	0.01 (0.057)	0.08 (0.096)	-0.02 (0.069)	-0.04 (0.097)	-0.05 (0.052)	0.03 (0.132)
Observations	4,706	3,897	4,700	3,870	4,451	3,608	4,643	3,799
R-squared	0.299	0.344	0.232	0.145	0.353	0.297	0.250	0.239
Number of id	152	139	152	139	148	133	151	139
Decade FE	YES	YES	YES	YES	YES	YES	YES	YES
Country FE	YES	YES	YES	YES	YES	YES	YES	YES
Mean VA Sh.	5.28 %	7.74 %	2.23 %	4.14 %	1.93 %	2.49 %	.953 %	2.27 %
Slope at Min.	7.59***	4.40**	2.43***	4.10***	3.69***	4.54***	-.47	-1.51
Slope at Max.	-8.51***	-6.32***	-1.64***	-3.51***	-2.91***	-3.60***	2.52***	5.43**
L-M p-value	.000	.008	.008	.000	.001	.004	.176	.149

Table A3. Regression coefficients for employment and value added shares from Eq. (1). Robust standard errors in parentheses
*** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$.

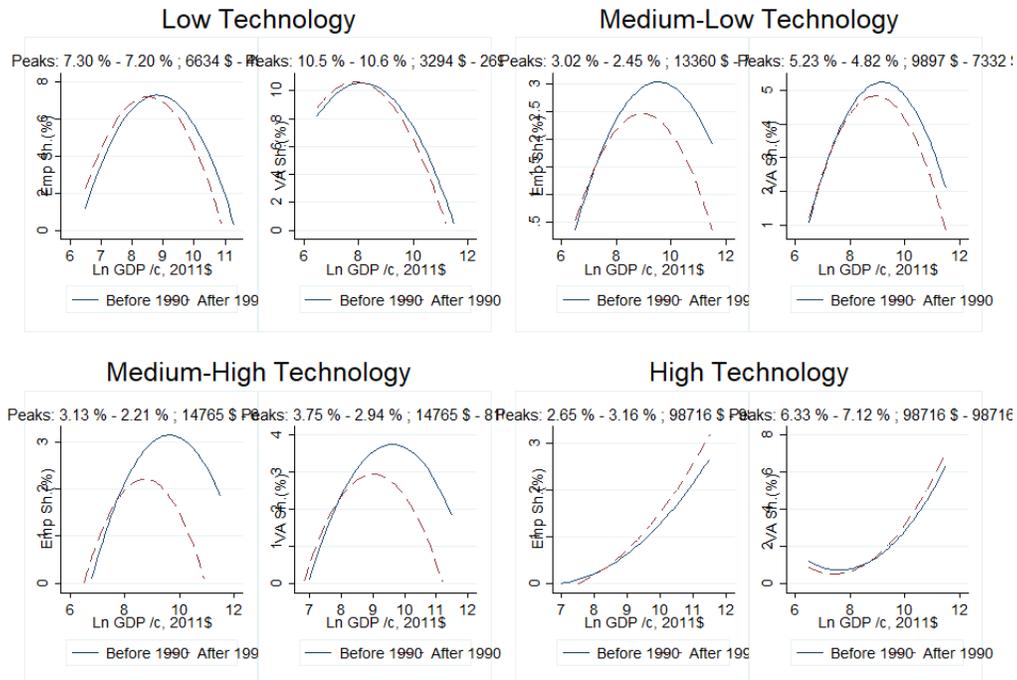


Fig. A2. Predicted employment and value added shares by OECD Technological Classes from regression in Eq. (2). Country fixed effect is averaged and population size is set to the sample average.

	Low		Medium-Low		Medium-High		High	
	<i>emp</i>	<i>va</i>	<i>emp</i>	<i>va</i>	<i>emp</i>	<i>va</i>	<i>emp</i>	<i>va</i>
Ln GDP /c	20.13*** (4.391)	14.46*** (4.631)	5.55*** (1.259)	10.63*** (2.138)	7.23*** (1.599)	10.28*** (2.525)	-1.44 (1.411)	-5.67 (4.075)
Ln GDP /c Sqr.	-1.14*** (0.246)	-0.89*** (0.246)	-0.29*** (0.071)	-0.58*** (0.115)	-0.38*** (0.085)	-0.53*** (0.131)	0.11 (0.086)	0.37 (0.237)
Ln Pop.	-1.66 (4.525)	-0.82 (1.140)	2.72 (1.797)	0.44 (0.460)	4.37** (2.072)	0.54 (0.523)	2.09 (1.762)	1.02 (0.772)
Ln Pop. Sqr.	0.04 (0.143)	-0.22 (0.201)	-0.07 (0.057)	0.02 (0.102)	-0.12* (0.067)	-0.18* (0.097)	-0.06 (0.053)	0.03 (0.125)
Ln GDP /c # Period	0.67*** (0.182)	0.41 (0.254)	0.23*** (0.066)	0.19 (0.136)	0.55*** (0.139)	0.52*** (0.137)	-0.11** (0.046)	-0.20 (0.145)
Ln GDP /c Sqr. # Period	-0.08*** (0.019)	-0.05** (0.025)	-0.03*** (0.007)	-0.03* (0.014)	-0.07*** (0.015)	-0.06*** (0.014)	0.01*** (0.005)	0.02 (0.015)
Obs.	4,706	3,897	4,700	3,870	4,451	3,608	4,643	3,799
R-squared	0.292	0.315	0.216	0.137	0.359	0.322	0.201	0.205
N. of Countries	152	139	152	139	148	133	151	139
Country FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Mean VA Sh.	5.28 %	7.74 %	2.23 %	4.14 %	1.93 %	2.49 %	.953 %	2.27 %

Table A4. Regression coefficients for employment and value added shares from regression in Eq. (2). Robust standard errors in parentheses *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$.

