Centre for Economic Performance

Discussion Paper ISSN 2042-2695

No. 2063
December 2024

Industry concentration in Europe: Trends and methodological insights

Sara Calligaris
Miguel Chaves
Chiara Criscuolo
Josh De Lyon
Andrea Greppi
Oliviero Pallanch

LSE The London School of Economics and Political Science

UKRI Economic and Social Research Council

Abstract
Concentration – the share of an industry's output accounted for by its largest firms and a frequently used proxy of competition – has increased in European countries. This paper provides evidence about this development by introducing several methodological refinements in the cross-country measurement of concentration: it defines industries at a disaggregated level, mostly 3-digit; it takes into account the geographic level at which competition takes place - domestic, European or global; and it accounts for linkages between firms within the same domestic and multinational business group in the relevant geographic region of competition. It then applies these improvements to representative data for fifteen European countries, showing that average concentration increased by about 5 percentage points over the period 2000-2019, from 26% to more than 31%. Third, the paper investigates how each of the methodological improvements affects the levels and trends of concentration.

Key words: concentration, competition, market power
JEL: L11; L22; F14; F60; D22

This paper was produced as part of the Centre's Growth Programme. The Centre for Economic Performance is financed by the Economic and Social Research Council.

The authors would like to thank Andreas Bovin, Thomas Deisenhofer, Adriaan Dierx, Fabienne Ilzkovitz, and Vincent Verouden from DG Competition for their invaluable inputs, feedback, and support. The authors also thank Christian Abele, Mâtej Bajgar, Angela De Martiis, Isabelle Desnoyers-James, Hanna-Mari Kilpelainen, Guy Lalanne, Jens Lundsgaard, Jenniffer Solorzano Mosquera, and Costanza Tomaselli for their important contributions to this project, and Márcio Carvalho, Shai Somek, and Charles-Édouard Van De Put for technical assistance. The authors are grateful for comments from Delegates from the WPIA and CIIE for their valuable input and insightful comments, in particular from Andrea Linarello, Francesca Lotti, Phil Luck, Mika Maliranta, and Chiara Peroni who discussed early stages of this project in previous presentations. The authors thank Martin Borowiecki and Jan Stráský for their careful reading of the paper and insightful comments. The authors would like to thank Agnes Cimper, Peter Gal, Peter Horvat, Francesco Losma, Natia Mosiashvili, Gueram Sargsyan, Nori Yamano for sharing useful data and their expertise. The project is in part financed by the European Union under the Single Market Programme. The findings and results of this document have been used in "Protecting Competition in a Changing World", European Commission (2024).

Sara Calligaris, Andrea Greppi and Oliviero Pallanch all OECD. Miguel Chaves, NIPE, University of Minho. Chiara Criscuolo, OECD, International Finance Corporation and Centre for Economic Performance at London School of Economics. Josh De Lyon OECD Centre for Economic Performance at London School of Economics.

Published by
Centre for Economic Performance
London School of Economic and Political Science
Houghton Street, London WC2A 2AE

1. Introduction

Competition plays a pivotal role in fostering innovation, efficiency, and enhancing consumer welfare. Consequently, there is widespread recognition across academic and policy circles of the importance of promoting a dynamic competitive environment (Babina et al. (2023[2]); Aghion et al. (2005[3]); Buccirossi et al. (2013[4]); Backus (2020[5])).

While competition is conceptually straightforward and theoretically well-defined, there is no direct empirical method to measure it. Therefore, the economic literature has developed alternative proxies to assess the degree of competition. Market concentration, defined as the output share of a market accounted for by the largest firms, serves as a key proxy for competition (or lack thereof) and it is widely used by governments, competition authorities, and researchers.[1] The underlying idea is that a high level of concentration may reflect excessive barriers to entry or collusion of market leaders.

In recent years, a lively debate has emerged regarding whether market concentration has increased across OECD countries (Bajgar et al. (2023[6])) and to what extent this reflects a decrease in competition. While most studies document a rise in concentration (Bighelli et al. (2023[7]); Autor et al. (2020[8]); Barkai (2020[9]); Covarrubias et al. (2019[10]); Furman and Orszag (2015[11])), a few have found concentration to be flat or even decreasing (Amiti and Heise (2024[12]); Benkard et al. (2021[13]), Gutierrez and Phillippon (2023[14]); Kalemli-Özcan et al. (2023[15])). These differences come either from the type of data used, the country or countries covered, or the methodological decisions taken to measure concentration.

Although the concept of concentration is conceptually intuitive, several methodological decisions must be made to measure it (OECD, 2018[16]). One of the most critical decisions is defining the relevant market. The antitrust and competition communities identify the latter as the bundle of products that consumers perceive as substitutes, based on product characteristics

[1] The findings and results of this document have been used in European Commission (2024[1]).

and the geographic region from which customers can purchase.[2] This seemingly simple definition, however, often poses significant empirical challenges. There is no clear definition of which products should be included within a single market and, due to data limitations, most studies rely on industry-level data instead of product-level data. Moreover, while competition can occur at local, national, or international levels depending on the specific market characteristics, all existing studies typically consider a single level of aggregation to define the geographic boundaries of a market. Additionally, researchers and practitioners must decide which firm boundaries to consider – for example single firms or business groups.

As a result, the literature is characterised by differences between alternative approaches, leading to different concentration measures. These variations have sometimes resulted in opposing conclusions about the evolution of concentration over time. Understanding whether and how different assumptions shape the measurement of concentration is crucial for drawing accurate conclusions about the state of competition.

This paper measures concentration using the four-firm concentration ratio (CR4), which is the ratio between the total output of the four largest business groups and the total output in the market. The paper makes three important contributions to the existing literature. First and foremost, it introduces several methodological refinements in the measurement of concentration: it defines industries at a more disaggregated level; it calculates concentration at the geographic level at which competition takes place; and it accounts for linkages between firms within the same business group in the geographic region of competition. Second, it applies such improvements to representative data for 15 OECD countries and shows that average concentration increased by about 5 percentage points (henceforth, p.p.) over the period 2000-2019, from 26% to slightly more than 31%. Third, the paper conducts a thorough investigation to assess if and how each of the methodological improvements affects the levels and the trends of concentration. Overall, this study provides robust evidence on the evolution of concentration in OECD countries and informs future analyses of concentration on the importance of robust measurement decisions.

The first methodological innovation consists of defining industries, in a cross-country study, at a more granular level compared to much of the existing literature. The finer granularity

[2] The concept allows to identify the boundaries of the market in which companies compete and is used to assess market power for competition policy purposes. See European Commission (1997[53]).

of industry detail brings the concentration measure closer to the true definition of a market in the absence of data at the product market level. Unlike previous representative cross-country studies that measure concentration at the 2-digit level (for example, Bajgar et al. (2023[6]) and Kalemli-Özcan et al. (2023[15])), industries are defined in this study mostly at the 3-digit NACE Rev.2 level (with some higher aggregations due to data constraints). Additionally, supplementary analysis is conducted at the 4-digit level for selected manufacturing industries (where data quality allows). [3]

Second, for each industry, concentration is measured at the geographic level at which competition takes place. The geographical boundaries of each industry are determined by a taxonomy of geographic competition developed in Calligaris et al. (2024[17]). Industries may differ in their extent of international integration due to differences in their physical characteristics, technology, and policies. For example, industries such as "Manufacture of communication equipment" operate in global markets in which firms from all over the world can potentially reach any customers. In contrast, industries like "food services" are constrained to serving local customers only. Understanding the degree of internationalisation across industries allows to characterise the geographical boundaries of markets, which is crucial from a competition perspective. By comparing the value of European trade, extra-European trade, and the value of domestic sales in each industry, the taxonomy developed by Calligaris et al. (2024[17]) defines the geographical dimension at which each industry competes. Industries are classified as competing mainly at the domestic level (i.e., nationally), at the European level, or

[3] Note that the IMF (2019[32]) use data at the 4-digit industry level. However, they rely for both the numerator and the denominator of the concentration measure on the Orbis dataset. Orbis is not representative of the entire population of firms, has poor coverage for small firms, and heterogenous coverage among countries (see Bajgar et al. (2020[52]) and Bajgar et al. (2019[43]) for further details). Therefore, Orbis is not a suitable dataset to compute the denominator of the concentration measure especially when comparing countries over time. This paper overcomes this difficulty by obtaining production data at the 3-digit level that are representative of the firm population and comparable across countries and, as such, suitable to capture the overall size of the industry output (e.g., the denominator of the concentration measure). A detailed description on how the internationally comparable production data at such a level of granularity, can be found in Calligaris et al. (2024[17]).

global level.[4,5] The three geographical categories will be also referred to, throughout the paper, as geographical buckets.

Importantly, by aggregating firms' activities at the relevant geographic level determined by the taxonomy, the present concentration measure accounts for firms' sales to other countries within the same geographical market region. Put differently, the taxonomy indirectly accounts for international trade within the relevant geography, even in the absence of firm-level trade data, which are not available in a cross-country setting. Recent research by Amiti and Heise (2024[12]) shows that concentration in the United States has not increased when accounting for the growth of import penetration. The present analysis accounts for such trade within the relevant geographic level of competition by aggregating activities across countries within the relevant geographic region.[6] In addition, robustness checks are conducted considering the role of imports and exports from outside the geographic region and from countries not included in the sample.

In what follows, the term market will refer to the combination of the industry dimension – mainly defined at 3-digit level – and the geographical scope implied by the taxonomy developed by Calligaris et al. (2024[17]). The baseline concentration measure is computed at the corresponding level.

Third, relying on the procedure of Bajgar et al. (2023[6]), ownership linkages between firms are accounted for, and all the subsidiaries belonging to the same business group and active in the same market are treated as a unique entity. This paper makes a step further with respect to Bajgar et al. (2023[6]) by also relying on the taxonomy developed in Calligaris et al. (2024[17]) to define the geographical boundaries of markets. Bajgar et al. (2023[6]) considered all industries as competing domestically or, alternatively, at the European level, and summed

[4] In this cross-country setting, data on sub-national activities of firms are not available. Rossi-Hansberg et al. (2021[36]) show that concentration measured at the local level in the United States may have declined. However, Autor et al. (2023[39]) show that only local employment concentration fell, while local sales concentration increased.

[5] The taxonomy is developed from a European standpoint, drawing on data from 15 European countries. However, the distinction between whether an industry is tradeable or non-tradeable is likely to have broadly applicability, albeit with some caution, to non-European countries. See Calligaris et al. (2024[17]) for further details.

[6] As an example: industry "052" (Mining of lignite) is classified as domestic and, therefore, industry "052" in France and industry "052" in Germany represent two distinct markets. Industry "132" (Weaving of textiles) is instead classified as European and, therefore, industry "132" is considered as a unique market across European countries.

the activities of firms within business groups accordingly. The present paper, instead, accounts for business group linkages across countries included within the relevant geographic market of each industry. Therefore, concentration is computed at a different level of geographical aggregation depending on the industry considered and the activities of business groups are aggregated accordingly. This is a key contribution of the paper as it allows to account for the overall sales of multinational firms in any relevant market, and to consider the role of mergers and acquisitions in driving concentration trends. Neglecting these ownership linkages, as well as their heterogenous geographical extension across industries, may lead to underestimating concentration. The propensity to consolidate in groups can indeed be different across industries and could for example depend on the geographical level at which competition takes place.

These three important methodological improvements are then applied to analyse concentration trends in a cross-country setting. To do so, significant efforts have been made to construct the final dataset, which includes 127 industries, mostly at the 3-digit level, for fifteen European countries over the period 2000-2019, as detailed in Calligaris et al. (2024[17]).[7]

The final dataset has been obtained by merging detailed data collected from numerous sources. The total value of production for each industry-country-year is a key variable used to compute the denominator of the concentration measure. Production data is collected from National Accounts (NA, henceforth), the Structural Analysis (STAN) database, and Eurostat's Structural Business Statistics (SBS). The primary source for firms' production and ownership linkages information, needed to compute the numerator of the concentration ratio, comes from Moody's Orbis database, which is supplemented with data from the Orbis M&A database of Mergers and Acquisitions (M&As). The comparison and robustness exercises leverage on international trade flows data from OECD Inter-Country Input-Output (ICIO) tables, the "Base pour l'Analyse du Commerce International" (BACI) dataset, and the Trade in Services by Partner (TISP) data.[8]

The study finds that, on average and over the period considered, concentration increased by 5 p.p., from an initial level of 26% to 31%. Interestingly, the taxonomy of

[7] For industries competing at global level, three additional non-European countries are included in the analysis (Japan, Korea and USA).

[8] For further details on the data, and on the methodological steps required to prepare them, please refer to the companion paper by Calligaris et al. (2024[17]).

geographic competition also enables concentration trends to be examined separately for industries competing mainly at the domestic, European, and global levels. Industries competing at the domestic level have higher levels of concentration, on average 40% over the period considered, and have experienced the highest increase in concentration, of approximately 6 p.p. over the period analysed. Industries competing at the European and global levels exhibit lower levels of concentration, on average approximately 28%, and have experienced a slightly less pronounced increase in concentration of approximately 4 p.p. These findings, obtained in a large cross-countries sample and applying relevant methodological improvements, are in line with several analyses in the literature suggesting a decline of competition in the markets (Autor et al. (2020[8]); Barkai (2020[9]); Covarrubias et al. (2019[10]); Furman and Orszag (2015[11])).

As an additional key contribution, the paper provides – in a harmonised empirical framework – a careful examination of the impact of each of the methodological improvements discussed above on concentration trends. Even though there has been a lively debate in recent years about the evolution of concentration across many economies and despite the relevance of the topic from a policy perspective, no unanimous conclusion has been reached on whether concentration has been rising. This lack of consensus is primarily due to differences in methodologies and data.[9] Given the importance of promoting competition and understanding its evolution, it is important to assess the impact of these methodological choices, which this paper does in a harmonised cross-country setting.

Thus, the paper first assesses the choice of the level of industry aggregation.[10] Concentration is computed at different levels of industry aggregation (2-digit vs. 3-digit for most industries and, only for few selected industries, at 3-digit vs. 4-digit). Narrowly defined industries are more closely related to consumer product markets, thus providing a more accurate measure of concentration perceived by consumers. Implementing this step required

[9] For example, some papers argued that concentration has been flat in the United States when accounting for international trade (Amiti and Heise (2024[12])), when defining the market at a detailed product market level (Benkard et al. (2021[13])), or at local geographical level (Rossi-Hansberg et al. (2021[36])).

[10] It is beyond the scope of the present paper to provide a comparison between product market concentration and industry concentration. While this is conceptually important, it relies on access to product market data and is, thus, left for future work.

the development of a novel cross-country harmonised dataset.[11] The level of concentration is substantially higher when industries are more disaggregated, although trends remain similar. The finding that concentration is higher as markets are defined more narrowly is likely to be due to the following reason. At broad industry levels, such as the 2-digit level, even the largest business groups are unlikely to be active and sell in all the different 3-digit industries. In more narrowly defined, and hence specialised, industries, large players are more likely to participate in all the economic activities within the industry.

Second, concentration measures which consider the geographical level of competition for each industry, predicted by the taxonomy in Calligaris et al. (2024[17]), are compared to measures that do not account for the taxonomy. As explained above, when relying on the taxonomy, concentration is computed for each industry at a different geographical level: either domestic, European, or global.

Without the taxonomy, as standard in most of the literature, all industries are assumed to compete at the domestic (i.e., national) level. The exercise shows that neglecting the taxonomy and, therefore, the geographic boundaries of the market, leads to an overestimation of both the level and growth of concentration. When computing concentration using the taxonomy, the sales of the subsidiaries of a business group active in the same industry are summed over various countries.

A third exercise considers the role of international trade when measuring concentration, in the spirit of Amiti and Heise (2024[12]). Even though the taxonomy accounts for international trade within the region, it does not account for trade from countries outside the region.[12] Using industry-level trade data enables additional trade adjustments to be made. However, as international trade affects both the sales of the top four firms and the total market size (respectively, numerator and denominator of the concentration measure) while firm-level trade data are not available, strong assumptions must be made to implement trade adjustments to the numerator in this setting. Despite this, the results are informative, and suggest that adjusting for international trade dampens the rise in concentration, in line with Amiti and Heise

[11] In particular, it has required to develop a methodology (see Calligaris et al. (2024[17]) for the details) to apportion 2-digit level figures coming from National Accounts, which are representative of the population of firms and comparable across countries at the 3-digit.

[12] Due to data limitation, some European countries, such as Ireland, Luxembourg, and the Netherlands, are not included in the sample. Therefore, trade with these countries is not accounted for automatically by the taxonomy.

(2024[12]). In contrast to their paper, which finds that US concentration is flat after accounting for trade, average concentration still increases in the EU even after accounting for trade flows. However, the trade correction does dampen the increasing trend found for the baseline concentration measure, suggesting, in line with Amiti and Heise, an increase in import competition over the period considered.

The fourth exercise computes concentration trends neglecting the ownership linkages between firms. In the baseline concentration measure, the sales of all subsidiaries belonging to the same business group and active in the same market are summed up, and the sales of the four largest business groups (rather than firms) in market are summed up to get the numerator of the concentration measures. In the alternative measure, instead, each subsidiary is considered as a separate entity and, thus, the numerator is computed at the firm, rather than the business group, level. As expected, not considering the business group dimension leads to lower levels of concentration across all geographical buckets, with the difference being larger in the European market.

The results also show that ignoring the role of business groups and their dynamics would lead to underestimating the growth in concentration over the twenty years considered. Without accounting for business groups, concentration has been relatively flat in industries competing at the European and global level, whilst still rising in domestic industries. This suggests that the increase in concentration when accounting for business group linkages is at least partly driven by the expansion of business groups through their subsidiaries, or through acquiring new subsidiaries both within and across borders, rather than individual firms getting a larger share of market sales.

While the proposed refinements affect the average level of concentration and the magnitude of its increase over time, which explains the mixed findings in the literature, the direction of its evolution remain consistent, with concentration increasing across almost all specifications.

The structure of the paper is the following. While the remainder of this introduction reviews the relevant literature, Section 2 describes the various datasets used in the analysis. The methodology used to define concentration is outlined in Section 3. 4 presents trends of the baseline concentration measure, as well as several exercises showing how the different methodological decisions (geographic aggregation, industry aggregation, trade corrections, and

ownership structure of business groups) affect the aggregate concentration trends. Finally, Section 5 draws together all the analysis and discusses some implications.

1.1. Literature review

Concentration is considered a proxy for competition and has been used widely in academic and policy circles (OECD, 2021[18]).[13] For example, the European Commission considers a firm as dominant if its market share exceeds 40% of the relevant market. This notwithstanding, as it emerges from the industrial organisation literature, the theoretical relationship between competition and concentration is not univocal. In elementary models, it is possible to have competitive outcomes with high concentration (Bertrand) or with lower concentration (Cournot). The relationship becomes more nuanced when product differentiation is added to the models. The trade literature shows that an increase in competition from foreign firms due to a fall in trade costs can lead to a rise in the market share of the most productive domestic firms (Melitz, 2003[19]). For classic discussions on the topic, see Demsetz (1973[20]), who pointed out how high concentration can be the outcome of a competitive process, with firms gaining market shares outperforming rivals, and Schmalensee (1989[21]), who reviewed the early literature relating concentration with other proxies of competition. For more recent references see Berry et al. (2019[22]) and Syverson (2019[23]), who provide an overview of the advantages and drawbacks of using concentration to assess market power.

There is an extensive literature measuring trends in concentration. The works vary in the country covered, type of data used, and methodological decisions taken to define concentration, mostly concerning the boundaries of markets. Most studies are based on measures defined at the industry level, as industry-level data are more readily available. In contrast, few studies measure product market concentration, which is more directly in line with practices used by competition authorities in antitrust cases but often only cover a subset of the

[13] Note that this paper focuses only on one measure of concentration, CR4, neglecting alternative measures, such as the Hirschman-Herfindahl Index (HHI). Data limitations drive this choice. As explained later, the firm-level data used in this analysis are obtained from Orbis, which has good coverage of large firms (Bajgar et al., 2020[52]). HHI, instead, requires data on the population of firms. In addition, Orbis's sample of small and medium firms is not perfectly balanced over time, creating additional concerns about using HHI as a proxy for competition.

economy. Analyses also vary in the geographic dimension at which concentration is computed, varying from local (sub-national) to national and international levels. Some studies additionally account for international trade. This section provides an overview of the literature structured along the main contributions of this analysis discussed above.

First, the present paper contributes to the literature by reaching the 3-digit level of industry disaggregation in a harmonised cross-country setting under a host of methodological assumptions. Several analyses that measure concentration ratios at the national industry level, mainly focused on the United States, have documented an increase in national average concentration over the past few decades (Autor et al. (2020[8]); Barkai (2020[9]); Covarrubias et al. (2019[10]); Furman and Orszag (2015[11])). Grullon et al. (2019[24]) show that more than 75% of US industries have experienced an increase in industry concentration since the late 1990s. Looking over the past 100 years, Kwon et al. (2023[25]) documented a long-term rise in concentration. In addition to documenting rising concentration, Ganapati (2021[26]) showed that changes in industry concentration in the US are positively correlated with productivity and real output growth, uncorrelated with price changes and overall payroll, and negatively correlated with labour's share of revenue. Industry concentration also increased in Canada, with the largest rise in industries that were already more concentrated (Canada Competition Bureau, 2023[27]). Outside North America, studies have also found increasing industry concentration, albeit usually at a lower rate than in the United States. For example, Lashkari et al. (2019[28]) and De Ridder (2024[29]) both found rising industry concentration using administrative data for France, and De Loecker et al. (2022[30]) showed similar patterns for the United Kingdom.

Previous works performing cross-country analysis considered either industries at the 2-digit (or higher) level of aggregation (Bighelli et al. (2023[7]); Bajgar et al. (2023[6])), or reached a high level of disaggregation but only for a limited set of industries and countries (Koltay et al. (2023[31])). The IMF (2019[32]) documented rising average concentration at the 4-digit level across 27 economies, both advanced and developing. To reach such a disaggregated level of analysis, the concentration measure had to be adapted; it has been defined as the ratio between the sales of the top 4 and the sales of the top 20 firms in an industry. This measure is likely to overestimate concentration with respect to the more traditional concentration ratio (by underestimating the effective industry size). A complementary approach has been exploited in recent papers which focused on product market concentration. Affeldt et al. (2021[33]) constructed market shares starting from information available from EU Commission merger

cases. They found that concentration has increased over time, in line with Benkard et al. (2021[13]) who showed that product market concentration is higher than industry concentration. In Calligaris et al. (2024[34]), industry and product-level measures of concentration are shown to be correlated, suggesting that industry-level measures of concentration can be used as proxies for concentration in the narrow markets defined at the product-level.

Second, this paper contributes to the literature through the use of the taxonomy on the geographic level of competition and the implementation of trade corrections to capture the effective size of a market. Recent papers have sought to address issues surrounding the role of international trade on concentration measures. Amiti and Heise (2024[12]) highlighted that existing studies mostly measure industry concentration using only sales of firms located in the relevant country, rather than sales in the relevant market. That is, they do not account for international trade, and in particular for import competition – firms selling their product in foreign countries without establishing a local subsidiary. Their data allow them to cover the universe of US imports since 1992, construct the market shares of the foreign sellers in the United States, and correct for double counting of imports from US plants abroad.[14] Accounting for import competition, they showed that US industry concentration has been flat between 1992 and 2012 because foreign firms have increased their exports to the United States, even if their individual market shares tend to be small. Put differently, on average, foreign exporters increase the overall size of the market more than they impact the contribution of the top firms. Concentration mostly fell in industries with high initial import penetration, that also experienced the fastest growth in import competition. In the United Kingdom, the Competition and Markets Authority (2022[35]) (henceforth, CMA) showed that correcting concentration ratios for international trade causes a fall in the level of concentration, but there is still a slight increase over the period 1997-2018.[15]

While these studies have highlighted the importance of accounting for the international integration of markets, other studies have argued that for non-tradeable products the relevant geographic dimension is likely to be local (i.e., sub-national). Each of these studies examine

[14] Matched firm-level data of this nature are confidential and only available for one country at a time, not on a cross-country basis.

[15] Freund and Sidhu (2017[54]) find that an increase in the number of emerging market firms in an industry is associated with a decline in concentration, looking at both manufacturing and services industries.

local concentration within a single country; due to data constraints, there is no evidence on local concentration in a cross-country setting and such an analysis is also not possible in the present paper. The evidence on concentration trends at the local level is mixed. Rossi-Hansberg et al. (2021[36]) observed that the national trend of increasing industry concentration is not reflected in average local market concentration, which is declining in the United States. They explain the differing trends at the national and local levels by observing that large firms are expanding by opening establishments in new local markets. Relatedly, Hsieh and Rossi-Hansberg (2023[37]) documented how the "industrial revolution in services" – the increasing returns to fixed-cost-intensive technologies and changing management practices in services sectors – has led to the expansion into new markets and a reduction in local concentration. Rinz (2022[38]) found similarly decreasing trends in local industry concentration between 1976 and 2015. However, Autor et al. (2023[39]), still looking at the United States, found that only local employment concentration has decreased, while local sales concentration has increased. They explain the divergence in local and national employment concentration trends with the structural shift of the economy, with a reallocation of economic activity from relatively concentrated manufacturing sectors, where employment concentration is high, to relatively unconcentrated services sectors, where employment concentration is lower. Moreover, looking within industry-by-county cells, concentration has increased, even for employment. They suggested that the differences between their findings and those of Hsieh and Rossi-Hansberg (2023[37]) and Rinz (2022[38]) are explained by the use of alternative datasets. Finally, Benmelech et al. (2022[40]) and Smith and Ocampo (2022[41]) found increasing average local concentration, in line with Autor et al. (2023[39]), with the latter focusing on the retail sector.