D Cluster Analysis

To set up a k-means cluster analysis we need to set the number of groups (k), the n clustering variables that define the n -dimensional space in which the analysis is carried out, and the distance function on which the similarity concept is based. The most commonly used distance function is the Euclidean distance. Given two n -dimensional points $p = (p_1, p_2, \dots, p_n)$ and $q = (q_1, q_2, \dots, q_n)$ the distance (d) is given by:

$$d(p, q) = \sqrt{\sum_{i=1}^n (p_i - q_i)^2} \quad (4)$$

The K-means is initialised through the extraction of k random observations, that are used as initial centroids to group the data. The iterative algorithm proceeds in three steps.

First, it computes the distance between each point and each current centroids (\bar{x}):

$$d(x, \bar{x}_k) \quad \forall x, k = 1, \dots, K \quad (5)$$

Second, each observation is assigned to a cluster (G_k), selecting the centroids to which the point is closest:

$$G(x) = \arg \text{Min}_{k=1..K} d(x, \bar{x}_k) \quad \forall x \quad (6)$$

Finally, the centroids of the new clusters are updated:

$$\bar{x}_k = \frac{1}{\#(x \in G_k)} \sum_{x \in G_k} x \quad \text{for } k = 1 \dots K \quad (7)$$

The procedure is iterated until clusters stabilise, and centroids stop changing. It is important to notice that the k-means algorithm might converge to a local optimum of the total within cluster sum of squares (WSS) and not necessarily to the global one. Therefore, in order to avoid to be trapped in a local optimum we perform 1000 cluster analysis, and we then choose the one with the smallest WSS.

To perform the k-means algorithm we have to choose ex-ante the number of clusters (k). Finding the right k is still an open choice. On the one hand increasing the number of clusters mechanically reduces the within group sum of squares. On the other

This Appendix section draws on Hastie, Tibshirani, and Friedman (2001) (chapter 14 section 3, p.467-72)

hand, adding too many clusters reduces the informative content of the cluster analysis. Ultimately, the validity of the cluster analysis should be confirmed by the scope of the description.

To detect the optimal number of groups (k) the most common method is to compare the within cluster sum of square (WSS) for alternative values of k .

$$WSS_k = \sum_{x \in G_k} d^2(x, \bar{x}_k) \quad \forall k \quad (8)$$

Higher values of k correspond to a more fine-grained representation of the observations. Indeed, the WSS is generally decreasing in k . A commonly used method to choose the best fitting k is to find where the $WSS(k+1)$ is only slightly lower than $WSS(k)$. That is, when the derivative of $WSS(k)$ w.r.t. k changes slope.

An alternative criterion proposed in the literature consists in evaluating the Proportional Reduction of Error (PRE) coefficient:

$$PRE_k = \frac{WSS_{k-1} - WSS_k}{WSS_{k-1}} \quad \forall k \geq 2 \quad (9)$$

the PRE_k coefficient quantifies how much the WSS changes adding one more cluster as a percentage of the WSS for $k-1$. PRE_k is bounded between 0 and 1. A PRE_k coefficient equal to zero means that there is no benefit in adding one more cluster, while a coefficient equal to one means that we have the perfect clustering (since WSS_k will be zero).

We will employ those two coefficients (i.e. WSS and PRE) to compare the clustering performance for different values of k . Using scree plots allows to detect the point in which the derivatives of the coefficients slow down. This is the so-called elbow method consisting in detecting the kink in the scree plots. Generally, it is difficult to find well-defined kinks.

Figure A3 shows the two alternatives scree plots from our cluster analysis. The scree plot relative to WSS shows a mild kink around 4. Similarly, we see that the proportional reduction of error improves of less than 5% moving from $k=4$ to $k=5$. Therefore, we opted for $k=4$.

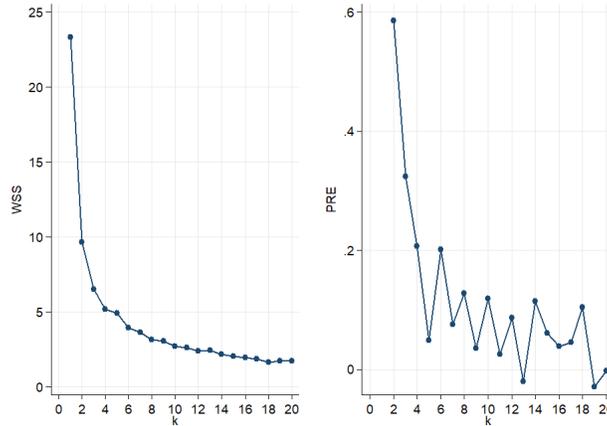


Fig. A3. Scree Plots