Other studies have argued that the relevant market is supra-national. Bajgar et al. (2023[6]) considered the EU as a single market and computed concentration measures both at the national and the European levels. Accounting for the cross-country subsidiaries of business groups, they documented a slight increase in concentration across 12 European countries included in their sample over the period 2000-2014, both at the national and European levels. Lyons et al. (2001[42]) estimated a model of industry concentration which endogenously allows for markets to be defined at either the national or European level, showing that the four countries studied – France, Germany, Italy, and the United Kingdom - varied in their integration with Europe. Affeldt et al. (2021[33]) used market definitions from EU horizontal merger cases to define the geographic scope, showing that concentration increased most in

worldwide markets. Note, though, that their sample is an unbalanced panel, and the relevant geography is not fixed over time, so the results could be driven by changing sample composition if more concentrated sectors become global over time.

Third, this paper combines official data sources (National Accounts and Structural business database) to measure the evolution of industry output across countries over twenty years and relies on detailed firm-level information on ownership structure of multinational business groups to reliably identify ownership linkages and their dynamics over this period. Gutierrez and Phillippon (2023[14]) reported trends in concentration relying on concentration measures constructed using total market size (the of measure) based on the commercial database Orbis as the denominator, which has increasing coverage of small firms over time and can consequently lead to a spurious increase in industry output and thus to flat industry concentration trends. Kalemli-Özcan et al. (2023[15]) did not account for connections between firms within a business group, which, as discussed above, leads to underestimating increases in concentration (see also Bajgar et al. (2019[43]) for a detailed discussion about these two points).[16]

[16] They find increasing concentration trends only when restricting the sample to firms reporting consolidated accounts for the entire business group level. Note that, as explained in details in Section 3 and in Bajgar et al. (2019[43]), the methodology followed in this paper does not use consolidated accounts and relies on unconsolidated accounts, meaning that the precise activities of each firm is recorded in the specific industry and country in which it operates, whilst also accounting for connections between business groups.

2. Data

Measuring concentration at a granular level in a cross-country setting requires merging detailed data collected from numerous sources.

There are three types of variables used in this paper. The first two are measures of production and international trade flows, both at the country-industry-year level. Industries are generally defined at the 3-digit level of aggregation; however, in some cases, industries must be aggregated due to data constraints, as described in this section. The total value of production is a key variable used in the construction of the denominator of the concentration, which captures the total size of a market.[17] Trade flows data are used to carry out robustness exercises in the computation of the concentration measures. Using cross-country data at this level of granularity combined with international trade data is a key innovation of this paper. Finally, the numerator of the concentration measure is constructed using sales data at the firm level, as well as ownership data to reconstruct the structure of business groups.

The final sample covers 15 European countries (Belgium, Denmark, Finland, France, Germany, Greece, Hungary, Italy, Norway, Poland, Portugal, Slovenia, Spain, Sweden and the United Kingdom).[18] In terms of sectoral coverage, the analysis focuses on mining, manufacturing, non-financial market services (excluding real estate) and utilities sectors, following the NACE Rev. 2 classification at the 3-digit industry level.[19] The final sample

[17] Note that production data are also used in combination with international trade flows data to capture the geographical level at which competition takes place developed in the geographic taxonomy of industries available in the companion paper by Calligaris et al. (2024[17]) and used here to compute concentration.

[18] The analysis also includes data for three extra-European countries (Japan, Korea, and the United States) exclusively when looking at industries that compete at the global level, as defined by the taxonomy (Calligaris et al., 2024[17]). Since the analysis is conducted from a European perspective, these three countries are included only for industries belonging to the global bucket.

[19] NACE is the "statistical classification of economic activities in the European Community" and is the subject of legislation at the EU level, which imposes the use of the classification uniformly within all the Member States. The present NACE Rev. 2 is the new revised version of the NACE Rev. 1 and of its minor update NACE Rev. 1.1.

consists of 127 industries.[20] Out of these, 112 (88%) are at the 3-digit level, 10 (8%) at the 2-digit level, and 5 (4%) aggregating two or more 2-digit industries.[21] The period covered is 2000-2019.

2.1. Production data

To measure concentration at the most disaggregated industry level possible, particular attention has been given to measure a denominator that reflects the "true" value of industry output and represents, consistently across countries and over time, the population of firms in an industry. This requires data on the value of production, defined as gross output in millions of euros, at the 3-digit NACE Rev. 2 level which must be internationally comparable and, as such, cannot rely on aggregating commercial firm-level data that might not represent the entire population of firms and whose coverage might not only differ across countries but also within countries over time (see a detailed discussion in Bajgar et al. (2019[43])). Instead, data is collected from official sources, that rely on commonly agreed standards of collection designed to be representative of the business population and comparable across countries.

The main data source for gross output is the Eurostat National Accounts (NA), which is the primary dataset used by countries to measure Gross Domestic Product (GDP) and other key economic variables. NA data are representative of the whole population of firms and comparable across countries.[22] Eurostat publishes NA data, including value of production, at the A*64 level of the industry classification NACE Rev.2, comprising of aggregations of 2-digit level activities. These data are therefore supplemented with data from Eurostat's Structural Business Statistics (SBS, henceforth), which contain information on economic activity of all economic sectors excluding agriculture and personal services and provide data

[20] The data limitations and cleaning steps are fully outlined in 0. Among these steps, it has been decided to exclude industries for which there are less than 4 firms – which are required to properly compute the concentration measure. This is a conservative approach which may lead to concentration being underestimated if the observed small number of firms accurately reflects the market, rather than representing data limitations. For example, the tobacco industry is excluded due to limited data availability but is known to be a highly concentrated sector.

[21] For further details on the different levels of aggregation in the sample, refer to 0.

[22] These data are collected according to the System of National Accounts (SNA), the internationally agreed standard set of recommendations on how to compile measures of economic activity. The SNA describes a coherent, consistent, and integrated set of macroeconomic accounts in the context of a set of internationally agreed concepts, definitions, classifications and accounting rules.

on the value of production at the 3-digit level of aggregation.[23] SBS measure production at a more disaggregated level than NA, and the data are representative of the economy within countries but not across countries (due to different methodologies adopted by National Statistical Agencies to collect them).

Therefore, to get production data at the 3-digit level, the 3-digit SBS data are used to construct the share that each 3-digit industry represents within its own 2-digit industry. These shares are then used as weights to apportion each 2-digit production value from the NA to the corresponding 3-digit industry. The 3-digit production data obtained are therefore consistent with NA – and thus representative of the population of firms and comparable across countries – and at the same time available at the desired level of granularity.

More formally, $GO_{Sct}^{2d\,NA}$ is the value of gross output for the 2-digit industry S in country c and year t obtained from the NA data. Then, $GO_{sct}^{3d\,SBS}$ is the value of gross output for the 3-digit industry s in country c and year t obtained from SBS. The apportioned values of gross output at the 3-digit level $GO_{sct}^{3d\,NA}$ are calculated as:

$$GO_{sct}^{3d\,NA} = \frac{GO_{sct}^{3d\,SBS}}{GO_{Sct}^{2d\,SBS}} GO_{Sct}^{2d\,NA}, \qquad \textbf{Equation 1}$$

where the first term of the right-hand side are the weights calculated from SBS data, with $s \in S$ representing each 3-digit industry contained in the associated 2-digit industry S, such that $\sum_{s \in S} GO_{sct}^{3d\,SBS} = GO_{Sct}^{2d\,SBS}$.

A detailed explanation of the steps required to get the final production dataset is reported in Calligaris et al. (2024[17]).[24]

[23] Additionally, data for extra-EU countries included in the analysis (Japan, Korea, and the United States) are required for industries that compete at the global level, as defined by the taxonomy. Data for the extra-EU countries come from the OECD STAN database, which is also based on National Accounts, complemented with data from the Korean Statistical Information Service (Korea), the Ministry of Economy, Trade and Industry (Japan), and the Bureau of Economic Analysis (the United States) over the period 2000-2019. In most cases, these data are available at the desired level of disaggregation in the original data sources used and, as such, no additional cleaning or data preparation was needed.

[24] There are two additional obstacles to overcome to obtain consistent 3-digit level data on production over the long period considered. First, the SBS dataset contains missing values. Second, the classification of economic activity changed in 2008 (from NACE Rev. 1.1 to NACE Rev. 2). Therefore, an imputation procedure and a conversion from the old to the new NACE classification system are required to have a time series at industry NACE Rev. 2 level from 2000 to 2019.

2.2. Trade data

International trade data are collected from three main sources. The primary data source is the OECD ICIO tables, which provide import and export flows between 76 countries and 45 industries, over the period 1995 to 2020, and can therefore distinguish trade with European and non-European countries for each of the fifteen countries considered. However, ICIO data are only available at the 2-digit at its most disaggregated level. These data are therefore supplemented with highly disaggregated information on trade flows from the Centre d'études prospectives et d'informations internationals (CEPII) "Base pour l'Analyse du Commerce International" (BACI) database for physical goods, and the OECD "Trade in services by partner economy" (TISP) database for non-financial market services. These data are reported at the product level, so crosswalks have been used to convert the product-level data to disaggregated industry level. Following this matching exercise, it is possible to recover goods trade data for all manufacturing and mining industries at 3-digit level, while industries belonging to utilities and non-financial market services have, in some cases, been aggregated to a higher level.

The detailed trade data are used to construct weights at the country-partner-industry-year level, where the partner is usually aggregated to EU or non-EU countries, and an apportioning procedure in the same fashion of that used for the production data is carried out for the trade data. The weights are constructed using BACI and TISP data and are used to distribute the more aggregated trade values from ICIO to more detailed industries. This method maintains the desirable characteristics of the ICIO trade flows data while providing a more disaggregated industry classification.[25]

2.3. Firm-level financial data and ownership information

Sales data at the firm-level are required to compute the numerator of the concentration measure. The main firm-level dataset used is the 2021 vintage of the Moody's Orbis database. Numerous steps are undertaken to clean the data, closely following Gal (2013[44]) and Kalemli-Özcan et al. (2023[15]). Financial information within Orbis is available at the business group

[25] For further details about the international trade datasets, please see the companion paper (Calligaris et al., 2024[17]).

level (consolidated financial data aggregated across all subsidiaries of a firm) and at the individual firm level (unconsolidated information referring to individual firms' activities). In this paper, unconsolidated accounts have been used and can therefore be aggregated to the business group level according to the relevant market, as discussed in Section 3.[26]

Balance sheet information from Orbis is complemented with Worldscope, a cross-country firm-level commercial database of listed firms provided by Thomson Reuters. This dataset covers 95% of the global stock market capitalisation, substantially increasing the coverage of listed firms included in the sample.[27]

Following the methodology developed by Bajgar et al. (2023[6]), the sum of the gross output of all subsidiaries of a business group within a market (defined as the combination of an industry and a geographical bucket, i.e., domestic, European or global) is used as key variable of interest in the analysis on concentration . To do so, the worldwide structure of business groups needs to be reconstructed, and ownership linkages between headquarters and all their subsidiaries need to be identified.

Orbis data contain detailed ownership information, with each firm being linked to its global ultimate owner. The global ultimate owner is defined as the firm owning at least 50.01% of total shares of a subsidiary. However, Orbis ownership information only covers the period 2007-2020. To get this information for the years pre-2007 and to further complement and quality-check the existing information, Orbis is supplemented with data from the Orbis M&A database, also provided by Moody's. This database reports information on the M&A activities of firms around the world. It captures both domestic and cross-border M&As and covers deals

[26] As explained in detail in Section 3, the measures of concentration built in this paper use business group activities. This approach fundamentally relies on unconsolidated data of the individual subsidiaries within a business group, as the objective is to identify the precise industry and location of all the subsidiaries belonging to a group, and to correctly apportion the group sales to the markets in which the business group is active. In 0, additional details on the data cleaning and preparation are provided, especially on when and how consolidated accounts are used to supplement data on unconsolidated accounts.

[27] These additional data are especially important in providing relevant information for non-European countries, which are usually less represented in Orbis. As explained in Bajgar et al. (2019[43]), Orbis data are generally well suited to analysing industry concentration in Europe, since it has a good coverage of medium and large enterprises in these countries. In contrast, Orbis generally has poorer coverage of subsidiary-level information for non-European countries, especially for US firms. Note that this concern does not apply to business group-level information (consolidated accounts), since Orbis and WorldScope together cover close to the universe of listed firms.

involving target and acquiring firms across all industries, enabling changes in ownership to be tracked over the years and the identification of business group structures pre-2007.

A detailed discussion of the methodology is provided in 0, together with further details on the process of cleaning and harmonising multiple data sources. For a detailed explanation of the methodology used to build the business group structure, see Bajgar et al. (2023[6]).

3. Methodology

This section defines the baseline measure of concentration and describes four aspects of its computation: first industry aggregation; second, the role of markets' geographical boundaries and, relatedly, the impact of adjusting for international trade. Finally, it discusses the importance of considering business group activities rather than those of individual firms.

3.1. Concentration measure

The concentration ratio captures the share of gross output accounted for by the largest business groups in a market. Its measurement involves several crucial methodological decisions. Specifically, it requires: understanding what are the boundaries of the "relevant market"; what is meant with "largest business groups"; properly capture the boundaries of "business groups"; define how to measure output (for both business groups and industries).

Analytically, the baseline concentration ratio (CR4) is defined as follows:

$$CR^{4}_{s,g,t} = \frac{\sum_{f \in Top\ 4} S_{f,s,g,t}}{S_{s,g,t}}. \qquad \textbf{Equation 2}$$

The gross output of the four largest business groups (numerator) is defined as $\sum_{f \in Top\ 4} S_{f,s,g,t}$.[28,29] The overall size of a market (denominator), $S_{s,g,t}$, is defined as the total

[28] Note that throughout the paper "top 4 firms" refers to the 4 *business groups* (not firms) with the largest gross output in each market. The term "firm" has been preferred to "business group" for simplicity of explanation.

[29] The numerator is obtained by adding the sales (as a proxy of gross output) of the largest business groups in the relevant market, while the denominator is defined using the measured gross output of an industry. Using at the denominator the sum of sales of all the firms contained in the Orbis dataset would not provide an accurate representation of a market in this setting, since Orbis is not representative of all firms, and therefore does not capture the overall economic activity produced in a market (Bajgar et al., 2020[52]). This would lead to an underestimation of the denominator, especially in the initial years of the sample, which in turn would introduce a downward bias in concentration trends as the sampling of smaller firms increased over time (Bajgar et al., 2023[6]). Note that, for most industries, sales and gross output are very similar and, as such, the term "gross output" will also be used when discussing the numerator. However, in certain industries, such as Wholesale and Retail, there might be differences. These concerns are addressed in various robustness exercises (see subsection on robustness

gross output in an industry *s*, in its relevant geographic region of competition *g*, at time *t.* As already mentioned, industries *s* are mainly defined at the 3-digit level, reaching a higher level of disaggregation than previous representative cross-country analyses. Geographic regions of competition *g* are defined as domestic, European, global using the taxonomy of industry tradeability (Calligaris et al., 2024[17]). To define the contribution of the leading firms to the market, the analysis aggregates all activities of relevant business groups within a given industry-geographical region of competition ("market" in this setting). This means that concentration is calculated at different levels of geographic aggregation for different industries.

The remainder of this sub-section provides details on the methodology used to compute this concentration measure. In particular, it outlines the methodological choices related to the level of industry aggregation, the geographical scope of the market considered, and the boundaries of the firms considered.

3.2. Industry aggregation

When examining concentration, the "relevant" market must be defined. Competition authorities, in assessing cases, typically identify relatively narrow markets – based on product rather than industry classification systems. Empirical cross-country analyses are restricted by the availability of data, which are usually collected at the industry, not product, level. In addition, existing industry data used to construct concentration measures are normally available at relatively aggregate levels, which depart from the notion of relevant market.[30] As a result, existing cross-national studies rely on data at the 2-digit (or higher) industry level (for example, Bajgar et al. (2023[6]) and Kalemli-Özcan et al. (2023[15])). Studies within single countries may have information at a more detailed industry level but are often unable to account for the cross-border activities of firms and to be easily comparable with analyses in other countries. For these reasons, there have been concerns that measuring concentration at the industry level does

in Section 4). It is important to note that both the possible discrepancy between sales and gross output concepts and the difficulty of measuring industry output at the 3-digit level of aggregation might lead to cases in which measured concentration exceeds 1.

[30] For example, "manufacture of food products" is a 2-digit industry, but consumers are unlikely to consider meat products and oils and fats (both three-digit industries) as substitutable when making their purchasing decisions.

not accurately capture its actual level (Berry et al. (2019[22]); Benkard et al. (2021[13]), Shapiro (2018[45]), Werden and Froeb (2018[46])).

This paper tries to overcome some of these concerns by developing measures of concentration for more detailed industries – typically 3-digit level, whenever possible – while incorporating a cross-country dimension and allowing for international activities of firms (see below). In total, industry concentration can be calculated for 127 industries across mining, manufacturing, utilities and non-financial market services sectors. Manufacturing industries are almost all defined at the 3-digit level, while some services industries are slightly more aggregated to allow an accurate match with the trade data, necessary to apply the taxonomy on geographic scope, discussed next. For a complete list of the sample of industries used in this paper see Table A 1.

3.3. Geographic scope

The second dimension that defines the relevant market is its geographic scope. Technological, physical, and policy-related factors – which differ for each industry – determine the geographic boundaries of each market. For instance, consumers in markets in which it is not costly to supply products across borders are able to purchase products sourced outside their home country.

In this paper, the geographic level at which each industry competes is determined by the taxonomy of industries developed in (Calligaris et al., 2024[17]). By comparing the value of international trade with the value of production in each industry, the taxonomy distinguishes between industries that compete mainly at the domestic level (non-tradeable) and those that compete internationally (tradeable). When trade is high relative to the value of production, an industry is considered tradeable. For industries that are identified as tradeable, a further distinction is made between whether the industry competes at the European level or at a global level.[31] As a result, all the industries considered are classified as competing at the domestic, the European or the global level. Note that the taxonomy is fixed across time and countries.

[31] Among the 127 industries considered in this paper, the taxonomy allocates 27 of them in the domestic bucket, 80 in the European bucket, and 20 to the global bucket.

The taxonomy allows concentration to be computed at the appropriate level of geographic competition and, thus, advances on previous measures of industry concentration. The rationale is the following. For non-tradeable industries, competition takes place mainly domestically. Consequently, concentration measures should be computed within each country-industry. For industries that compete at the European level, the activities of business groups are aggregated across all European countries (noting that the activities of non-European subsidiaries are excluded, as discussed in detail in the next sub-section). For example, when defining concentration, the numerator includes the gross output of the biggest four firms across all European countries in each industry and the denominator aggregates industry level gross output across all European countries. Similarly, for industries defined as competing globally, both business groups and industry-level gross output are aggregated over all countries in the sample.[32]

Aggregating tradeable industries across countries offers a significant advantage by capturing the entirety of trade activities among countries within the geographic region, even in the absence of firm-level data on international trade. Given the importance of trade in shaping concentration trends (Amiti and Heise, 2024[12]), and the lack of detailed firm-level trade data across countries, this is a very important benefit. For example, in global markets, the full activities of firms in each industry are accounted for, regardless of where production and consumption take place (although note that with the data limitations, some countries remain excluded). By aggregating firm and industry-level gross output across countries within each geographical bucket, there is no need for intra- region trade adjustments. Therefore, the taxonomy provides a conceptually sound and empirically feasible solution to account for the globalised nature of highly tradeable industries when computing concentration.[33]

[32] Recall that, outside the EU, data are available for three additional countries: Korea, Japan, and USA.

[33] In principle, when looking for example at industries competing at domestic level, only the gross output sold domestically by the top firms should be considered in the numerator of concentration, subtracting therefore their exports. On the contrary, when looking at industries competing at the European level, all exports of top 4 firms to other European countries are correctly included in the numerator of the CR4.

3.4. Accounting for international trade

While the taxonomy indirectly accounts for international trade between countries within the geographic region in which the industry competes, it does not account for trade with countries outside that region (or those not included in the sample). Trade can affect both the numerator and the denominator of the concentration ratio. Imports increase the overall size of a market (the concentration ratio's denominator), and to the extent that foreign firms serving a market only through imports could be among the top firms, trade could also increase the numerator. The value of exports would decrease the overall size of the market i.e. the denominator of the concentration ratio and likely also its numerator, as exports are likely to be disproportionally represented by an industry's top firms. The overall effect of accounting for trade is therefore ambiguous as is the role of any adjustment on concentration trends. Amiti and Heise (2024[12]) use confidential firm-level micro data for the United States to show that industry concentration, which is increasing if the import correction is not made, becomes flat after accounting for the gross output of foreign exporters.

It is important to reiterate the interaction between the taxonomy and trade corrections. The taxonomy is defined using data on international trade. Industries with high values of trade will be classified either as competing at the European or the global level, and therefore only imports from outside the region are added in the denominator. For industries competing at the domestic level, imports comprise a smaller share of production. Hence, in the presence of the taxonomy, the import correction is less impactful.[34] Furthermore, while Amiti and Heise (2024[12]) have access to confidential firm-level micro data on sales of firms in the US, this is not available in a cross-country setting, yet the taxonomy can still indirectly account for much of international trade.

Despite this, as a robustness check, international trade from outside a geographic region g is incorporated into the concentration measures. The adjusted measure can be written as:

$$CR^4_{s,g,t} = \frac{\sum_{f \in Top\ 4} S_{f,s,g,t} - \alpha X_{s,g,t}}{S_{s,g,t} + M_{s,g,t} - X_{s,g,t}}. \qquad \textbf{Equation 3}$$

[34] An important caveat is that, due to data availability, there are only three additional countries (Japan, Korea, and USA) included in the global category. Therefore, the import correction always accounts for imports from other non-European countries, such as the People's Republic of China.

the denominator is adjusted to account for imports into the geographic region ($M_{s,g,t}$) and exports from it ($X_{s,g,t}$). Note that imports and exports between countries within a region, e.g., gross output of Italian firms in Spain for an industry defined to compete at the European level, do not need to be corrected as they are included in the production value of the exporting country (Italy, in this example).[35]

A further correction could be made in the numerator to account for the value of the top 4 firms' gross output that is exported to different markets. However, as firm-level data on trade are not available, an assumption must be made on the share of total exports that is accounted for by the top 4 firms, denoted by α. For example, α is assumed to be equal to the share of total production accounted for by the top 4 firms, this is likely to underestimate their share as larger firms are more likely to export (Bernard et al., 2012[47]). Additionally, it is assumed that firms that only sell in a market across borders– i.e., via exports – but are located abroad are not large enough to be in the top 4 (supported by the findings of Amiti and Heise (2024[12])). [36]

In this paper, the robustness checks performed to capture the impact of exposure to international competition on concentration are implemented in two ways. The first only makes a correction for imports in the denominator, while the second makes a correction for both imports and exports in the denominator and relies on the proportionality assumption between export and sales shares for top four firms to estimate the "export-adjusted" numerator.

Given the lack of the necessary firm-level data, these trade corrections are only used as robustness checks and not incorporated in the baseline, while in the baseline the taxonomy is relied on to account for the role of trade integration for concentration measures. This further highlights the key advantage of the taxonomy already mentioned: by aggregating sales across countries in tradeable industries, the inclusion of the taxonomy in computing concentration

[35] Note that not all European countries are included in the sample, so also their imports are incorporated in the adjusted denominator.

[36] Given that, unlike Amiti and Heise (2024[12]), firm-level international trade data are not available, if the foreign exporters rank amongst the top 4 firms, then the measure of concentration would be downward biased. Amiti and Heise (2024[12]) show that exporters into the US market tend to be smaller, which explains why accounting for increases in imports flattens the concentration trend.

measures allows to account for trade within the relevant geographic region at which each industry competes, reducing the importance to correct for trade flows.[37]

3.5. Identifying activities of business groups

Firms operating in the same market may be part of the same business group (Altomonte et al., 2021[48]). This is likely more relevant when concentration is measured at the global or European level, because business groups often serve each country through a different firm entity. Market power is likely to be determined at the business group level, because the activities of firms within the same group can be coordinated, and groups may acquire firms in the same industry to consolidate their market share. Therefore, it is crucial to account for the gross output of business groups when measuring concentration, and not to focus on individual firms or any other economic entity.

Building on Bajgar et al. (2023[6]) and Bajgar et al. (2019[43]) (who, in turn, build on Bloom et al. (2013[49])), this paper focuses on the activities of business groups. When measuring concentration, different subsidiaries belonging to the same business group and active in the same market are treated as a unique entity, as neglecting these ownership linkages would lead to an underestimation of concentration. Bajgar et al. (2023[6]) show that considering the business group dimension has a substantial impact on concentration measures. Therefore, accounting for linkages between firms in the same business group is important when measuring concentration, which the paper does taking also into account the level at which the specific industry competes. The use of the taxonomy developed in Calligaris et al. (2024[17]) complements such a methodology. The taxonomy allows to aggregate business groups' activities to the level at which the specific industry competes (rather than assuming that all industries competed at either the national or the European level, as in Bajgar et al. (2023[6]) and Bajgar et al. (2019[43]).

[37] For example, consider an industry that competes at the European level. For this industry, the relevant measure of concentration is based on the sales of the top 4 firms across European countries. Therefore, it is not necessary to subtract a firm's exports to other European countries to have the correct numerator, i.e., the total sales in European countries of each top 4 business group. Similarly, the denominator for an industry that competes at the European level is given by the sum of gross output of all European countries in that particular 3-digit industry. Since trade flows are symmetric, all imports/exports between European countries cancel out by construction, and as such the correction for trade flows within Europe is not needed.

To construct the relevant activities of a business group in each market, the unconsolidated gross output of each subsidiary is aggregated across all subsidiaries that operate in the relevant geographic region and industry.[38,39]

Figure 1 illustrates the methodology using as an example a fictitious business group composed of: i) the headquarter, and ii) five subsidiaries, operating in three different industries and five different countries. Assume that the market for industry 1 is global, for industry 2 is European, and for industry 3 is domestic. The business group is active in *four* different markets (each represented by a different colour in the figure): i) industry 1, where it competes at the global level, with total gross output equal to 80 (headquarter 70 + subsidiary A 10); ii) industry 2, where it competes at the European level, with total gross output equal to 55 (subsidiary C 25 + subsidiary E 30); iii) industry 3 in Spain, where it competes at domestic level, with total gross output equal to 50 (subsidiary D); and iv) industry 3 in France, where it competes at domestic level, with total gross output equal to 15 (subsidiary B). These are the different aggregation of the firm sales that would be compared with the total sales in each market. For further details, see 0 and Bajgar et al. (2023[6]), (2019[43]).[40]

[38] Note that in Orbis all firms (at the unconsolidated level) are registered as active only in their main industry of activity, although their activities could well span across different industries. Prioritising the use of unconsolidated rather over consolidated accounts partly reduces this issue because it uses the industry code of individual firms rather than of the headquarter. This notwithstanding, the consequence of registering all the activities in one single industry might result in an overestimation of the numerator of concentration, and potentially in concentration exceeding one and more so at more granular levels of industry aggregation.

[39] Consolidated accounts are used only to correct the unconsolidated information in cases where the total subsidiary sales exceed group sales (due to inter-company transactions) or where unconsolidated data are missing. See 0 for further details.

[40] Except for Bajgar et al. (2023[6]), previous studies have mainly followed two approaches to deal with cross-ownership linkages. A first approach is to neglect business groups and focus only on unconsolidated information of individual firms. This method underestimates concentration if multiple firms in the same market are part of the same group. A second approach is to neglect subsidiaries and focus only on the consolidated accounts of the headquarters. This method attributes the entire activity of the business group to the headquarters, overestimating (underestimating) the concentration in the headquarters' (subsidiaries') market. See Bajgar et al. (2019[43]) for additional explanations on these alternative approaches.

4. Trends

This section begins by presenting the evolution of the baseline concentration measure. Concentration has increased by 5 percentage points over the period, from 26% to 31%. It then compares the baseline to measures of concentration computed under alternative specifications. First, when industries are more disaggregated, the average level of concentration and its growth over time are higher. Second, accounting for the international dimension at which an industry may compete decreases the average concentration level and growth. Third, correcting for international trade leads to a lower growth of concentration. Finally, neglecting the ownership structure of business groups lowers the average level of concentration and its growth, especially in industries competing at the European and at the global level. Importantly, while the different assumptions affect both the average concentration level and its evolution over time, concentration remains increasing across almost all these different specifications.

4.1. Trends of the baseline concentration measure

The baseline concentration trends are first presented at the aggregate level and then separately for each geographical bucket. For both types of exercises, the evolution of the baseline concentration measure is presented by plotting both levels and the cumulative unweighted average change since the year 2000 (normalised to 0). In each year, the average yearly change is computed across all industries within a bucket for the European and the global buckets, and across all country-industries pairs within domestic industries. The cumulative change is computed by summing up the yearly changes starting from the base year 2000. Moreover, the percentage change of cumulative growth (where cumulative change is divided

by the initial level of average concentration) is also discussed.[41,42] In the baseline trends, the unweighted average is taken across industries for numerous reasons, which are discussed in detail later in the section, and comparisons to the unweighted trends are conducted and presented.

Figure 2 shows the average level (top panel) and cumulative changes (bottom panel) in concentration across all markets. To derive average concentration trends, i.e., averaging across domestic, EU, and global industries, it is important to note the following. Concentration for industries competing at the domestic level is computed for each country-industry-year. Concentration measures for industries competing at the European and global levels are computed at the industry-year level, as country-level data have been aggregated to the relevant level of geographic competition. Therefore, appropriate weighting is adopted to ensure that the trends do not over-represent domestic industries.[43]

Average concentration rose, from around 26% in 2000 to above 31% in 2019 (top panel). In line with this, the bottom panel of Figure 2 shows that concentration increased by 5 percentage points (p.p., henceforth) between 2000 and 2019. This is equivalent to an increase in average concentration of about 21% in percentage terms. This result aligns well with other findings in the literature. Bajgar et al. (2023[6]) find that concentration over 2000-2014

[41] This is because average concentration can start from different initial levels in the various exercises of this section and comparing only cumulative changes might be misleading. There may be minor differences between the trends in the cumulative changes and the average level of concentration, as the former holds the sample constant between periods while the latter may be affected by a small number of cases whereby industries enter or exit the sample.

[42] Concentration trends are presented both in levels and as cumulative growth. Note that, given that the sample used in the analysis is not fully balanced, the cumulative average changes control for any changes in sample composition over time, while trends in levels can potentially exhibit jumps caused by industries with substantially different levels of concentration compared to the average level entering and exiting. Therefore, trends of cumulative changes are more methodologically stable and, as such, are the preferred option to show the evolution of concentration over time. More generally, concentration levels should always be considered with caution due to data limitations. Levels of concentration might be affected by specific data issues for several reasons, including: different data sources and definitions for the main variables for the numerator and the denominator; output measures being potentially volatile and difficult to measure at the 3-digit level in certain industries; attribution of all firms' sales to a unique industry, which may lead to an overestimation of the numerator, as discussed in Footnote 38. Each of these caveats might lead to measured concentration in a specific market being higher than 1 (see also footnotes 38 and 40 for further details).

[43] For industries competing at the European and at the global levels, each industry-year observation is counted 15 times (the number of countries in the data) to ensure they are weighted equally with industries competing at the domestic level (otherwise, industries competing at the domestic level would disproportionally influence the statistics). This approach effectively assigns each country the same value of concentration within an industry-year for industries competing at the European and global levels.

increased by about 2 p.p. in Europe and by 7 p.p. in North America. Autor et al. (2020[8]), focusing on US data, find that concentration increased across all broadly-defined sectors of the economy within a range of 5 to 15 p.p. between 1980 and 2012. Furman and Orszag (2015[11]) found that concentration increase, over 1997-2007 in three-fourths of the Census Bureau defined broad sectors.

Figure 3 reports concentration levels (top panel) and cumulative unweighted average changes (bottom panel) for each of the three geographical buckets, disentangling the information contained in the aggregate graphs of Figure 2 (a discussion on the choice of the weighting scheme adopted is provided in the next sub-section). The average level of concentration is higher in industries competing at the domestic level than in those competing at the European and global levels, with the latter two having similar levels of concentration. Specifically, the top 4 firms comprise around 43% of the total gross output of the industry in industries competing at the domestic level, and around 26% in the other geographical regions (on average, over the period considered).

The bottom panel of Figure 3 shows that industry concentration has increased across all geographic buckets. Industries that compete at the domestic level had the greatest increase in average concentration, by around 6 p.p. between 2000 and 2019. Industries that compete internationally – either at the European or global level – experienced an increase in concentration of approximately 4 p.p. However, while industries competing at the domestic and European levels see a smooth increase over the period considered, in industries competing at the global level concentration is relatively stable up to 2012, and then starts to increase from 2012 onwards.[44]

4.1.1. Robustness checks of the baseline concentration measure

This sub-section first discusses different robustness checks performed on the baseline measure. First, the choice of the weighting scheme adopted when computing average concentration. Second, the robustness of the results across subsamples of countries and

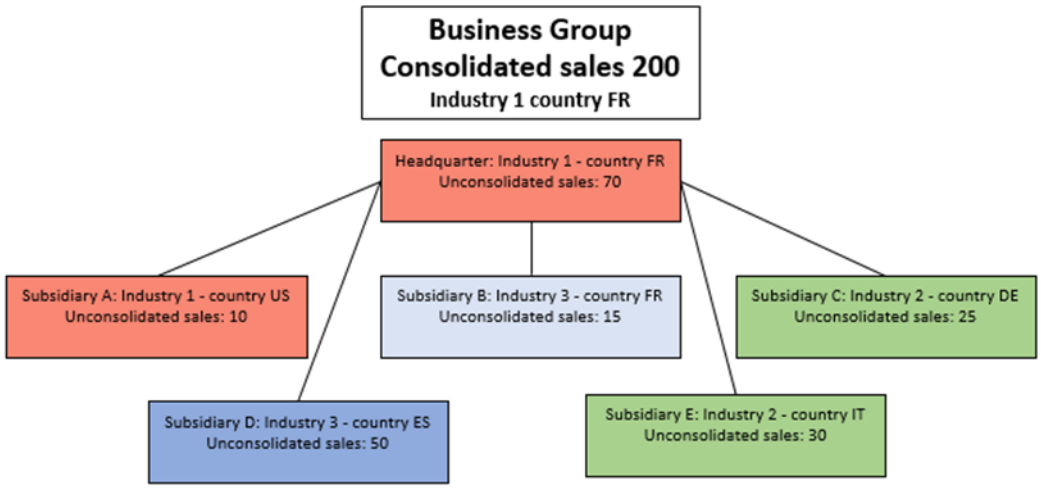
[44] Note that, in each year, the global bucket includes 20 observations (industries), the European one 80 observations (industries), while the domestic buckets includes 405 observations (27 industries for 15 countries). The trend for the global bucket is therefore slightly more volatile than the trends for other buckets: since it covers fewer observations, it is more sensitive to single industries' changes in concentration.

industries. Third, it reports an alternative concentration measure based on the top 8 business groups rather than the top 4 (CR8 instead of CR4).

The baseline analysis does not weight by the size of the market; instead, it takes the unweighted averages across geographies and industries. The unweighted average is chosen as the baseline for numerous reasons. First, when examining overall concentration trends in each bucket, weighting causes the trends to be strongly determined by just a few big countries and industries (or geography-industry pairs), whereas it is preferable in this context to provide evidence that is more equally representative across countries and industries. Second, due to data limitations, there is variation in the level of aggregation of industry classifications (mostly 3 digit, some 2 digit, and some aggregations of 2 digits). More aggregated industries would mechanically get more weight. Third, once each industry is assigned to a geographic bucket, its weight within the bucket does not necessarily correspond to its share of the domestic economy and, thus, to its relative importance within the country, making the interpretation of the weighted trends challenging. Fourth, the variable available in the data that could be used for weighting is gross output, not value added. Gross output includes the value of intermediates, so weighting by gross output would give more weight to downstream industries relative to upstream industries.[45] Despite these drawbacks, output weighted trends for all three buckets are presented as a robustness checks in the Figures and Tables

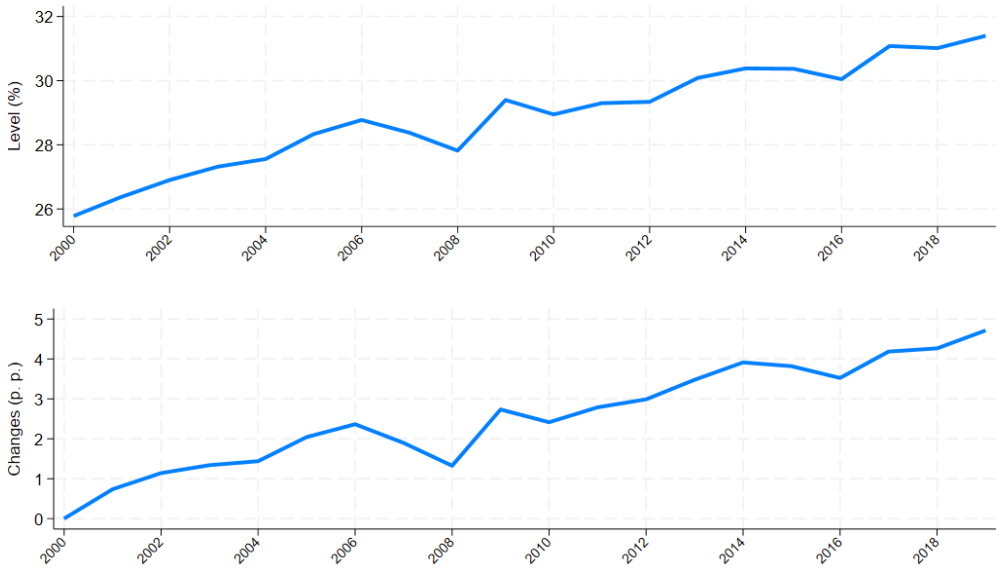
[45] Take the example of two industries with the same value added, but different value of intermediates. At the same value added, the industry with higher value of intermediates will have higher gross output, and hence a larger weight.

Figure 1. Example of the apportioning technique of business group activities



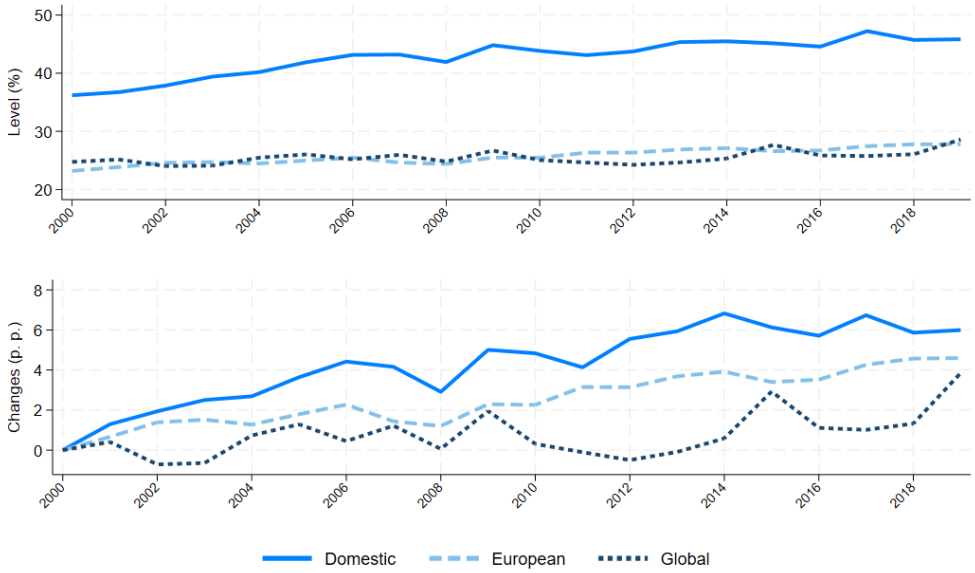
Note: The figure depicts an example of a hypothetical group consisting of a parent company from France and operating: i) in industry 1 (global), with a US subsidiary in the same industry; ii) in industry 2 (European) with two subsidiaries, one from Germany and one from Italy; and finally, iii) in industry 3 (domestic) with a subsidiary in France and one in Spain. The different colours identify the four different markets in which the group is active: the global market in industry 1 (with total gross output of 80), the European market in industry 2 (with total gross output of 55), the French market in industry 3 (with total gross output of 15), and the Spanish market in industry 3 (with total gross output of 50).

Figure 2. Concentration aggregating across geographical buckets



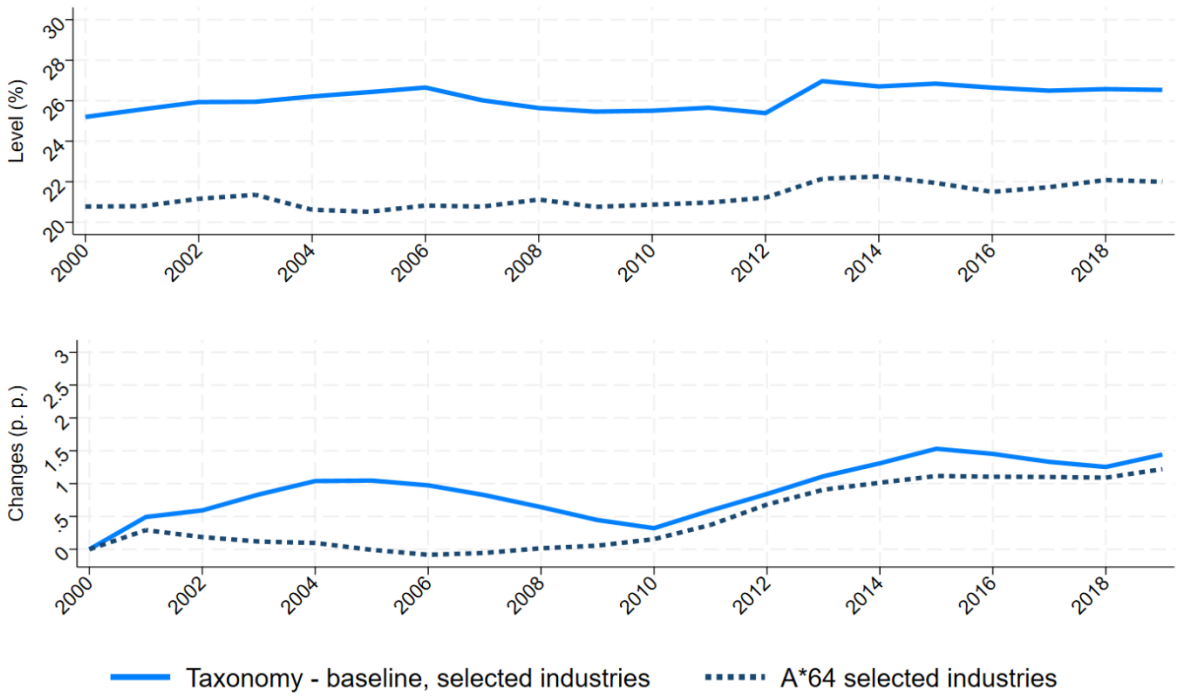
Note: The chart shows the weighted average across industry-geography combinations of CR4 levels (top panel) and cumulative growth (bottom panel). A weight of 1 is attributed to country-industries cells for domestic industries, while a weight of 15 is attributed to industries belonging to the European and the global buckets. Industries included in the analysis are a mix of 2 and 3-digit industries belonging to mining, manufacturing, utilities, and non-financial market services and are classified as either domestic, European, or global, depending on the taxonomy developed by Calligaris et al. . The countries included in the sample are BEL, DEU, DNK, ESP, FIN, FRA, GBR, GRC, HUN, ITA, NOR, POL, PRT, SVN, and SWE for the domestic and European buckets, while in the global one also JPN, KOR, and USA are included.
Source: Authors' calculations.

Figure 3. Concentration across geographical buckets



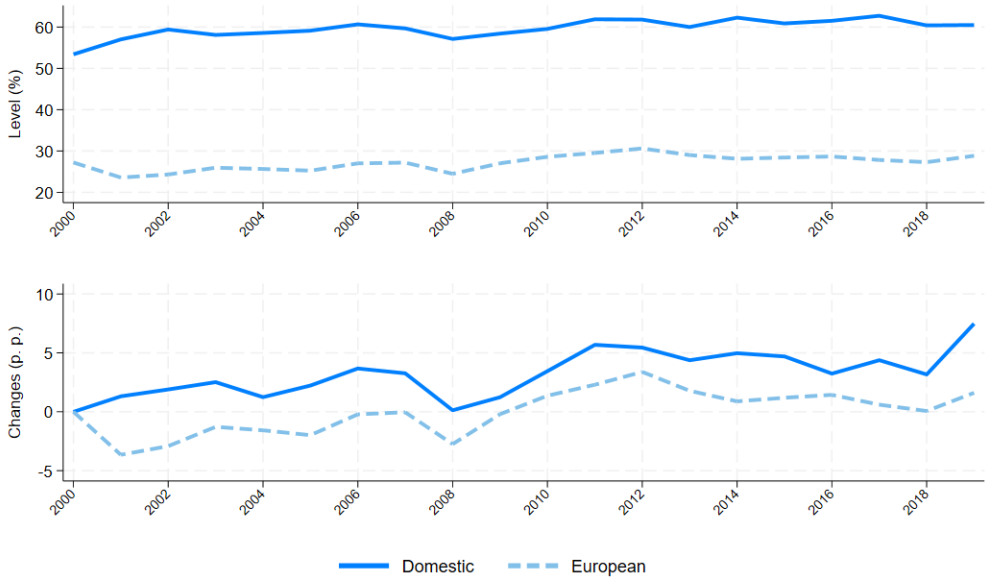
Note: The chart shows the unweighted average across industries (and countries, for the domestic bucket) of CR4 levels (top panel) and cumulative growth (bottom panel). Industries included in the analysis are mix of 2 and 3-digit industries belonging to mining, manufacturing, utilities, and non-financial market services and are classified as either domestic, European, or global, depending on the taxonomy developed by Calligaris et al. . The countries included in the sample are BEL, DEU, DNK, ESP, FIN, FRA, GBR, GRC, HUN, ITA, NOR, POL, PRT, SVN, and SWE for the domestic and European bucket, while in the global one also JPN, KOR, and USA are included.
Source: Authors' calculations.

Figure 4. Concentration with different industry aggregation levels (selected industries)



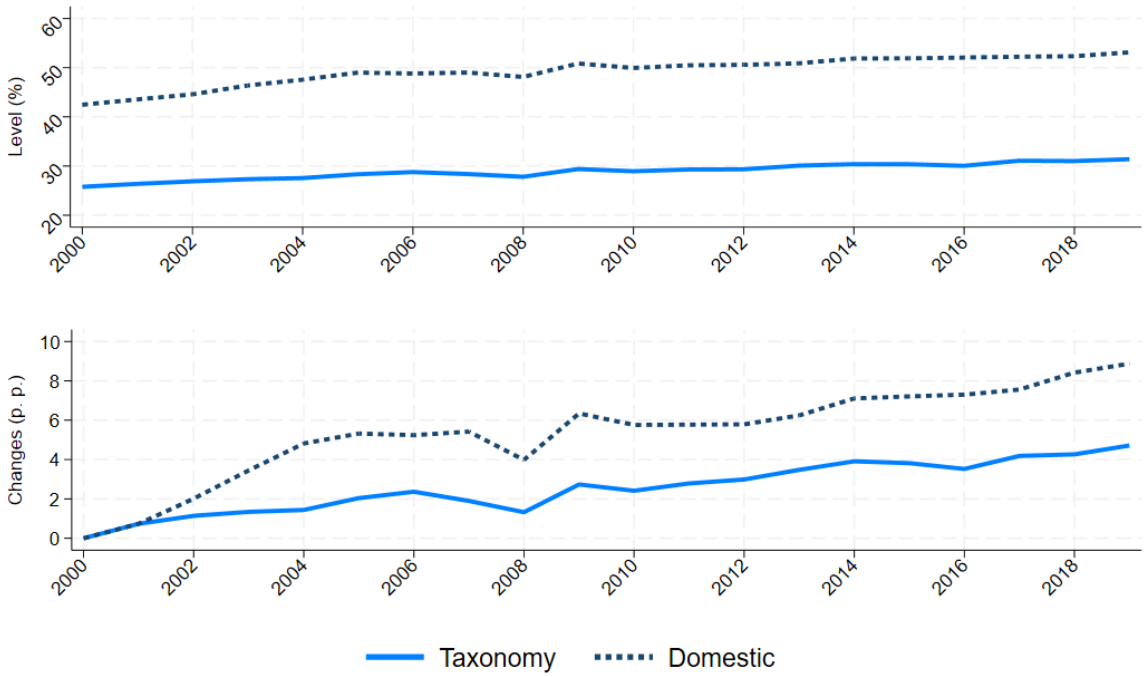
Note: The chart shows the average across industry-geography combinations of CR4 levels (top panel) and cumulative growth (bottom panel) in a selection of industries when using the industry aggregation of the baseline measure (continuous blue line, mostly 3-digit) and when using a higher-level industry aggregation (dotted dark line, mostly 2-digit). A weight of 1 is attributed to country-industries cells for domestic industries, while a weight of 15 is attributed to industries belonging to the European and the global buckets. Industries included in the analysis include all the 2-digit industries (following the STAN A*64 classification) whose 3-digit sub-industries belong to a unique geographical bucket. The 2-digit industries are, for the domestic bucket: C33, D35, D36, D37, H55, J59, and J61; for the European bucket: C16, C17, C19, C20, C22, C24, C27, C28, C29, G45, G46, G47H25, H53, J62, M71, M72, M73, N77, N78, and N79. The countries included in the sample are BEL, DEU, DNK, ESP, FIN, FRA, GBR, GRC, HUN, ITA, NOR, POL, PRT, SVN, and SWE.
Source: Authors' calculations.

Figure 5. Concentration for selected 4-digit industries



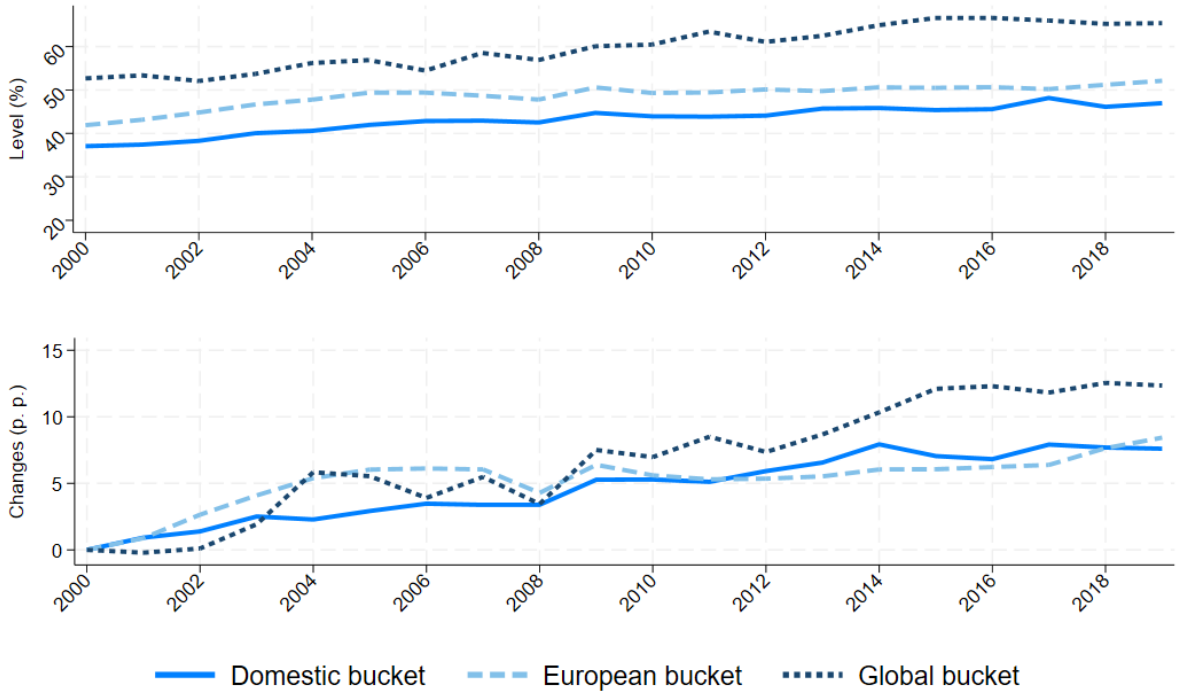
Note: The chart shows the unweighted average across industries (and countries, for the domestic bucket) of CR4 levels (top panel) and cumulative growth (bottom panel). Industries included in the analysis are all the 4-digit industries (except for 1042 and 1089 due to data limitations) belonging to the 2-digit industries from C10 (manufacture of food products) and C11 (manufacture of beverages). The countries included in the sample are BEL, DEU, DNK, ESP, FIN, FRA, GBR, GRC, HUN, ITA, NOR, POL, PRT, SVN, and SWE.
Source: Authors' calculations.

Figure 6. National concentration aggregating over different geographical regions



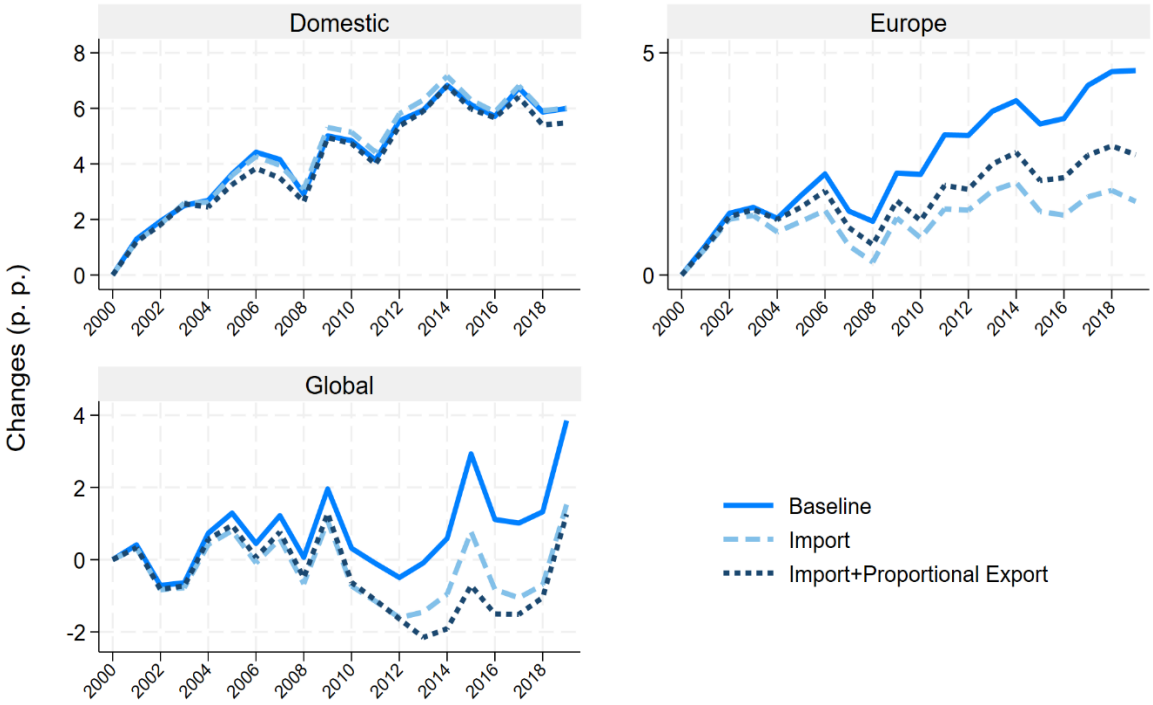
Note: The chart shows the average across industry-geography combinations of CR4 levels (top panel) and cumulative growth (bottom panel) when the taxonomy is adopted (continuous blue line) and with all industries considered as domestic ones (dotted dark line). When the taxonomy is used, a weight of 1 is attributed to country-industries cells for domestic industries, while a weight of 15 is attributed to industries belonging to the European and the global buckets. Industries included in the analysis are mix of 2 and 3-digit industries belonging to mining, manufacturing, utilities, and non-financial market services. The countries included in the sample are BEL, DEU, DNK, ESP, FIN, FRA, GBR, GRC, HUN, ITA, NOR, POL, PRT, SVN, and SWE for the domestic and European bucket, while in the global one also JPN, KOR, and USA are included.
Source: Authors' calculations.

Figure 7. National concentration across geographical buckets (all industries as domestic)



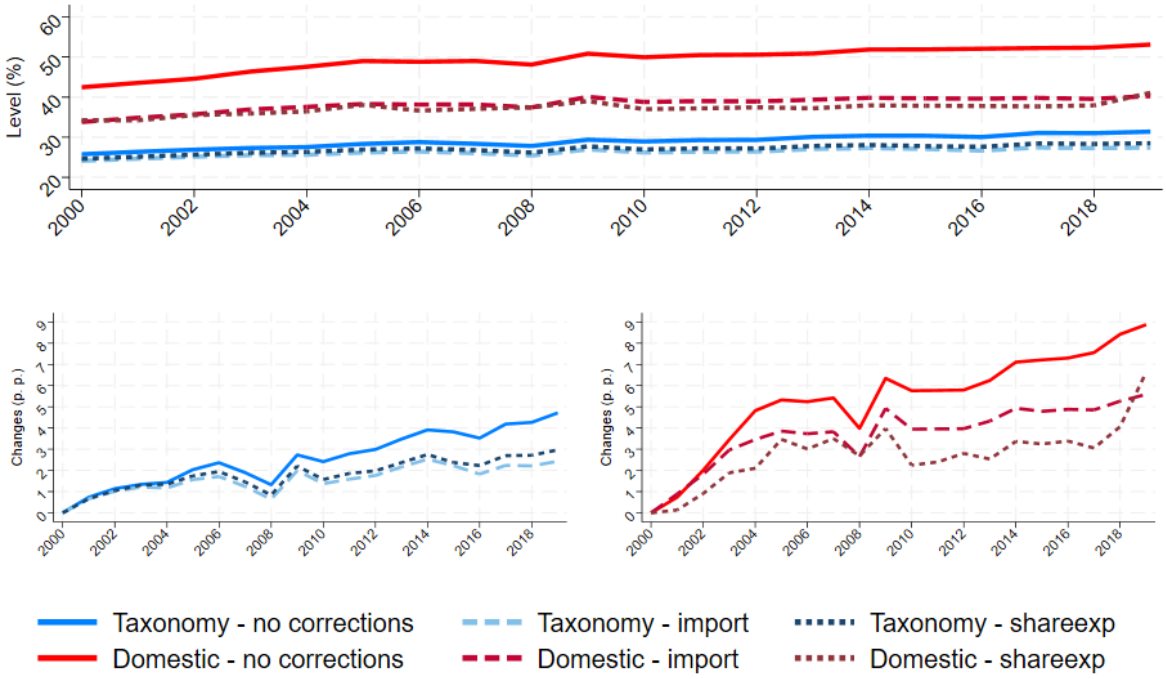
Note: The chart shows the unweighted average across industries and countries and within geographical buckets of CR4 levels (top panel) and cumulative growth (bottom panel), computed treating all industries as if they were competing at the domestic level. Industries included in the analysis are mix of 2 and 3-digit industries belonging to mining, manufacturing, utilities, and non-financial market services. The countries included in the sample are BEL, DEU, DNK, ESP, FIN, FRA, GBR, GRC, HUN, ITA, NOR, POL, PRT, SVN, and SWE for the domestic and European bucket, while in the global one also JPN, KOR, and USA are included.
Source: Authors' calculations.

Figure 8. Different trade adjustments



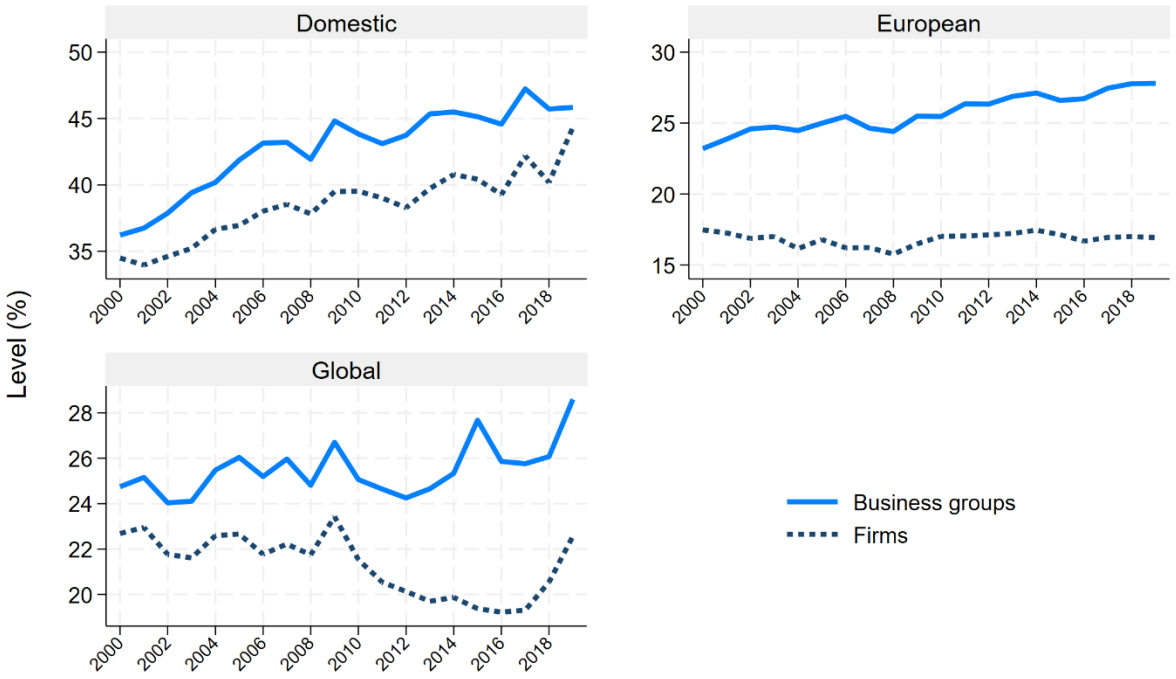
Note: The chart shows the unweighted average across industries and countries of cumulative change in CR4 across geographical buckets for different types of trade adjustments. The solid blue lines refer to the baseline average cumulative change (no corrections). The light blue dashed ones to the correction obtained by adding import at the denominator. The dark blue dotted ones to the correction obtained by adding import and subtracting total exports at the denominator, and by subtracting at the numerator a fraction of export given by the share of gross output accounted for by the four largest firms in the market. Industries are a mix of 2 and 3-digit industries belonging to mining, manufacturing, utilities, and non-financial market. The countries included in the sample are BEL, DEU, DNK, ESP, FIN, FRA, GBR, GRC, HUN, ITA, NOR, POL, PRT, SVN, and SWE. Industries competing at the global level also include JPN, KOR, and USA.
Source: Authors' calculations.

Figure 9. Concentration level different trade corrections (taxonomy vs. all as domestic)



Note: the chart shows the unweighted average across industries and countries of CR4 levels for different types of trade adjustments. Blue lines refer to concentration computed using the taxonomy developed by Calligaris et al., red lines to concentration computed with all industries considered as domestic. Solid lines refer to the baseline average level (no corrections). Dashed ones to the correction obtained by adding import at the denominator. Dotted ones to the correction obtained by adding import and subtracting total exports at the denominator, and by subtracting at the numerator a fraction of export given by the share of gross output accounted for by the four largest firms in the market. Industries are a mix of 2 and 3-digit industries belonging to mining, manufacturing, utilities, and non-financial market. The countries included in the sample are BEL, DEU, DNK, ESP, FIN, FRA, GBR, GRC, HUN, ITA, NOR, POL, PRT, SVN, and SWE (plus JPN, KOR, and USA for industries belonging to the global bucket).
Source: Authors' calculations.

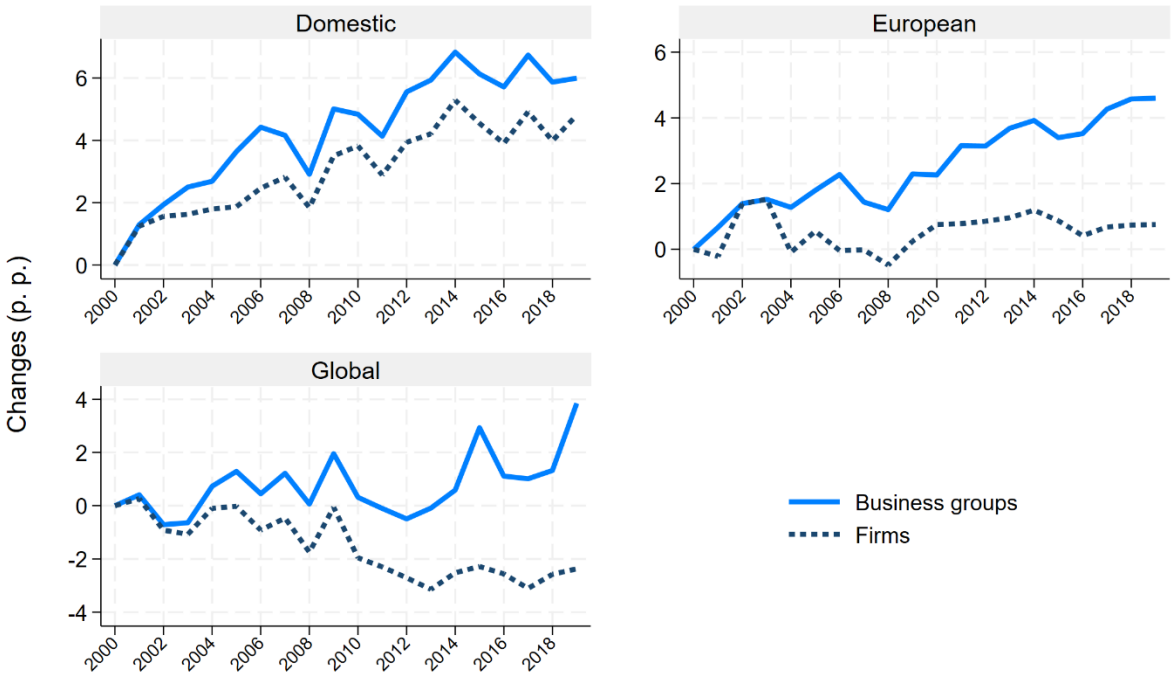
Figure 10. Concentration levels, firms vs. business groups



Note: The chart shows the unweighted average across industries and countries of CR4 levels across geographical buckets. The solid blue lines refer to the baseline average cumulative change. The dashed ones represent concentration when the ownership structure of business groups is neglected. Industries are a mix of 2 and 3-digit industries belonging to mining, manufacturing, utilities, and non-financial market. The countries included in the sample are BEL, DEU, DNK, ESP, FIN, FRA, GBR, GRC, HUN, ITA, NOR, POL, PRT, SVN, and SWE. Industries competing at the global level also include JPN, KOR, and USA.

Source: Authors' calculations.

Figure 11. Concentration cumulative changes, firms vs. business groups



Note: The chart shows the unweighted average across industries and countries of cumulative change in CR4 across geographical buckets. The solid blue lines refer to the baseline average cumulative change. The dashed ones represent concentration when the ownership structure of business groups is neglected. Industries are a mix of 2 and 3-digit industries belonging to mining, manufacturing, utilities, and non-financial market. The countries included in the sample are BEL, DEU, DNK, ESP, FIN, FRA, GBR, GRC, HUN, ITA, NOR, POL, PRT, SVN, and SWE. Industries competing at the global level also include JPN, KOR, and USA.
Source: Authors' calculations.

Data Appendix.

Figure A 1 shows concentration levels (top panel) and cumulative changes (bottom panel) weighted by market gross output (country-industry for the domestic bucket, industry for the European and global buckets) within each geographical bucket. The comparison of unweighted levels of concentration (Figure 3) with the weighted ones (Figure A 1) reveals that weighting by market size (in terms of gross output) reduces the level of aggregate concentration in industries competing at the domestic level, suggesting that concentration is higher in smaller markets (country-industries pairs in this case). On the contrary, the weighting procedure increases the aggregate concentration level in the global bucket and, to a lower extent, also in the European bucket, indicating that in international markets concentration is higher in relatively bigger industries. In addition, the comparison of unweighted and weighted

concentration cumulative changes (bottom panels of Figure 3 and Figure A 1) shows that, when weighting for the relative size of the markets, concentration looks overall flat over the period 2000-2019 for industries competing both domestically and at the European level. In industries competing at the global level, the weighted trend is decreasing. The combination of these results suggests that for the domestic and European buckets, the increase in concentration mainly occurs in relatively small markets (in terms of gross output), while for the global bucket, the decrease occurs in relatively big industries.

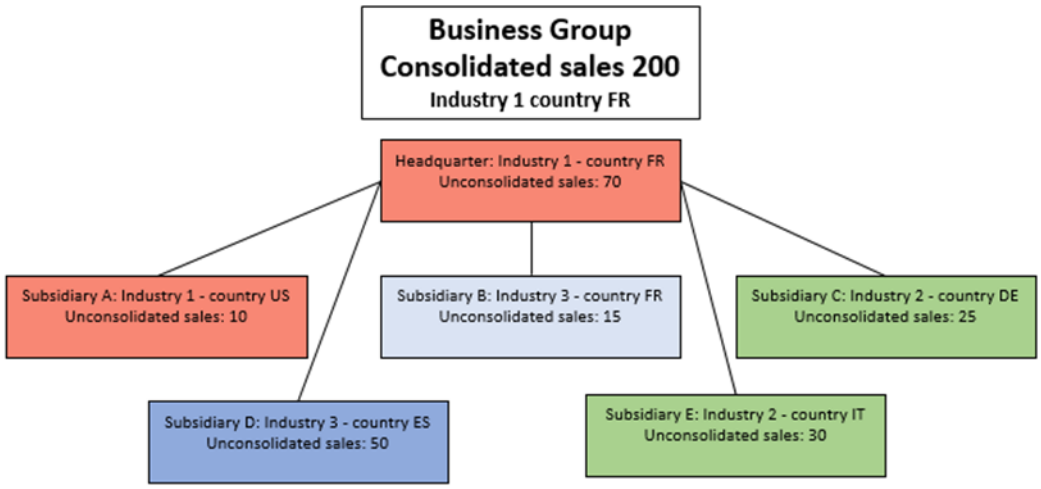
The aggregate trends reported in Figure 2 and Figure 3 can hide substantial heterogeneity across industries (and countries, for the domestic bucket). Therefore, extensive sensitivity tests have been performed to determine whether these trends are driven by a few observations. The main exercise performed to study heterogeneity and robustness across industries involved removing industries from the sample one at a time to assess their relative impact on the cumulative changes trends.[46] In addition, for industries competing at the domestic level, where the country dimension can also be investigated, a similar exercise has been conducted by removing one country at a time.[47] All in all, the trends appear to be robust to the exclusion of single industries and countries, as no single countries or industries drive the cumulative changes of concentration in any of the three geographical buckets.

As a final robustness check the Figures and Tables

[46] Domestic bucket: only three industries marginally affect the average cumulative change trend when removed, either by decreasing it by 1 p.p. (091, Support activities for petroleum and natural gas extraction; 352, Manufacture of gas, distribution of gaseous fuels through mains) or by increasing it, again, by 1 p.p. (353, Steam and air conditioning supply). European bucket: four industries that make the cumulative change in concentration increase by about 1 p.p. each when removed from the sample (232, Manufacture of refractory products; 242, Manufacture of tubes, pipes, hollow profiles and related fittings, of steel; 781 Activities of employment placement agencies; 783 Other human resources provision) and one that makes the cumulative change in concentration decrease by almost 2.5 p.p. (262, Manufacture of computers and peripheral equipment). Global bucket: only two industries marginally drive the trends in two opposite directions. Dropping from the sample industry 151 (Tanning and dressing of leather; manufacture of luggage, handbags, saddlery, and harness; dressing and dyeing of fur) decreases the cumulative change of the global bucket by about 1.5 p.p., while dropping industry 303 (Manufacture of air and spacecraft and related machinery) increases the trend by about 2 p.p.

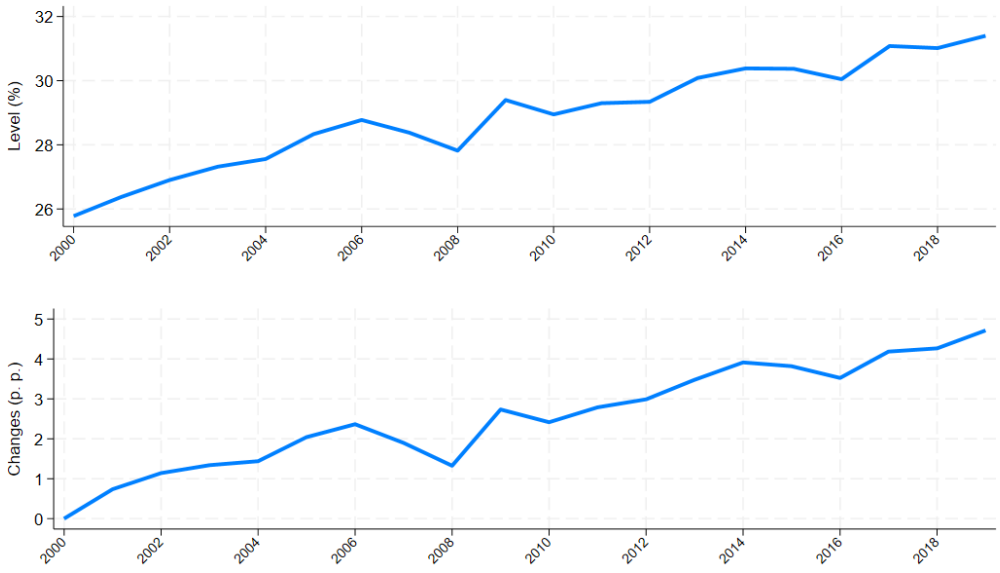
[47] Only Poland, Portugal, Sweden, and the United Kingdom can be considered as marginally driving the overall growth in the domestic bucket (excluding Poland or Sweden decreases the cumulative change by about 1 p.p. each, while for Portugal and the United Kingdom the opposite is true, excluding them would increase cumulative change in concentration by 1 p.p.).

Figure 1. Example of the apportioning technique of business group activities



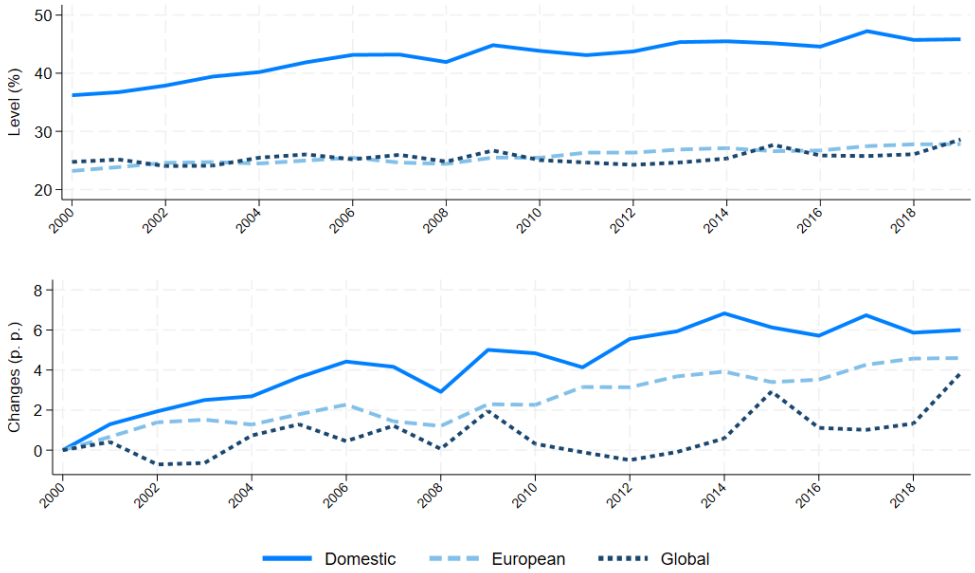
Note: The figure depicts an example of a hypothetical group consisting of a parent company from France and operating: i) in industry 1 (global), with a US subsidiary in the same industry; ii) in industry 2 (European) with two subsidiaries, one from Germany and one from Italy; and finally, iii) in industry 3 (domestic) with a subsidiary in France and one in Spain. The different colours identify the four different markets in which the group is active: the global market in industry 1 (with total gross output of 80), the European market in industry 2 (with total gross output of 55), the French market in industry 3 (with total gross output of 15), and the Spanish market in industry 3 (with total gross output of 50).

Figure 2. Concentration aggregating across geographical buckets



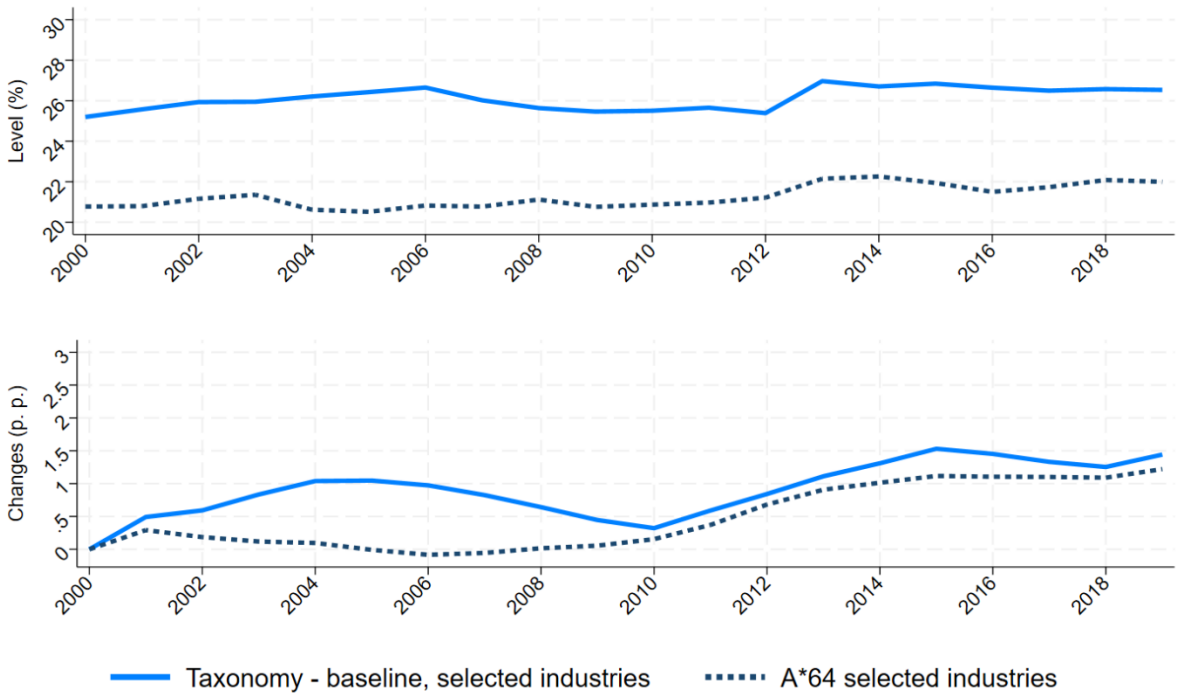
Note: The chart shows the weighted average across industry-geography combinations of CR4 levels (top panel) and cumulative growth (bottom panel). A weight of 1 is attributed to country-industries cells for domestic industries, while a weight of 15 is attributed to industries belonging to the European and the global buckets. Industries included in the analysis are a mix of 2 and 3-digit industries belonging to mining, manufacturing, utilities, and non-financial market services and are classified as either domestic, European, or global, depending on the taxonomy developed by Calligaris et al. . The countries included in the sample are BEL, DEU, DNK, ESP, FIN, FRA, GBR, GRC, HUN, ITA, NOR, POL, PRT, SVN, and SWE for the domestic and European buckets, while in the global one also JPN, KOR, and USA are included.
Source: Authors' calculations.

Figure 3. Concentration across geographical buckets



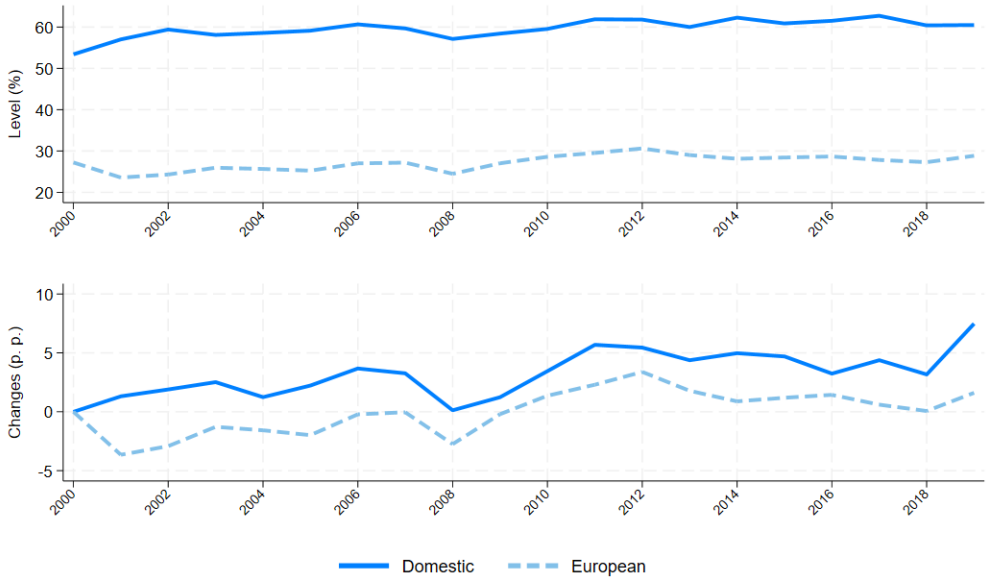
Note: The chart shows the unweighted average across industries (and countries, for the domestic bucket) of CR4 levels (top panel) and cumulative growth (bottom panel). Industries included in the analysis are mix of 2 and 3-digit industries belonging to mining, manufacturing, utilities, and non-financial market services and are classified as either domestic, European, or global, depending on the taxonomy developed by Calligaris et al. . The countries included in the sample are BEL, DEU, DNK, ESP, FIN, FRA, GBR, GRC, HUN, ITA, NOR, POL, PRT, SVN, and SWE for the domestic and European bucket, while in the global one also JPN, KOR, and USA are included.
Source: Authors' calculations.

Figure 4. Concentration with different industry aggregation levels (selected industries)



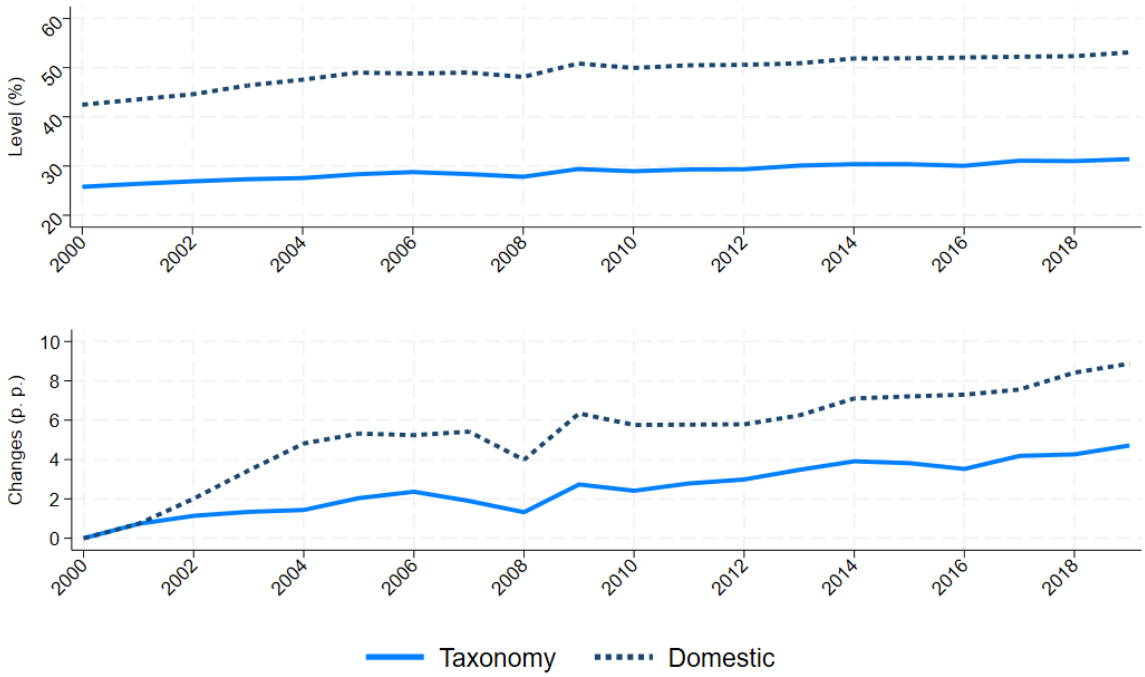
Note: The chart shows the average across industry-geography combinations of CR4 levels (top panel) and cumulative growth (bottom panel) in a selection of industries when using the industry aggregation of the baseline measure (continuous blue line, mostly 3-digit) and when using a higher-level industry aggregation (dotted dark line, mostly 2-digit). A weight of 1 is attributed to country-industries cells for domestic industries, while a weight of 15 is attributed to industries belonging to the European and the global buckets. Industries included in the analysis include all the 2-digit industries (following the STAN A*64 classification) whose 3-digit sub-industries belong to a unique geographical bucket. The 2-digit industries are, for the domestic bucket: C33, D35, D36, D37, H55, J59, and J61; for the European bucket: C16, C17, C19, C20, C22, C24, C27, C28, C29, G45, G46, G47H25, H53, J62, M71, M72, M73, N77, N78, and N79. The countries included in the sample are BEL, DEU, DNK, ESP, FIN, FRA, GBR, GRC, HUN, ITA, NOR, POL, PRT, SVN, and SWE.
Source: Authors' calculations.

Figure 5. Concentration for selected 4-digit industries



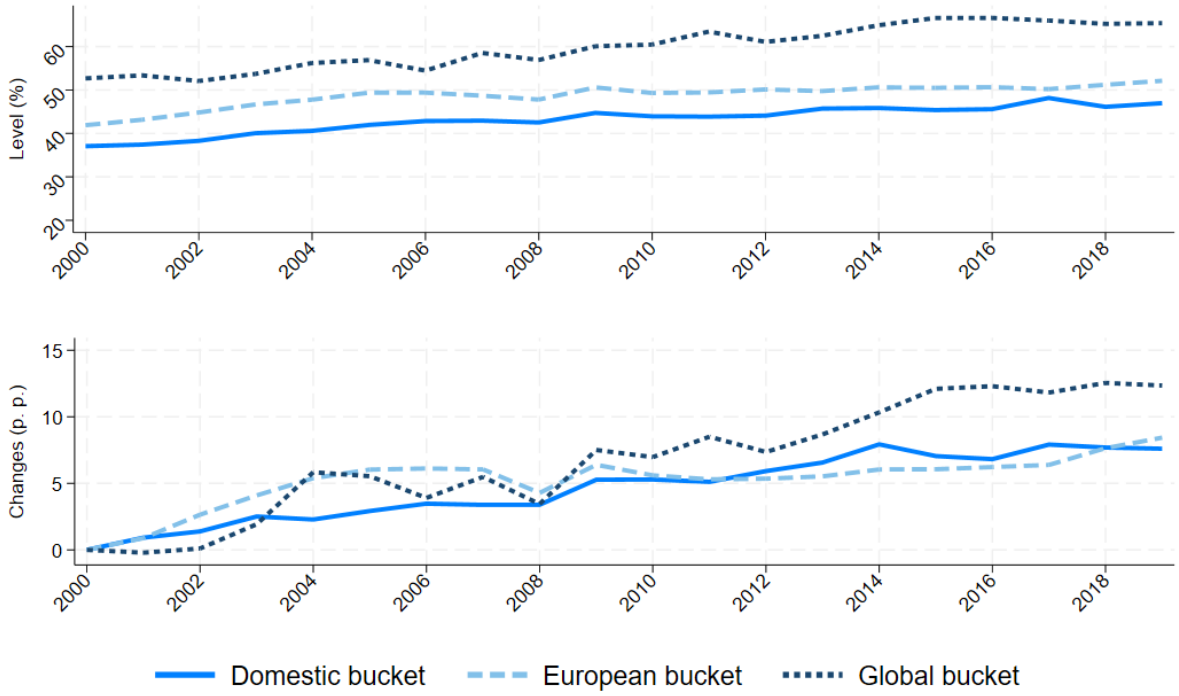
Note: The chart shows the unweighted average across industries (and countries, for the domestic bucket) of CR4 levels (top panel) and cumulative growth (bottom panel). Industries included in the analysis are all the 4-digit industries (except for 1042 and 1089 due to data limitations) belonging to the 2-digit industries from C10 (manufacture of food products) and C11 (manufacture of beverages). The countries included in the sample are BEL, DEU, DNK, ESP, FIN, FRA, GBR, GRC, HUN, ITA, NOR, POL, PRT, SVN, and SWE.
Source: Authors' calculations.

Figure 6. National concentration aggregating over different geographical regions



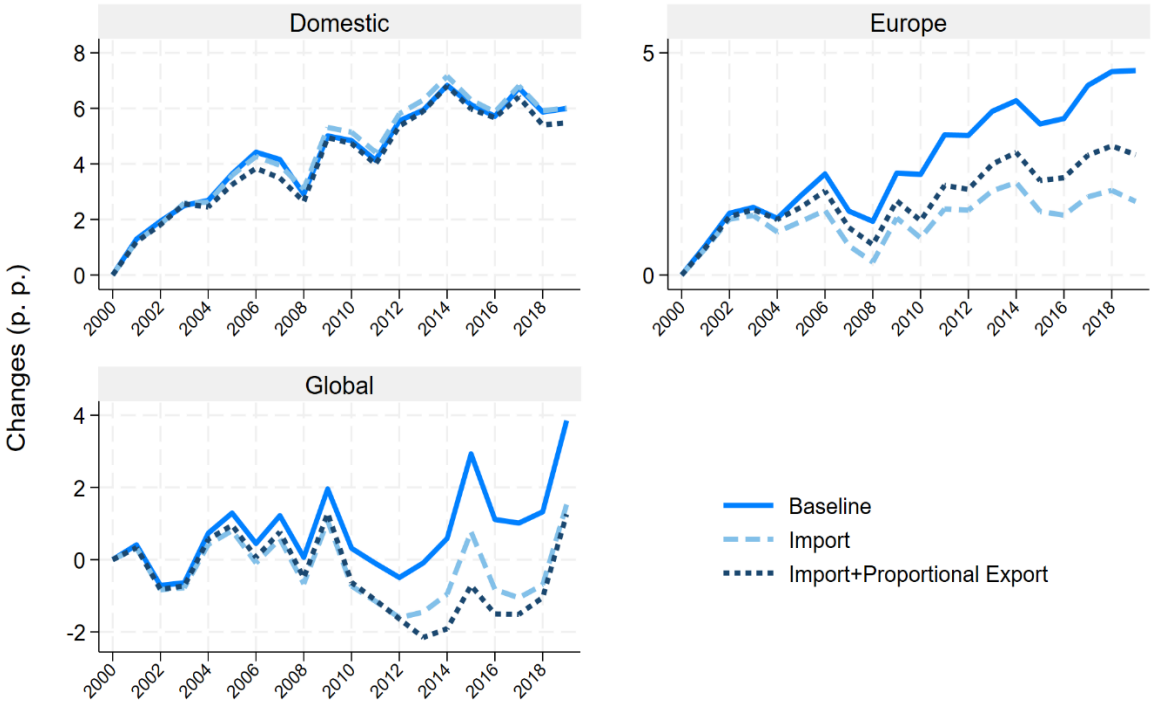
Note: The chart shows the average across industry-geography combinations of CR4 levels (top panel) and cumulative growth (bottom panel) when the taxonomy is adopted (continuous blue line) and with all industries considered as domestic ones (dotted dark line). When the taxonomy is used, a weight of 1 is attributed to country-industries cells for domestic industries, while a weight of 15 is attributed to industries belonging to the European and the global buckets. Industries included in the analysis are mix of 2 and 3-digit industries belonging to mining, manufacturing, utilities, and non-financial market services. The countries included in the sample are BEL, DEU, DNK, ESP, FIN, FRA, GBR, GRC, HUN, ITA, NOR, POL, PRT, SVN, and SWE for the domestic and European bucket, while in the global one also JPN, KOR, and USA are included.
Source: Authors' calculations.

Figure 7. National concentration across geographical buckets (all industries as domestic)



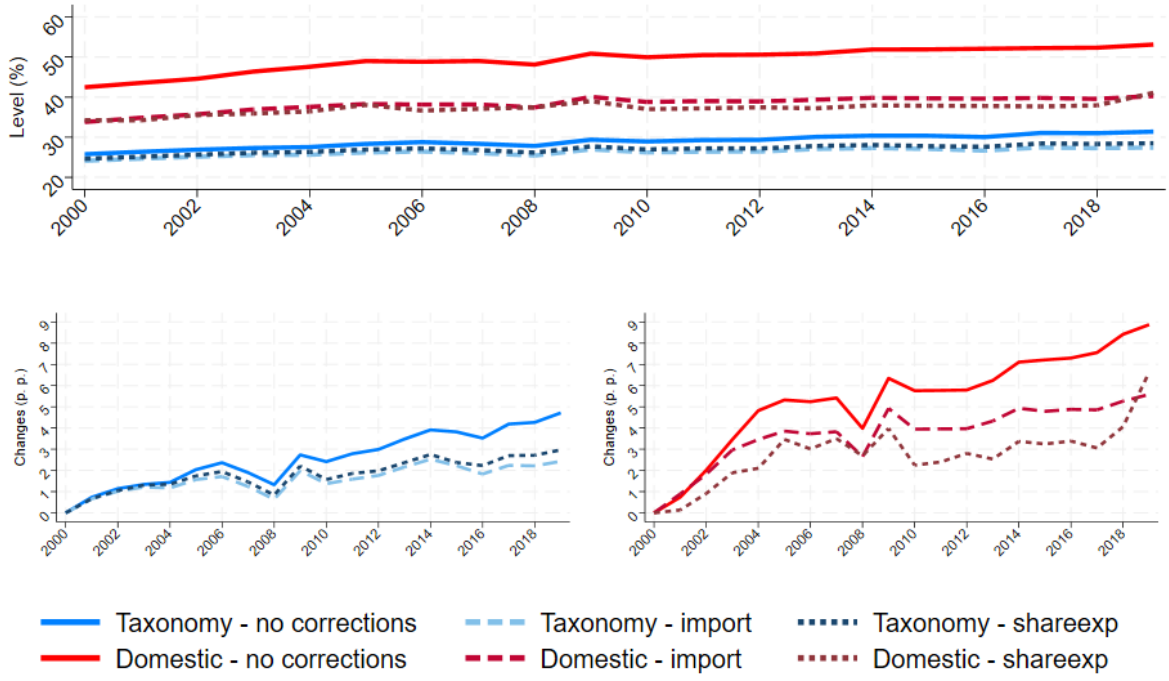
Note: The chart shows the unweighted average across industries and countries and within geographical buckets of CR4 levels (top panel) and cumulative growth (bottom panel), computed treating all industries as if they were competing at the domestic level. Industries included in the analysis are mix of 2 and 3-digit industries belonging to mining, manufacturing, utilities, and non-financial market services. The countries included in the sample are BEL, DEU, DNK, ESP, FIN, FRA, GBR, GRC, HUN, ITA, NOR, POL, PRT, SVN, and SWE for the domestic and European bucket, while in the global one also JPN, KOR, and USA are included.
Source: Authors' calculations.

Figure 8. Different trade adjustments



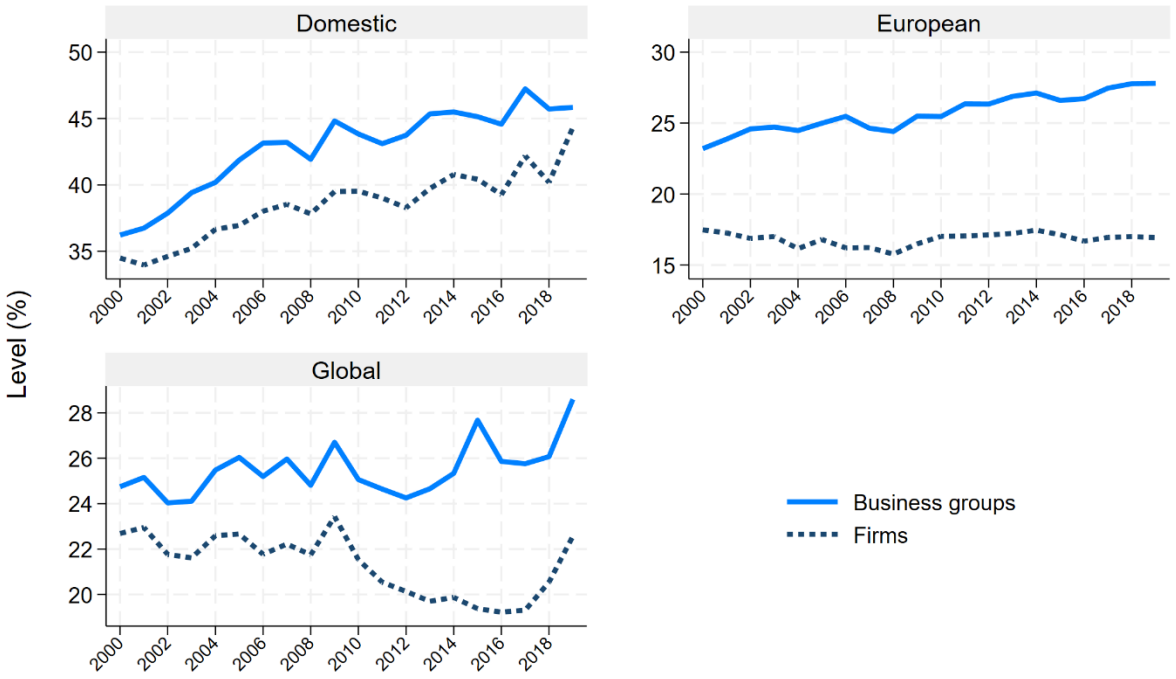
Note: The chart shows the unweighted average across industries and countries of cumulative change in CR4 across geographical buckets for different types of trade adjustments. The solid blue lines refer to the baseline average cumulative change (no corrections). The light blue dashed ones to the correction obtained by adding import at the denominator. The dark blue dotted ones to the correction obtained by adding import and subtracting total exports at the denominator, and by subtracting at the numerator a fraction of export given by the share of gross output accounted for by the four largest firms in the market. Industries are a mix of 2 and 3-digit industries belonging to mining, manufacturing, utilities, and non-financial market. The countries included in the sample are BEL, DEU, DNK, ESP, FIN, FRA, GBR, GRC, HUN, ITA, NOR, POL, PRT, SVN, and SWE. Industries competing at the global level also include JPN, KOR, and USA.
Source: Authors' calculations.

Figure 9. Concentration level different trade corrections (taxonomy vs. all as domestic)



Note: the chart shows the unweighted average across industries and countries of CR4 levels for different types of trade adjustments. Blue lines refer to concentration computed using the taxonomy developed by Calligaris et al., red lines to concentration computed with all industries considered as domestic. Solid lines refer to the baseline average level (no corrections). Dashed ones to the correction obtained by adding import at the denominator. Dotted ones to the correction obtained by adding import and subtracting total exports at the denominator, and by subtracting at the numerator a fraction of export given by the share of gross output accounted for by the four largest firms in the market. Industries are a mix of 2 and 3-digit industries belonging to mining, manufacturing, utilities, and non-financial market. The countries included in the sample are BEL, DEU, DNK, ESP, FIN, FRA, GBR, GRC, HUN, ITA, NOR, POL, PRT, SVN, and SWE (plus JPN, KOR, and USA for industries belonging to the global bucket).
Source: Authors' calculations.

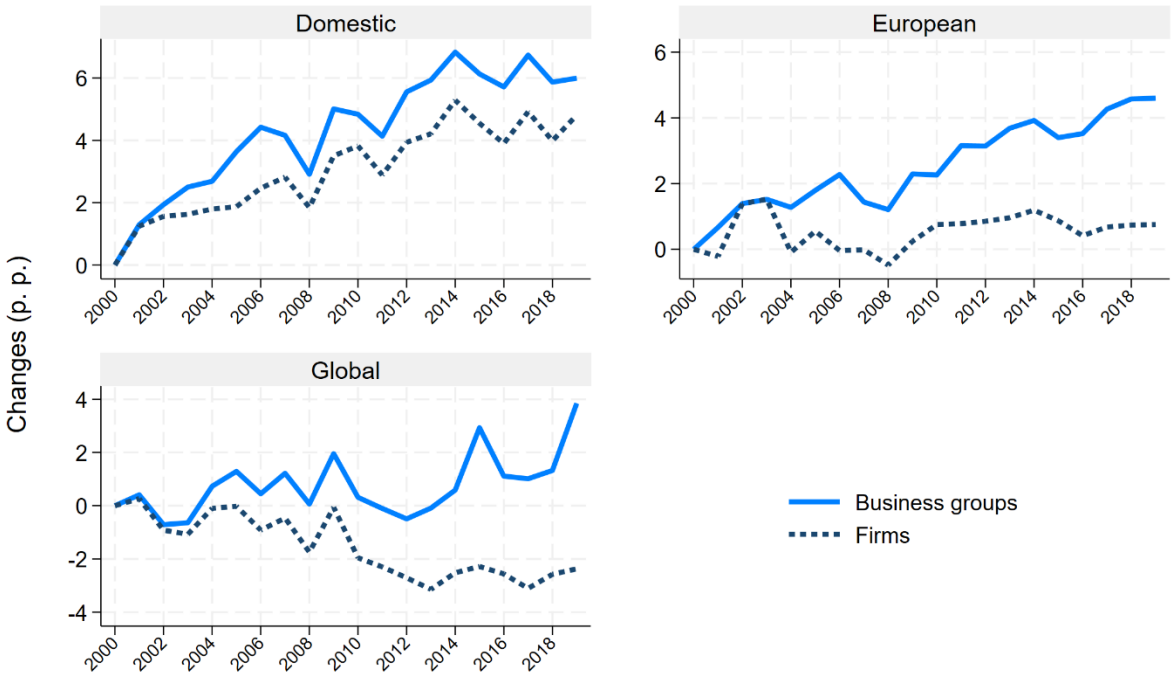
Figure 10. Concentration levels, firms vs. business groups



Note: The chart shows the unweighted average across industries and countries of CR4 levels across geographical buckets. The solid blue lines refer to the baseline average cumulative change. The dashed ones represent concentration when the ownership structure of business groups is neglected. Industries are a mix of 2 and 3-digit industries belonging to mining, manufacturing, utilities, and non-financial market. The countries included in the sample are BEL, DEU, DNK, ESP, FIN, FRA, GBR, GRC, HUN, ITA, NOR, POL, PRT, SVN, and SWE. Industries competing at the global level also include JPN, KOR, and USA.

Source: Authors' calculations.

Figure 11. Concentration cumulative changes, firms vs. business groups



Note: The chart shows the unweighted average across industries and countries of cumulative change in CR4 across geographical buckets. The solid blue lines refer to the baseline average cumulative change. The dashed ones represent concentration when the ownership structure of business groups is neglected. Industries are a mix of 2 and 3-digit industries belonging to mining, manufacturing, utilities, and non-financial market. The countries included in the sample are BEL, DEU, DNK, ESP, FIN, FRA, GBR, GRC, HUN, ITA, NOR, POL, PRT, SVN, and SWE. Industries competing at the global level also include JPN, KOR, and USA.
Source: Authors' calculations.

Data Appendix reports the levels and trends for CR8. The average levels (top panel of Figure A 2) are higher by construction, the trends are qualitatively similar to those reported for CR4: the domestic bucket displays the highest level of concentration, with the top 8 firms accounting on average for about 50% of the market output, followed by the European and the global buckets, where average concentration level is around 35%. When looking at the cumulative changes (bottom panel of Figure A 2) the differences with CR4 are slightly more pronounced, especially for industries competing at the European level: CR8 increased more than CR4 in all the geographical buckets. It increases similarly, about 8 p.p., in the domestic and the European bucket (vs. 6 and 4 p.p. respectively, for CR4), and by about 4.4. in the global bucket (similarly to CR4).

4.2. The impact of different methodological assumptions on measured concentration

This sub-section describes the impact of different methodological choices made in the construction of the concentration measure. These alternative measures depart from the baseline's assumptions in terms of geographical aggregation, industry aggregation, trade corrections, and firm's boundaries.

4.2.1. The role of industry aggregation

A key contribution of this paper is the computation of concentration at a highly disaggregated industry level, aimed at approximating as close as possible the "relevant" market. To test the effect of changing the level of industry aggregation on aggregate trends and compare the baseline results of the paper with those in previous cross-country literature (see, for instance, Bajgar et al. (2023[6])), concentration has been also computed at the 2-digit level (following the STAN A*64 classification). This exercise considers all the 2-digit industries whose 3-digit sub-industries belong to a unique geographical bucket.[48] This selection is required to isolate the effect of changing industry aggregation, while keeping the geographical dimension of the market definition fixed. Therefore, this exercise is necessarily based on a restricted sample of industries with respect to the ones considered elsewhere and, as a result, the aggregate figures are qualitatively different from the trends of the preferred baseline measure.[49]

The continuous blue line in the top panel of Figure 4 represents the baseline average concentration level (that is, using the selected industries at the 3-digit level), while the dotted line represents the average concentration computed at the 2-digit industry level of aggregation. Defining concentration at a more disaggregated level raises the average level of concentration from 21% to about 29%. A possible interpretation of these differences is the following. When looking at broad industry levels, for example at the 2-digit level, even the largest business

[48] For example, it was possible to consider the 2-digit industries C33 (Repair and installation of machinery and equipment), whose 3-digit industries (C331 and C332) all belong to the domestic bucket, and C29 (Manufacture of motor vehicles, trailers and semi-trailers), whose 3-digit industries (C291, C292, and C293) all belong to the European bucket. Conversely, it was not possible to include mining and quarrying (A08), since one 3-digit industry competes at the European level and the other at global level.

[49] In particular, the 2-digit industries considered are the following (in terms if NACE Rev.2 classification); for the domestic bucket: C33, D35, D36, D37, H55, J59, and J61; for the European bucket: C16, C17, C19, C20, C22, C24, C27, C28, C29, G45, G46, G47H25, H53, J62, M71, M72, M73, N77, N78, and N79.

groups are unlikely to be active and sell in all the different activities performed in the industry. When industries are more disaggregated, and therefore more specialised, large players are more likely to produce in all the economic activities within the narrow industry. Therefore, at higher levels of disaggregation concentration is likely to be higher. The bottom panel of Figure 4 shows that in terms of cumulative changes in aggregate concentration, the trends remain similar across different levels of industry aggregation. Due to the smaller sample, cumulative changes in this exercise are more volatile than in the baseline measure.

In a second attempt to understand the importance of industry aggregation when computing concentration, an analysis at the 4-digit industry level has been performed for some selected industries. Due to data limitations, the exercise has focused only on the manufacture of food and beverages (C10 and C11 at the 2-digit level of aggregation). These industries are a particularly relevant component of the manufacturing sector, both in terms of employment and value added, and in recent years they have been under scrutiny by the European Commission (EC) with the aim of improving the state of competition.[50] In addition, the 4-digit industries belonging to the food and beverages industries reach a granular level of detail which are very similar to the equivalent product markets (confirmed both by EC Competition Directorate and from comparing with product market classification systems).

According to the taxonomy of sectors, all the 4-digit industries belonging to the 3-digit industries C101, C105, C107, C109, and C110 are classified as competing at the domestic level, while those belonging to the 3-digit C102, C103, C104, C106, C108 are classified as competing at the European level.[51] Concentration is therefore computed at the corresponding geographical level for each 4-digit industry.

The top panel of Figure 5 shows the average concentration level for these industries in the domestic and European buckets. The average level in the European bucket (about 27%) is higher than the average level found in the corresponding industries at the 3-digit level of aggregation (about 20%); the average level in the domestic bucket is higher (around 59%) than in the corresponding 3-digit (42%).

[50] Please refer to the webpage https://competition-policy.ec.europa.eu/sectors/agriculture/food_en for a discussion of these topics and further links to relevant work.

[51] A table with the description of the industries considered is provided in the Annex (see Table A 1 and Table A 2).

The bottom panel of Figure 5 reports the average cumulative growth across these industries. As in the baseline results, industries belonging to the domestic geographical buckets experienced, on average, a higher concentration growth than those competing at the European level in the years between 2000 and 2019.

Note that the conclusions drawn from the 4-digit data must be interpreted with some caution, given that with this exercise data have been stretched to their maximum. For example, the range of activities of firms at this level of disaggregation might indeed easily fall into different 4-digit industries, but in Orbis all firms are registered as active only in their main industry of activity. While conceptually this can happen even at higher levels of aggregation, this issue is much more likely to happen when disaggregating at the 4-digit level. Registering all activities in one single 4-digit industry might result in higher market share, and potentially in concentration becoming higher than 1. At the same time, denominators are obtained following the same methodology outlined in Section 3. As explained above, several assumptions and imputations have been made to obtain data of gross output consistent across countries and industries over the years at such disaggregated levels.[52] With this caveat in mind, this exercise is interesting because it allows comparing industry concentration with a measure of concentration even closer to the product level.

[52] For a more thorough discussion on the relationship between industry concentration and product market concentration please refer to Calligaris et al. (2024[34]).

4.2.2. The role of geography

A key contribution of this paper is to compute concentration while considering the geographical boundaries of the market. In most of the previous studies, concentration is measured by assuming that all industries compete at the domestic level, i.e., at the national level.[53]

The top panel of Figure 6 compares the aggregate baseline concentration levels (continuous blue line, also portrayed in Figure 2) with the aggregate concentration computed by assuming that all industries at the domestic level (dotted line). The average level is consistently higher when concentration is computed as if all industries were domestic, with an average over the period considered of about 49%, against an average level of about 29% when the taxonomy is applied. This is in line with the findings showing that concentration levels are higher when computed within more disaggregated markets, as discussed in the previous sub-section. When computing concentration using the taxonomy, the sales of the subsidiaries of a business group active in the same industry are summed over the countries where they are active. The numerator of the concentration ratio includes the activities of the largest four groups, with their activities summed across all countries even if their subsidiaries are not in the top four in every market. In contrast, when concentration is computed at the national level, the activities of the largest four business groups in each country are considered (so the same business group in different countries is considered as two distinct ones), but their activities in subsidiaries from outside the country are not considered unless each subsidiary is also in the top four in other countries.[54] Therefore, if the contribution of including the subsidiaries is less important than the contribution of including the four largest firms in each country, the average of domestic concentration is likely to be larger than concentration computed at the international level.

[53] Three notable exceptions being Rossi-Hansberg et al. (2021[36]) and Autor et al. (2023[39]) who distinguish between local and national trends, and Bajgar et al. (2023[6]), who also compute concentration assuming that all industries in European countries compete at the European level.

[54] For example, suppose business group X has in a certain industry a headquarter X-A which is in the top four in France and a subsidiary X-B that is not in the top four Spain. Business group X is in the top four at the European level and the contributions of both X-A and X-B would be considered in the numerator. However, at the domestic level in Spain, a different company, Y, is in the top four in the same industry. When computing domestic concentration for this industry, the activities of Y are considered but X-B are not. The overall comparison of concentration at the domestic and European level will depend on the trade-off between these two factors.

The bottom panel of Figure 6 compares the cumulative growths over the period considered. Concentration increased by about 5 p.p. in the baseline framework, and about 9 p.p. when all industries are considered as competing domestically. This corresponds to a percentage increase: of 25% when all industries are considered as competing domestically, and of about 21% in the baseline framework. Thus, the definition of the relevant market influences the average level of concentration, with concentration being higher when markets are considered only domestic. Concentration is found to be increasing in both exercises, though more significantly when all industries are classified as competing domestically.

The impact of the level of geographical aggregation on the concentration measures is further explored by computing concentration at the national level for all industries but averaging over the different regions implied by the taxonomy. This means that concentration in domestic markets is the same as the baseline one in Figure 3 (because it was already computed averaging the country-industry levels). In contrast, for industries belonging to the European and the global buckets, in this exercise the average concentration across the relevant countries is considered.[55]

The results of the exercise are reported in Figure 7. Accounting for the effective geographical boundaries of a market affects both levels and trends. Figure 7 shows that, relative to the baseline in Figure 3, average national level concentration is higher in European and global markets relative to domestic ones, with an average level of 49% and 60%, respectively. The cumulative changes, reported in the bottom panel of Figure 7, are consistent with the overall view that concentration has increased independently of the geographical boundaries of markets. Differently from the baseline measure, concentration computed at the national level has increased relatively more in industries classified as competing globally (about 12 p.p.), followed by those competing at the European and then domestic level (respectively, around 8.5 and 7.6 p.p.). In terms of percentage growth rates, domestic concentration increases rather uniformly in the three geographical buckets.[56] The fact that the ranking of concentration levels and growth by geographical bucket undergoes a reversal with respect to the baseline when

[55] As for the baseline taxonomy measure, JPN, KOR, and USA are accounted for in global industries but excluded for domestic and European ones.

[56] Domestic, European, and global industries increased, on average, of about 26% (domestic) and 24% (European and global).

computing concentration at domestic level for all industries, highlights the importance of accounting for integration of markets thanks to the taxonomy.

These results also suggest that the largest firms competing in European and global markets are playing an increasingly important role in national economies, and more so than firms competing on national markets. The observation that firms in globally-competing industries have both the highest level and growth of concentration is evidence in support of Melitz (2003[19]), whereby international competition causes a reallocation of resources to the largest and most productive firms.

4.2.3. The role of trade corrections

As mentioned, some recent literature (Amiti and Heise, 2024[12]) shows that accounting for international trade might significantly affect concentration trends in the US.

Therefore, a robustness check that accounts for the role of imports and exports is performed (see Section 3 for details). However, this robustness check should be interpreted with caution since, due to data limitations, some important assumptions are made. While Amiti and Heise (2024[12]) have access to confidential firm-level data on sales to the US market of both domestic and foreign firms, this type of data is not available in a cross-country setting.[57] Thus, as detailed in Section 3, this robustness analysis needs to rely on strong assumptions to estimate the amount of trade conducted by the top four firms in a market and carry out these robustness checks. The lack of firm-level trade data, combined with the fact that the taxonomy already indirectly accounts for most of international trade in each geographic region, are the main reasons why the paper's baseline measure does not include the trade corrections presented in this robustness exercise. Nonetheless, the effects of additional trade adjustments on concentration are presented.[58]

Figure 8 reports the results of the trade correction exercise across the three geographical buckets (domestic in the top-left panel, European in the top-right panel, and global in the bottom-left one). In all the panels: i) the solid blue lines report the baseline specification with no correction for international trade (same as the baseline trends reported in Figure 3); ii) the light blue dashed lines include the import correction in the denominator, but no export correction; iii) the dark blue dotted lines account for both the import and the export corrections. Specifically, in the import correction exercise, imports in the same industry from outside the geographical region (domestic, European, global) are added to the denominator, as are imports

[57] In many OECD countries information on trade flows is not available at the firm level as it is collected at the product level by customs and not linked to firm level financials. Note also that this information is often only available for goods trade and not for services trade as it is the case for Amiti and Heise (2024[12]).

[58] Also note that, as described above, while the taxonomy accounts for trade between all countries included in the relevant geography, not all European and (especially) global countries are included in the sample. If data were available for every country in the world, there would be no need for a trade correction in global industries. Trade with countries not included in the sample is accounted for in the trade adjustments, explaining why the further adjustments have a non-trivial effect on the trends (in addition to the caveat that only aggregate, not firm-level, data on trade flows is available so these adjustments are an approximation).

from countries that are in the region but not included in the sample due to data constraints.[59] In the exercises in which exports are also accounted for, exports in the same industry to the rest of the world (with respect to the geographic region considered) are fully subtracted from the denominator (to account only for the gross output that is consumed in each market), while in the numerator a share of exports corresponding to the share of gross output accounted for by the four largest firms in the market is subtracted (since data do not provide information on the export flows at the firm level). This assumption is likely to underestimate the share of trade accounted for by the top four firms, as larger firms tend to export more (Bernard et al., 2012[47]). This means that the level of concentration with this correction may be overestimated, but it is not clear what the effect on the trend would be, as it depends on the evolution of the relative contribution of the top four firms to exports compared with the one of their overall market shares.

In line with Amiti and Heise (2024[12]), trade corrections tend to dampen the rise of concentration. As expected, trade corrections affect relatively more the trends of industries that compete at the European and global levels, which are more tradeable (see the top-right panel and the bottom-left one in Figure 8). Being more exposed to international trade, these industries are likely to have higher trade flows from outside the geographic region of competition. The growth of concentration in the domestic bucket (see top- left panel of Figure 8) remains virtually unchanged because trade flows are small. This adds further support to the validity of the taxonomy developed by Calligaris et al. (2024[17]). Moreover, the diverging trends of the baseline measure with respect to the trade-corrected measures suggest an increase in import competition over the period considered.

[59] While there are likely to be many foreign firms that sell in a market, they are likely to be small on average in terms of their sales relative to firms with a domestic presence (noting that foreign-owned firms with domestic subsidiaries would be considered as domestic firms). Therefore, they are likely to impact the total market size, so imports are added to the denominator to capture the full size of a market. However, as they are small on average, they are unlikely to enter the top four and are therefore not considered in numerator of this measure. This is a necessary assumption due to data constraints and is in line with the evidence of Amiti and Heise (2024[12]).

In a second exercise trade corrections are applied to aggregate concentration, both computed using the taxonomy and by considering concentration at the national level for all markets.[60] This exercise allows to test the robustness of the baseline concentration measure to applying indirect (taxonomy) and direct trade corrections. If the taxonomy accounts for most of the relevant trade, trade corrections should have a lower impact on the baseline measure than on the alternative concentration measure computed at the national level independently of whether the industry competes at the domestic, European or global level.

The top panel of Figure 9 reports the results of the exercise for the average level of concentration: the continuous blue line represents the baseline concentration measure based on domestic, European and global concentration levels, the continuous red line represents national level concentration for all industries, independently on whether they compete at the international level according to the taxonomy. The dashed lines (red and blue following the distinction above) represent concentration when the import correction, but not the export one, is accounted for in the denominator. The dotted lines account for both imports and exports, which are subtracted from both the denominator and (a share of) the numerator of the baseline concentration measure. Trade corrections have a much stronger effect when concentration is computed at the national level rather than when the taxonomy is used: the lines in blue are much closer to each other than the red ones. Such a differential impact is reassuring because it suggests that the taxonomy already accounts for the role of international trade in concentration measures. The bottom panels of Figure 9 report the analogous chart for cumulative changes instead of levels of concentration. Also, when looking at changes, trade corrections appear to have a stronger impact when concentration is computed at the national level in all industries (bottom-right panel in Figure 9). The trade corrections seem to slightly reduce concentration levels and their increase over time but both concentration measures are increasing over the period considered. Most importantly the results confirm the validity of the taxonomy to account for trade flows in the measurement of concentration levels and trends.

[60] See Section 4 on the role of industry aggregation for further details about the two measures.

4.2.4. Accounting for business groups activities

Finally, the last sub-section explores the impact on concentration of accounting for the ownership structure of business groups within each market. In the baseline concentration measure, subsidiaries belonging to the same business group and active in the same market are treated as a unique entity. First, even in domestic industries, neglecting the ownership structure would lead to an underestimation of concentration because subsidiaries belonging to the same business group in the same industry would otherwise be considered as independent entities. Second, in tradeable industries the expansion of a business group into new countries would not be captured as an increase in its market share (and, therefore, of concentration) via the new subsidiary would be ignored as the new subsidiary would be treated as an independent entity. Therefore, analysing business groups rather than firms helps capture potential increases in concentration due to the organic growth of a group's sales via growth of existing entities and greenfield investments as well as its expansion through M&A.

To test the impact of this assumption, the baseline concentration measure is compared with a measure of concentration that ignores the business group dimension. Instead of summing the activities of all the firms within the same business group in each market, this approach considers the activities of all the firms in a business group as if they were independent entities (single firms). Therefore, the numerator of this alternative concentration measure will be the sum of sales of the largest four firms (rather than business groups) in a market.

Figure 10 and Figure 11 show the comparison of the baseline measure (solid line) with the alternative measure of concentration, where the ownership structure of business groups is ignored (dashed line). Figure 10 shows that, as expected, neglecting the business group dimension leads to lower levels of concentration in every geographical bucket, with the difference being more evident in European markets. Moreover, Figure 11 shows that the cumulative changes of the two measures are diverging across the three geographical buckets, with the baseline measure showing the highest cumulative growth. Importantly, the divergence between the two measures is particularly evident in the European and global buckets. It is worth noting that while concentration remains increasing in the domestic and European buckets when using the alternative measure, concentration is decreasing in the global one. This is most likely due to the inability of this alternative measure of capturing the expansion of multinational activities across countries via greenfield or brownfield (i.e. via M&A activity) investments.

Overall, this exercise shows that neglecting the business group's structure leads to a significant underestimation of both levels and growth of concentration and to different conclusions on concentration trends, especially in industries competing internationally, as ignoring ownership structures would suggest flat (in European markets) or declining (in global markets) concentration trends. These results confirm the importance of considering the ownership structure of business groups when computing concentration as the expansion of business groups within industries and across countries for European and global industries is clearly prevalent. The greater increase in concentration over time when accounting for business groups suggests that this practice has clearly gained importance in the last two decades, especially across European and global markets.

5. Conclusions

Concentration is a key indicator of competition. Although conceptually intuitive, its computation involves several methodological decisions that can significantly impact the results. The main contribution of the paper is to propose several methodological improvements to its measurement, with the aim of accurately capture what competition authorities refer to as the "relevant" market and investigate how each of them affects the levels and trends of concentration. These methodological refinements allow to compute market concentration, defined as the market share of leading firms, in fifteen European countries from 2000 to 2019.

First, industries are narrowly defined – mainly at 3-digit industry level. This allows the market shares of leading firms to be computed in relatively detailed markets, and thus more likely to consider only goods perceived as substitutes by consumers.

Second, for each industry, the geographical level at which firms compete is considered. Leveraging on Calligaris et al. (2024[17]), each industry is categorised as competing either at the domestic, European or global level. Concentration measures are then computed for each industry at the geographical level indicated by this taxonomy. This approach is important for two main reasons. Differently from previous studies, which typically only measure concentration at the national level, i) it considers the relevant geographical dimension at which competition takes place; and ii) it indirectly accounts for the role of international trade within a region without requiring firm-level data on international trade, which are not available in a cross-country setting.

Third, building on the procedure developed by Bajgar et al. (2019[43]), ownership linkages between firms are considered, and all the subsidiaries belonging to the same business group and active in the same market are treated as a single entity. This allows the overall importance of business groups to be accounted for. Looking at business groups rather than at individual firms is particularly important in this context, since the propensity to consolidate into groups can vary across industries and may depend on the geographical level at which competition takes place. The integration of data on ownership linkages among firms with the

information on the geographical dimension of markets contained in the taxonomy developed in Calligaris et al. (2024[17]) is a key contribution of this paper.

These methodological improvements are then used to compute concentration trends for a sample of fifteen European countries. The study finds that average concentration increased by 5 p.p. between 2000 and 2019, from 26% to slightly more than 31%. Interestingly, industries competing at the domestic level show higher and faster-growing concentration with respect to industries competing at the European and global levels. These findings, obtained in a large cross-countries sample and by applying relevant methodological improvements, align with several works that hint an overall decline of competition.

The paper further contributes to the existing literature by studying the impact of each methodological decision on concentration levels and trends in a unified framework. Determining how to appropriately measure concentration is crucial for researchers and policymakers. There is a large literature – among academics, competition authorities, governments, and international organisations – that measures trends in concentration. These studies differ in the type of data used, the country of analysis, and methodological assumptions. Consequently, even studies within the same country often draw different conclusions about concentration trends; comparing them is complicated by differences in datasets and settings.

The first methodological decision examined concerns the level of industry aggregation. The baseline measure of concentration proposed in the paper is compared to a measure which considers industries at the 2-digit level of aggregation. The exercise shows that higher granularity leads to higher concentration levels but similar trends.

Second, the paper examines the choice of accounting for the geographical boundaries of a market when computing concentration. With respect to standard concentration measures, computed exclusively at national level, the baseline concentration measure, which accounts for the taxonomy, exhibits lower levels and growth of concentration. When including the taxonomy concentration increases by about 5 p.p., compared with nearly 9 p.p. when computing concentration at the national level for all industries.

Moreover, additional trade corrections are investigated to try to identify the actual size of each market, adjusting for the value of products that leave or enter the market. However, these corrections can only be implemented under strong assumptions as firm-level trade data are not available, so they are used as a robustness check but not incorporated in the baseline

measures. In line with the innovative work of Amiti and Heise (2024[12]), adjusting for trade dampens concentration growth. However, unlike their US-focused paper, in the present paper concentration still shows an increase under each of the trade adjustments considered.

Finally, the baseline measure of concentration, computed using business group information, is compared to an alternative measure that neglects ownership linkages and considers each firm as an independent entity, as standard in the literature. Neglecting the ownership linkages substantially underestimates concentration levels and leads to lower growth from 2000 to 2019.

There are several potential avenues for future work. First, this study, like most other studies on concentration, aggregates firm-level data at the industry level. Recently, some authors have questioned whether industries reflect consumer product markets, even when they are narrowly defined. Investigating the relationships between industry-level and product-level concentration would enhance the informativeness of industry-level analyses. Forthcoming work shows that concentration measured at the industry level and at the product level are strongly positively correlated. Second, to further understand the state of competition, concentration should be investigated alongside other proxies of competition, such as business dynamics and markups. A forthcoming OECD paper shows that there is a lack of business at the top of markets and that this has slightly declined over the period. Third, the possible causes and consequences of increasing concentration should be explored, though this is challenging empirically, due to the difficulty of identifying exogenous variations in concentration. Recent OECD work shows that higher concentration is positively correlated with the intensity of intangible assets, M&As by leading firms, and burdensome regulations in upstream sectors (Calligaris et al. (2024[34])). Fourth, this study lacks access to sub-national data, which may be more appropriate for computing concentration in some non-tradeable industries and warrants further investigation. Finally, as cross-country data availability improves, the usefulness of the taxonomy increases, as does the validity and representativeness of the findings, so the study could be repeated in the future with new data.

Overall, this paper contributes to the existing literature both methodologically and empirically. Methodologically, it proposes a more precise way to measure concentration. The methodological exercises demonstrate how different assumptions impact concentration measures in a harmonised framework, shedding some light on the sources of controversial findings in the literature. From an empirical viewpoint, the paper shows that overall

concentration increased from 2000 to 2019. While the proposed refinements affect the average concentration level and its evolution over time, concentration consistently rises across almost all specifications.

The fact that this result, which informs the debate about a possible weakening of competition, is robust to many specifications has important implications for policymakers. A holistic policy framework encompassing antitrust, industrial, innovation, and broader economic policies must be adopted to ensure healthy competition.

References

Affeldt, P. et al. (2021), "Market Concentration in Europe: Evidence from Antitrust Markets", *Discussion Papers*, No. DP 15699, CEPR, http://www.cepr.org. [33]

Aghion, P. et al. (2005), "Competition and Innovation: An Inverted-U Relationship", *The Quarterly Journal of Economics*, Vol. 120/2, pp. 701-728. [3]

Altomonte, C. et al. (2021), "Business groups as knowledge-based hierarchies of firms". [48]

Amiti, M. and S. Heise (2024), "U.S. Market Concentration and Import Competition", *The Review of Economic Studies*, https://doi.org/10.1093/restud/rdae045. [12]

Autor, D. et al. (2020), *The fall of the labor share and the rise of superstar firms*, Oxford University Press, https://doi.org/10.1093/qje/qjaa004. [8]

Autor, D., C. Patterson and J. Reenen (2023), "Local and National Concentration Trends in Jobs and Sales: The Role of Structural Transformation", *NBER WP*, No. 31130, NBER, https://www.dropbox.com/s/vdax1usi2ra8ifq/concentration_trends.xlsx?dl=0. [39]

Babina, T. et al. (2023), "Antitrust Enforcement Increases Economic Activity", *NBER Working Paper*, Vol. 31597. [2]

Backus, M. (2020), "Why Is Productivity Correlated With Competition?", *Econometrica*, Vol. 88/6, pp. 2415-2444, https://doi.org/10.3982/ECTA12926. [5]

Bajgar, M. et al. (2023), "Industry concentration in Europe and North America", *Industrial and Corporate Change*, https://doi.org/10.1093/icc/dtac059/6987147. [6]

Bajgar, M. et al. (2020), "Coverage and representativeness of Orbis data", *OECD Science, Technology and Industry Working Papers*, No. 2020/06, OECD Publishing, Paris, https://doi.org/10.1787/c7bdaa03-en. [52]

Bajgar, M. et al. (2019), "Industry concentration in Europe and North America", *OECD STI WP*, OECD. [43]

Barkai, S. (2020), "Declining Labor and Capital Shares", *Journal of Finance*, Vol. 75/5, pp. 2421-2463, https://doi.org/10.1111/jofi.12909. [9]

Benkard, C., A. Yurukoglu and A. Zhang (2021), *Concentration in Product Markets*, National Bureau of Economic Research, Cambridge, MA, https://doi.org/10.3386/w28745. [13]

Benmelech, E., N. Bergman and H. Kim (2022), "Strong Employers and Weak Employees How Does Employer Concentration Affect Wages?", *Journal of Human Resources*, Vol. 57/SpecialIssue 1, pp. S201-S250, https://doi.org/10.3368/jhr.monopsony.0119-10007R1. [40]

Bernard, A. et al. (2012), "The Empirics of Firm Heterogeneity and International Trade", *Annual Review of Economics*, Vol. 4/1, pp. 283-313, https://doi.org/10.1146/annurev-economics-080511-110928. [47]

Berry, S., M. Gaynor and F. Morton (2019), *Do increasing markups matter? Lessons from empirical industrial organization*, American Economic Association, https://doi.org/10.1257/jep.33.3.44. [22]

Bighelli, T. et al. (2023), "European Firm Concentration and Aggregate Productivity", *Journal of the European Economic Association*, Vol. 21/2, pp. 455-483, https://doi.org/10.1093/jeea/jvac040. [7]

Bloom, N., M. Schankerman and J. Van Reenen (2013), "Identifying technology spillovers and product market rivalry", *Econometrica*, Vol. 81/4, pp. 1347-1393. [49]

Buccirossi, P. et al. (2013), "Competition Policy and Productivity Growth: An Empirical Assessment", *The Review of Economics and Statistics*, Vol. 95/4, pp. 1324-1336. [4]

Calligaris, S. et al. (2024), "Measuring competition through different proxies: some OECD evidence", *Protecting competition in a changing world, EC report*, https://competition-policy.ec.europa.eu/about/reaching-out/protecting-competition-changing-world_en (accessed on 5 August 2024). [34]

Calligaris, S. et al. (2024), "Defining the Geographical Level of Competition: A Taxonomy of Industries", *OECD Science, Technology and Industry Working Papers, No. 2024/05*. [17]

Canada Competition Bureau (2023), *Competition in Canada from 2000 to 2020: An Economy at a Crossroads*, Competition Bureau, Gatineau QC. [27]

Competition and Markets Authority (2022), *The State of UK Competition*, http://www.nationalarchives.gov.uk/doc/open-government-. [35]

Covarrubias, M., G. Gutierrez and T. Philippon (2019), "From Good to Bad Concentration? US Industries over the Past 30 Years", in *2019 NBER Macroeconomic Annual*, NBER. [10]

Cravino, J. and A. Levchenko (2017), "Multinational Firms and International Business Cycle Transmission*", *The Quarterly Journal of Economics*, Vol. 132/2, pp. 921-962, https://doi.org/10.1093/qje/qjw043. [50]

De Loecker, J., T. Obermeier and J. Van Reenen (2022), *Firms and inequality*. [30]

De Ridder, M. (2024), "Market Power and Innovation in the Intangible Economy", *American Economic Review*, Vol. 114/1, pp. 199-251, https://doi.org/10.1257/aer.20201079. [29]

Demsetz, H. (1973), "Industry Structure, Market Rivalry, and Public Policy", *The Journal of Law and Economics*, Vol. 16/1, pp. 1-9, https://doi.org/10.1086/466752. [20]

European Commission (1997), "Commission Notice on the definition of relevant market for the purposes of Community competition law.", *Official Journal of the European Communities*, pp. 0005-0013. [53]

European Commission, D. (2024), "Protecting competition in a changing world – Evidence on the evolution of competition in the EU during the past 25 years", https://data.europa.eu/doi/10.2763/089949. [1]

Fons-Rosen, C. et al. (2021), "Quantifying productivity gains from foreign investment", *Journal of International Economics*, Vol. 131, p. 103456, https://doi.org/10.1016/j.jinteco.2021.103456. [51]

Freund, C. and D. Sidhu (2017), "Global Competition and the Rise of China", No. 17-3, Peterson Institute for International Economics. [54]

Furman, J. and P. Orszag (2015), "A Firm-Level Perspective on the Role of Rents in the Rise in Inequality for us". [11]

Gal, P. (2013), "Measuring total factor productivity at the firm level using OECD-ORBIS". [44]

Ganapati, S. (2021), "Growing Oligopolies, Prices, Output, and Productivity", *American Economic Journal: Microeconomics*, Vol. 13/3, pp. 309-327, https://doi.org/10.1257/mic.20190029. [26]

Grullon, G., Y. Larkin and R. Michaely (2019), "Are US industries becoming more concentrated?", *Review of Finance*, Vol. 23/4, pp. 697-743, https://doi.org/10.1093/ROF/RFZ007. [24]

Gutierrez, G. and T. Philippon (2023), "How European markets became free: a study of institutional drift", *Journal of the European Economic Association*, https://doi.org/10.1093/jeea/jvac071. [14]

Hsieh, C. and E. Rossi-Hansberg (2023), "The Industrial Revolution in Services", *Journal of Political Economy Macroeconomics*, Vol. 1/1, pp. 3-42, https://doi.org/10.1086/723009. [37]

IMF (2019), "The Rise of Corporate Market Power and Its Macroeconomic Effects", in *World Economic Outlook, April 2019: Growth Slowdown, Precarious Recovery*. [32]

Kalemli-Ozcan¸ Sebnem, B. et al. (2023), "How to Construct Nationally Representative Firm Level Data from the Orbis Global Database: New Facts on [15]

SMEs and Aggregate Implications for Industry Concentration", *American Economic Journal: Macroeconomics*.

Koltay, G., S. Lorincz and T. Valletti (2023), "Concentration and Competition: Evidence From Europe and Implications For Policy", *Journal of Competition Law & Economics*, https://doi.org/10.1093/joclec/nhad012. [31]

Kwon, S., Y. Ma and K. Zimmermann (2023), "100 Years of Rising Corporate Concentration", *SSRN Electronic Journal*, https://doi.org/10.2139/ssrn.4362319. [25]

Lashkari, D., A. Bauer and J. Boussard (2019), "Information Technology and Returns to Scale", Banque de France. [28]

Lyons, B., C. Matraves and P. Moffatt (2001), "Industrial Concentration and Market Integration in the European Union", *Economica*, Vol. 68/269, pp. 1-26, https://doi.org/10.1111/1468-0335.00230. [42]

Melitz, M. (2003), "The Impact of Trade on Intra-Industry Reallocations and Aggregate Industry Productivity", *Econometrica*, Vol. 71/6, pp. 1695-1725, http://ideas.repec.org/a/ecm/emetrp/v71y2003i6p1695-1725.html. [19]

OECD (2021), "Methodologies to Measure Market Competition", *OECD Competition Committee Issues Paper, OECD Publishing, Paris.*. [18]

OECD (2018), "Market Concentration", *OECD Competition Committee Issues Paper, OECD Publishing, Paris.*. [16]

Rinz, K. (2022), "Labor Market Concentration, Earnings, and Inequality", *Journal of Human Resources*, Vol. 57/Special Issue 1, pp. S251-S283, https://doi.org/10.3368/jhr.monopsony.0219-10025R1. [38]

Rossi-Hansberg, E., P. Sarte and N. Trachter (2021), "Diverging Trends in National and Local Concentration", *NBER Macroeconomics Annual*, Vol. 35, pp. 115-150, https://doi.org/10.1086/712317. [36]

Schmalensee, R. (1989), "Inter-industry studies of structure and performance", *Handbook of Industrial Organisation*, Vol. 2, pp. 951-1009. [21]

Shapiro, C. (2018), "Antitrust in a time of populism", *International Journal of Industrial Organization*, Vol. 61, pp. 714-748, https://doi.org/10.1016/j.ijindorg.2018.01.001. [45]

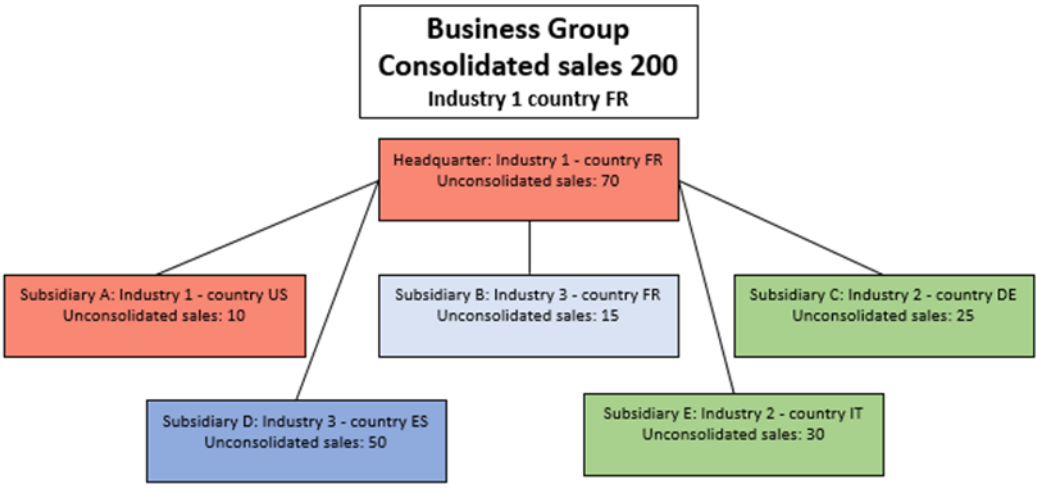
Smith, D. and S. Ocampo (2022), "The Evolution of U.S. Retail Concentration", Arxive, http://arxiv.org/abs/2202.07609. [41]

Syverson, C. (2019), "Macroeconomics and Market Power: Context, Implications, and Open Questions", *Journal of Economic Perspectives*, Vol. 33/3, pp. 23-43, https://doi.org/10.1257/jep.33.3.23. [23]

Werden, G. and L. Froeb (2018), "Don't Panic: A Guide to Claims of Increasing Concentration", *SSRN Electronic Journal*, https://doi.org/10.2139/ssrn.3156912. [46]

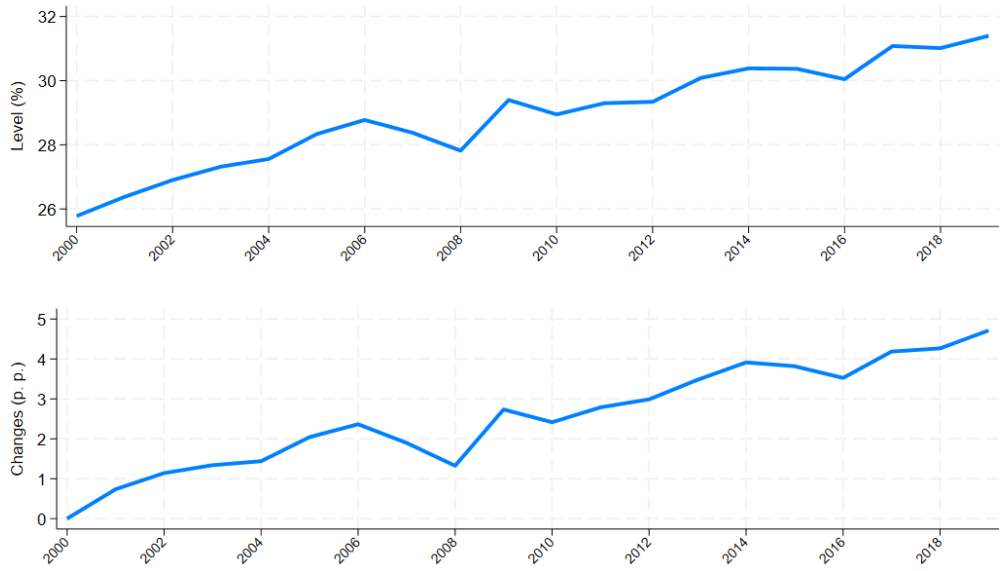
Figures and Tables

Figure 1. Example of the apportioning technique of business group activities



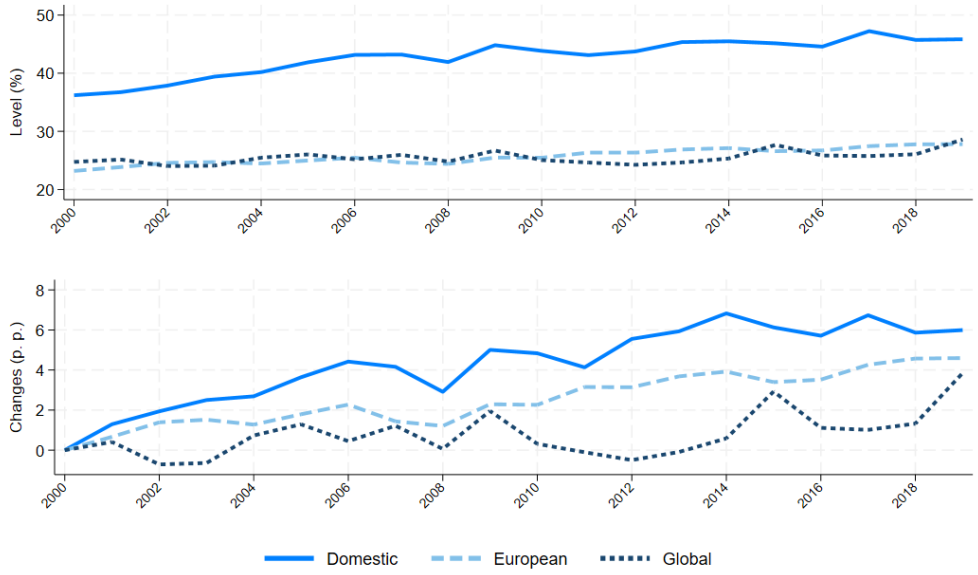
Note: The figure depicts an example of a hypothetical group consisting of a parent company from France and operating: i) in industry 1 (global), with a US subsidiary in the same industry; ii) in industry 2 (European) with two subsidiaries, one from Germany and one from Italy; and finally, iii) in industry 3 (domestic) with a subsidiary in France and one in Spain. The different colours identify the four different markets in which the group is active: the global market in industry 1 (with total gross output of 80), the European market in industry 2 (with total gross output of 55), the French market in industry 3 (with total gross output of 15), and the Spanish market in industry 3 (with total gross output of 50).

Figure 2. Concentration aggregating across geographical buckets



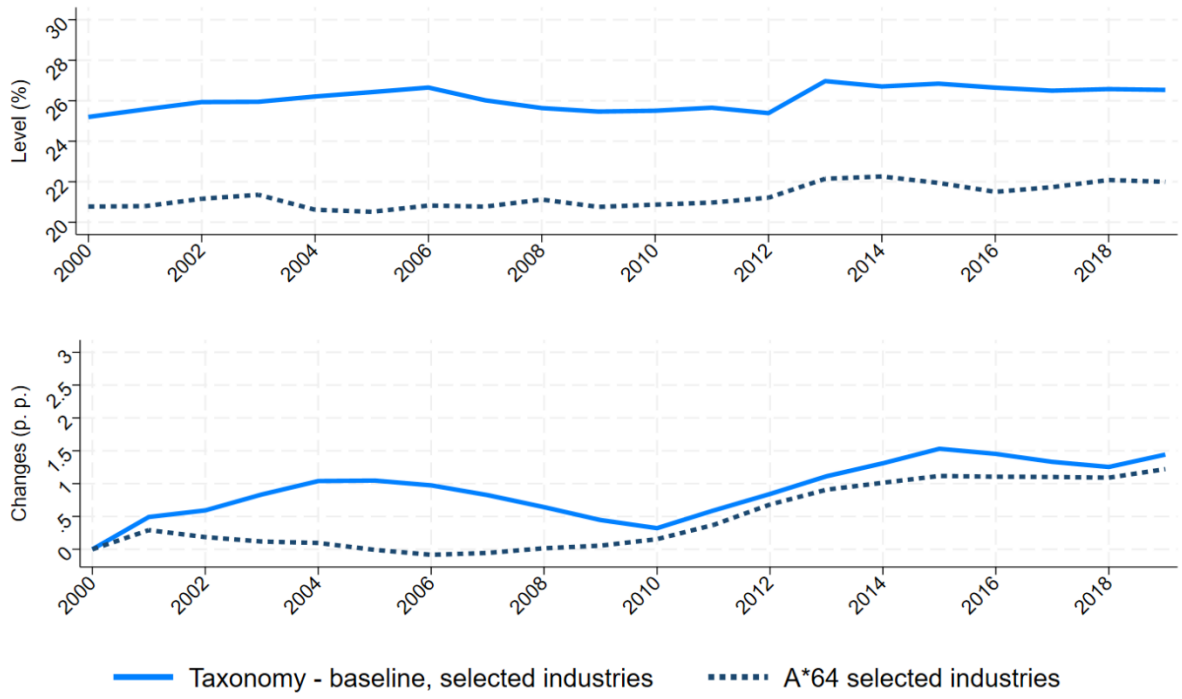
Note: The chart shows the weighted average across industry-geography combinations of CR4 levels (top panel) and cumulative growth (bottom panel). A weight of 1 is attributed to country-industries cells for domestic industries, while a weight of 15 is attributed to industries belonging to the European and the global buckets. Industries included in the analysis are a mix of 2 and 3-digit industries belonging to mining, manufacturing, utilities, and non-financial market services and are classified as either domestic, European, or global, depending on the taxonomy developed by Calligaris et al. (2024[16]). The countries included in the sample are BEL, DEU, DNK, ESP, FIN, FRA, GBR, GRC, HUN, ITA, NOR, POL, PRT, SVN, and SWE for the domestic and European buckets, while in the global one also JPN, KOR, and USA are included.
Source: Authors' calculations.

Figure 3. Concentration across geographical buckets



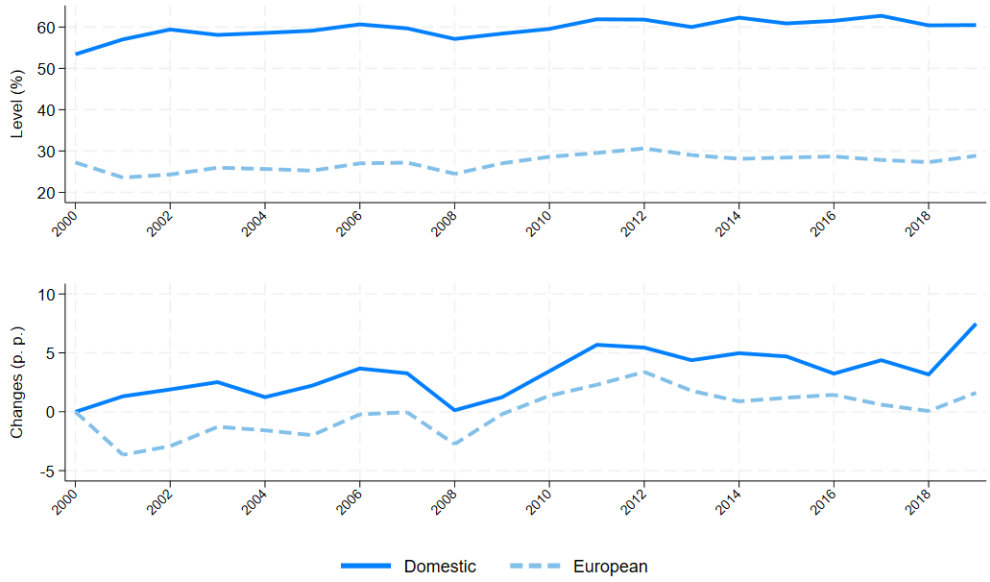
Note: The chart shows the unweighted average across industries (and countries, for the domestic bucket) of CR4 levels (top panel) and cumulative growth (bottom panel). Industries included in the analysis are mix of 2 and 3-digit industries belonging to mining, manufacturing, utilities, and non-financial market services and are classified as either domestic, European, or global, depending on the taxonomy developed by Calligaris et al. (2024[16]). The countries included in the sample are BEL, DEU, DNK, ESP, FIN, FRA, GBR, GRC, HUN, ITA, NOR, POL, PRT, SVN, and SWE for the domestic and European bucket, while in the global one also JPN, KOR, and USA are included.
Source: Authors' calculations.

Figure 4. Concentration with different industry aggregation levels (selected industries)



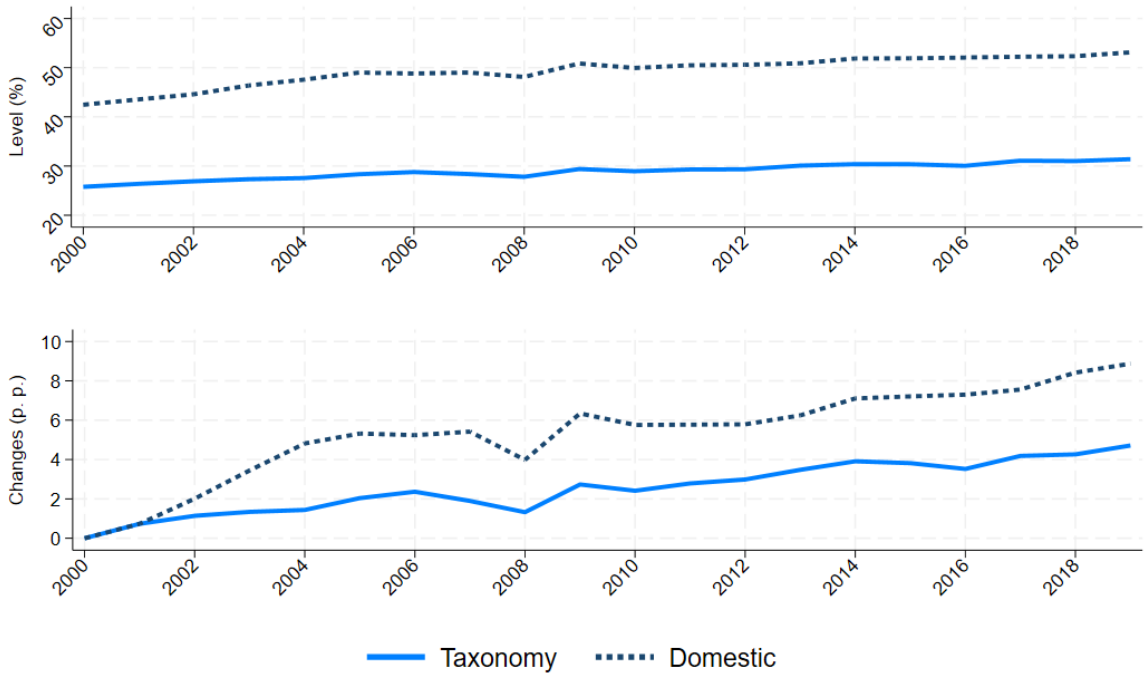
Note: The chart shows the average across industry-geography combinations of CR4 levels (top panel) and cumulative growth (bottom panel) in a selection of industries when using the industry aggregation of the baseline measure (continuous blue line, mostly 3-digit) and when using a higher-level industry aggregation (dotted dark line, mostly 2-digit). A weight of 1 is attributed to country-industries cells for domestic industries, while a weight of 15 is attributed to industries belonging to the European and the global buckets. Industries included in the analysis include all the 2-digit industries (following the STAN A*64 classification) whose 3-digit sub-industries belong to a unique geographical bucket. The 2-digit industries are, for the domestic bucket: C33, D35, D36, D37, H55, J59, and J61; for the European bucket: C16, C17, C19, C20, C22, C24, C27, C28, C29, G45, G46, G47H25, H53, J62, M71, M72, M73, N77, N78, and N79. The countries included in the sample are BEL, DEU, DNK, ESP, FIN, FRA, GBR, GRC, HUN, ITA, NOR, POL, PRT, SVN, and SWE.
Source: Authors' calculations.

Figure 5. Concentration for selected 4-digit industries



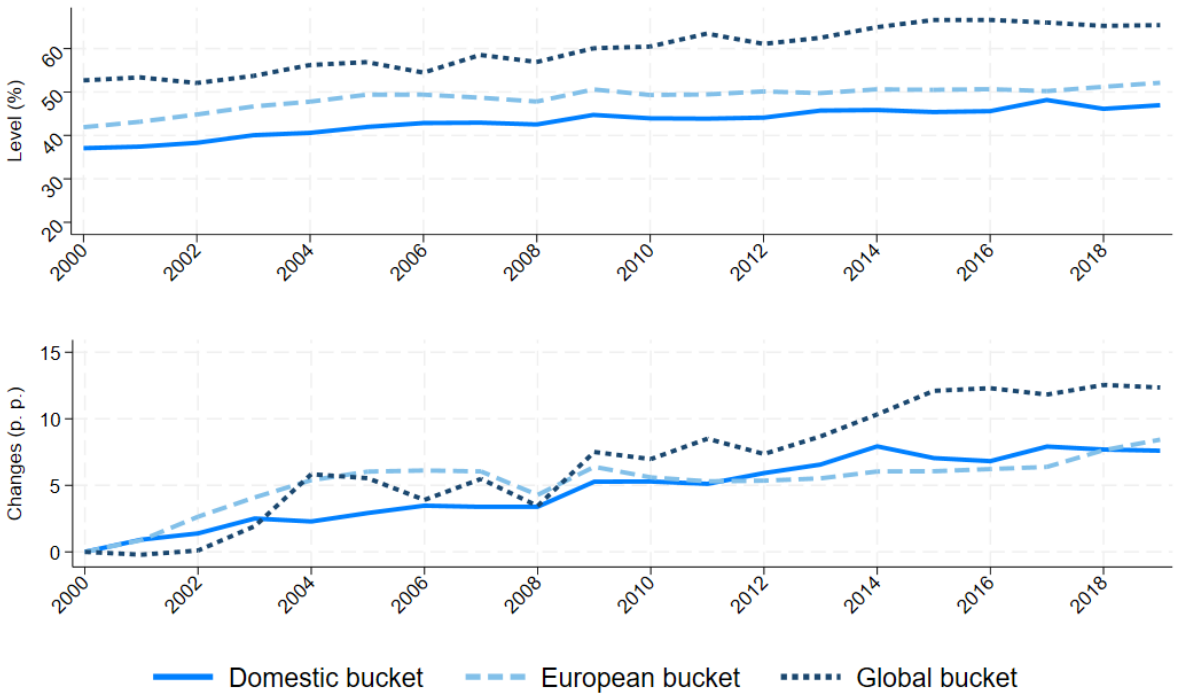
Note: The chart shows the unweighted average across industries (and countries, for the domestic bucket) of CR4 levels (top panel) and cumulative growth (bottom panel). Industries included in the analysis are all the 4-digit industries (except for 1042 and 1089 due to data limitations) belonging to the 2-digit industries from C10 (manufacture of food products) and C11 (manufacture of beverages). The countries included in the sample are BEL, DEU, DNK, ESP, FIN, FRA, GBR, GRC, HUN, ITA, NOR, POL, PRT, SVN, and SWE.
Source: Authors' calculations.

Figure 6. National concentration aggregating over different geographical regions



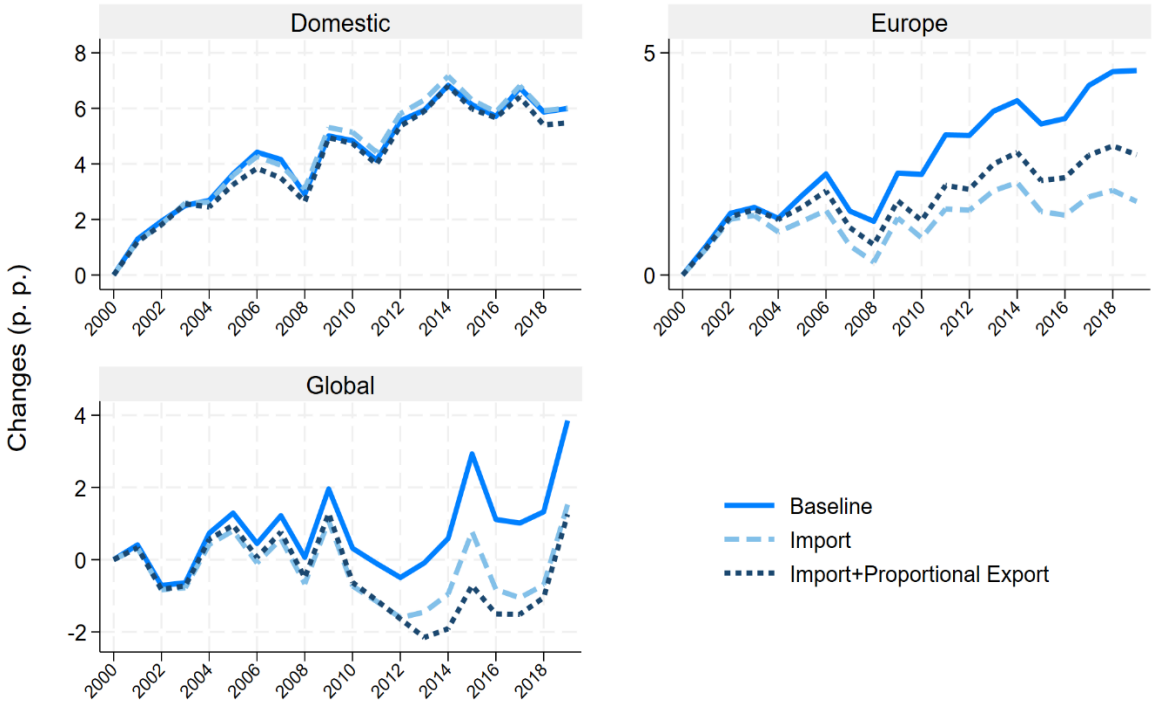
Note: The chart shows the average across industry-geography combinations of CR4 levels (top panel) and cumulative growth (bottom panel) when the taxonomy is adopted (continuous blue line) and with all industries considered as domestic ones (dotted dark line). When the taxonomy is used, a weight of 1 is attributed to country-industries cells for domestic industries, while a weight of 15 is attributed to industries belonging to the European and the global buckets. Industries included in the analysis are mix of 2 and 3-digit industries belonging to mining, manufacturing, utilities, and non-financial market services. The countries included in the sample are BEL, DEU, DNK, ESP, FIN, FRA, GBR, GRC, HUN, ITA, NOR, POL, PRT, SVN, and SWE for the domestic and European bucket, while in the global one also JPN, KOR, and USA are included.
Source: Authors' calculations.

Figure 7. National concentration across geographical buckets (all industries as domestic)



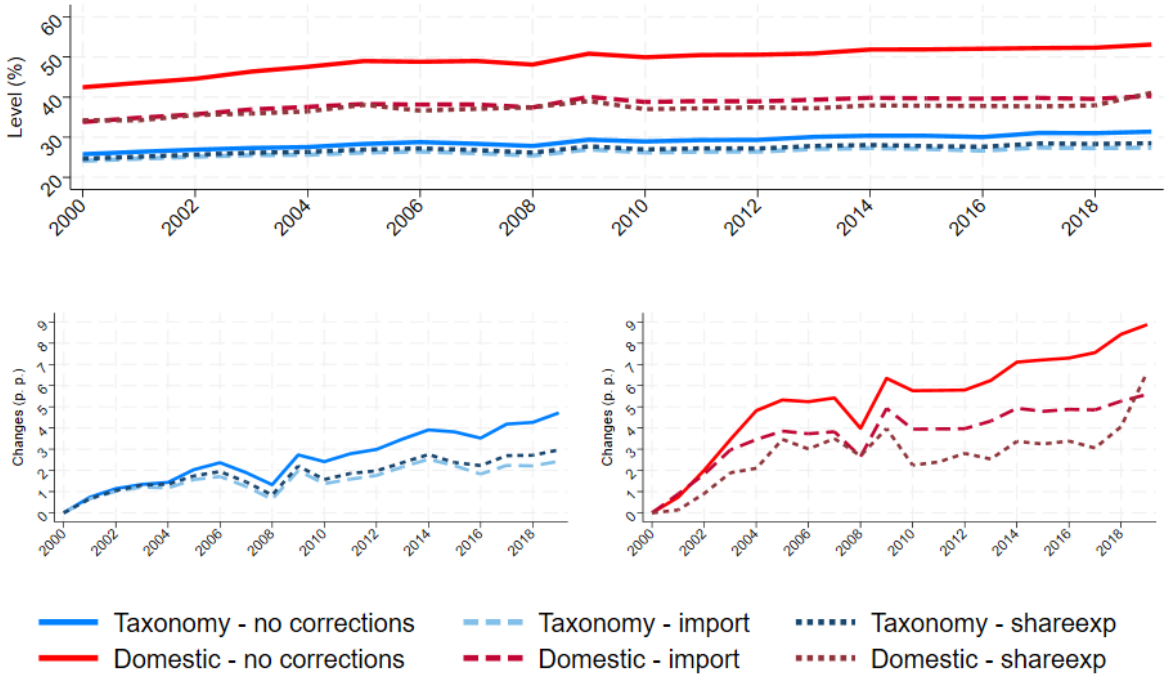
Note: The chart shows the unweighted average across industries and countries and within geographical buckets of CR4 levels (top panel) and cumulative growth (bottom panel), computed treating all industries as if they were competing at the domestic level. Industries included in the analysis are mix of 2 and 3-digit industries belonging to mining, manufacturing, utilities, and non-financial market services. The countries included in the sample are BEL, DEU, DNK, ESP, FIN, FRA, GBR, GRC, HUN, ITA, NOR, POL, PRT, SVN, and SWE for the domestic and European bucket, while in the global one also JPN, KOR, and USA are included.
Source: Authors' calculations.

Figure 8. Different trade adjustments



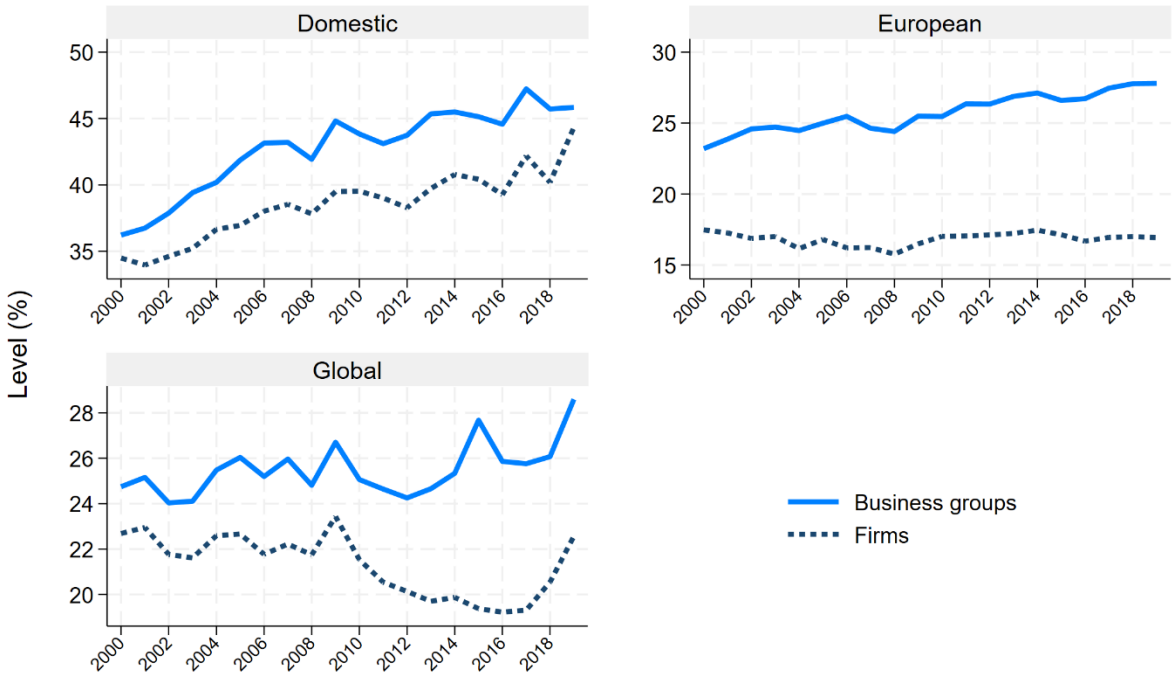
Note: The chart shows the unweighted average across industries and countries of cumulative change in CR4 across geographical buckets for different types of trade adjustments. The solid blue lines refer to the baseline average cumulative change (no corrections). The light blue dashed ones to the correction obtained by adding import at the denominator. The dark blue dotted ones to the correction obtained by adding import and subtracting total exports at the denominator, and by subtracting at the numerator a fraction of export given by the share of gross output accounted for by the four largest firms in the market. Industries are a mix of 2 and 3-digit industries belonging to mining, manufacturing, utilities, and non-financial market. The countries included in the sample are BEL, DEU, DNK, ESP, FIN, FRA, GBR, GRC, HUN, ITA, NOR, POL, PRT, SVN, and SWE. Industries competing at the global level also include JPN, KOR, and USA.
Source: Authors' calculations.

Figure 9. Concentration level different trade corrections (taxonomy vs. all as domestic)



Note: the chart shows the unweighted average across industries and countries of CR4 levels for different types of trade adjustments. Blue lines refer to concentration computed using the taxonomy developed by Calligaris et al. (2024[16]), red lines to concentration computed with all industries considered as domestic. Solid lines refer to the baseline average level (no corrections). Dashed ones to the correction obtained by adding import at the denominator. Dotted ones to the correction obtained by adding import and subtracting total exports at the denominator, and by subtracting at the numerator a fraction of export given by the share of gross output accounted for by the four largest firms in the market. Industries are a mix of 2 and 3-digit industries belonging to mining, manufacturing, utilities, and non-financial market. The countries included in the sample are BEL, DEU, DNK, ESP, FIN, FRA, GBR, GRC, HUN, ITA, NOR, POL, PRT, SVN, and SWE (plus JPN, KOR, and USA for industries belonging to the global bucket).
Source: Authors' calculations.

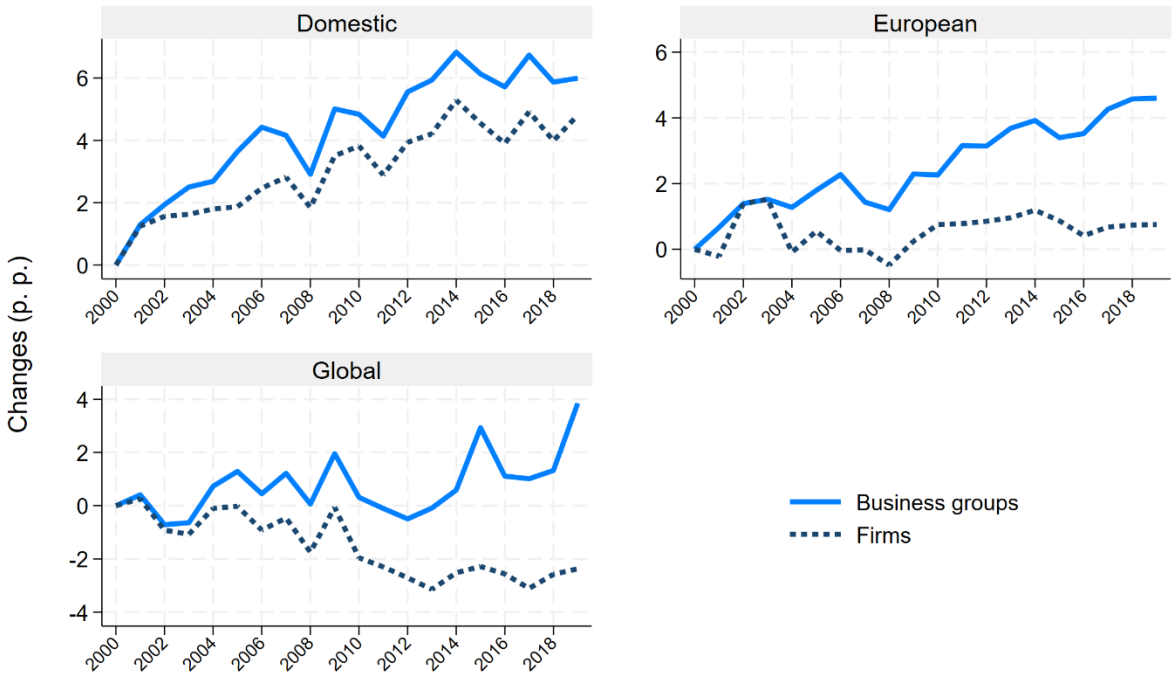
Figure 10. Concentration levels, firms vs. business groups



Note: The chart shows the unweighted average across industries and countries of CR4 levels across geographical buckets. The solid blue lines refer to the baseline average cumulative change. The dashed ones represent concentration when the ownership structure of business groups is neglected. Industries are a mix of 2 and 3-digit industries belonging to mining, manufacturing, utilities, and non-financial market. The countries included in the sample are BEL, DEU, DNK, ESP, FIN, FRA, GBR, GRC, HUN, ITA, NOR, POL, PRT, SVN, and SWE. Industries competing at the global level also include JPN, KOR, and USA.

Source: Authors' calculations.

Figure 11. Concentration cumulative changes, firms vs. business groups



Note: The chart shows the unweighted average across industries and countries of cumulative change in CR4 across geographical buckets. The solid blue lines refer to the baseline average cumulative change. The dashed ones represent concentration when the ownership structure of business groups is neglected. Industries are a mix of 2 and 3-digit industries belonging to mining, manufacturing, utilities, and non-financial market. The countries included in the sample are BEL, DEU, DNK, ESP, FIN, FRA, GBR, GRC, HUN, ITA, NOR, POL, PRT, SVN, and SWE. Industries competing at the global level also include JPN, KOR, and USA.
Source: Authors' calculations.

Annexe A. Data Appendix

Sample of industries

The final sample used for studying the concentration measure spans 20 years (2000-2019) and is composed by 127 distinct industries allocated to the three different geographical buckets (27 are domestic, 80 are European, and 20 are global). Out of these, 112 (88%) are 3-digit, 10 (8%) are 2-digit, and 5 (4%) are aggregation of two or more 2-digit. The difference from the number of industries included in the production sample is due essentially to combined data limitation either in Orbis at the firm level, in gross in gross output at the industry level or in the trade data. In particular, a number of sectors that might be relevant from a competition perspective - either because structurally oriented towards high concentration levels or because historically under the lens of competition authorities – had to be excluded due to severe data limitations both at the numerator and at the denominator of the concentration measure. Table A 1 reports the list of industries used in this paper and their associated geographic market.

Firm-level financial data

This section provides a summary of the cleaning procedures applied to the Orbis dataset used in the report. Orbis data have been used to identify the top firms in a market in order to compute concentration. While the initial database is the same, some different cleaning procedures have been adopted for the two purposes. Both are described in this Annex.

As explained in Section 2, financial information within Orbis is available both at the business group level (consolidated financial data aggregated across subsidiaries belonging to the same owner) and at individual firm level (unconsolidated information referring to an individual firm). In this work, unconsolidated accounts are used in the analysis.[61] Consolidated accounts are used only in two cases

[61] As explained in Section 3, the measure of concentration built in this report look at business group activities rather than at single firms. The approach adopted in this paper fundamentally relies upon unconsolidated data of the individual subsidiaries within a business group, to identify the precise industry and location of all the subsidiaries belonging to a group, and correctly apportion the group sales to the markets in which the business group is active.

in which they can be considered equivalent to unconsolidated accounts: for independent firms (i.e., firms that are not part of a business group), and for subsidiaries at the bottom of the ownership hierarchy (subsidiaries not owning further subsidiaries) that do not have unconsolidated accounts, because for such firms consolidated and unconsolidated accounts coincide.

In the following two sub-sections, additional details on the data cleaning preparation for the two different samples are provided.

Sample for concentration

The methodology outlined in Section 3, computing concentration requires good coverage of both business group and individual subsidiaries' financial information. To ensure that all economic activity of each group's subsidiaries is captured, information for firms of all sizes and in all sectors are used.[62]

Following Bajgar et al. (2023[6]), some steps are taken to improve the coverage of the data (see Bajgar et al. (2019[43]) for a discussion on the impact of these steps on the sample). First, the coverage of Orbis has been expanded by using available consolidated information to infer missing years in the unconsolidated information of the same firm and vice-versa. Second, to increase the coverage of consolidated accounts of listed firms, Worldscope database has been used. Worldscope is merged to Orbis through firms' International Securities Identification Number (ISIN) numbers, which uniquely identify listed firms. For some countries, such as the United States, Worldscope can improve the coverage of Orbis substantially. The same cleaning rules used for Orbis data have been applied to Worldscope data. Worldscope reports consolidated financial data and contains very similar values to Orbis consolidated data for observations present in both datasets.[63]

As discussed in Section 3, the methodology developed to aggregate sales across all subsidiaries operating in a given market only uses unconsolidated sales of each firm. Consolidated accounts are used only to correct the unconsolidated information in cases where the total subsidiary sales exceed

[62] Note that this sample is used at an initial stage of the data construction, in the attempt to consider the worldwide sales of business groups across all sectors in which they are active. This is important because it allows to have similar numbers when comparing the consolidated accounts of the headquarters and the sum of the unconsolidated sales of all their subsidiaries. Subsequently, as explained in Section 2, due mainly to data coverage and comparability across countries, the sample of countries and industries is restricted to those specified in the report.

[63] See Bajgar et al. (2019[43]) for a discussion on the comparability between the two sources for firms present in both databases.

group sales (presumably due to inter-company transactions) or where unconsolidated data are missing. In the latter case, if a headquarter company reports always consolidated accounts but unconsolidated accounts only in some years, the missing years in the unconsolidated accounts are interpolated using growth rates of the consolidated accounts and assuming a constant share of unconsolidated accounts relative to consolidated accounts.

Ownership data

As explained in Section 3, the business group structure is used to apportion the overall sales of the group across all the relevant market where it is active. To do so requires detailed ownership information on parent-subsidiary linkages. The primary source of firm ownership information is Orbis, which is supplemented with data from the Orbis M&A database of Mergers and Acquisitions. Importantly, both datasets are provided by Moody's and share a common firm identifier which allows to merge the two datasets.

Orbis contains comprehensive information on ownership linkages among firms, extensively used in the existing literature (Cravino and Levchenko (2017[50]), Fons-Rosen (2021[51])), which allow to detail ownership linkages between shareholders and their subsidiaries, as well as the identity of the global ultimate owner of subsidiaries (calculated at each calendar year from 2007 until 2020). The global ultimate owner is defined as the firm owning at least 50.01% of total shares of a subsidiary. This is a commonly used threshold for the definition of control of another firm and, thus, to understand whether the subsidiary's financial information is consolidated into the parent accounts.

To calculate ultimate owners, Orbis uses the tree of ownership linkages for each firm and year. They identify each for each firm its shareholder (the immediate owner), then the shareholder's shareholders and so on. So, for each firm, they start at the bottom and work up the tree of ownership linkages until they find a shareholder that is independent (not controlled by anyone) or controlled by an individual. That shareholder is classified as the ultimate owner of the subsidiary firm at the bottom of the tree.

However, in Orbis the data primarily start in 2007 and later for some firms. Thus, the main data source is complemented with the Orbis M&A database to measure earlier changes in ownership, enabling the construction of a series starting as early as 2000 whenever data allow, as discussed in the

following sub-sections.[64] The Orbis M&A database contains deal-level information on M&As from 1997 onwards for European firms, from 2000 onwards for North American firms, and for other geographic regions from 2003. Overall, Orbis M&A contains about 2 million M&A deals from 2000 to 2020.

In the following sub-sections, a summary of the methodology is provided. For further details and a more complete discussion, please refer to Bajgar et al. (2019[43]), (2023[6]).

Identifying business groups

The procedure to clean and harmonise Orbis and Orbis M&A relies on the work of (Bajgar et al. (2023[6])). Following their approach, several steps are undertaken to expand the coverage of the ultimate owner from Orbis. The first step is to use Orbis M&A to identify changes in immediate (rather than global ultimate) owners not available from Orbis M&A. For each deal, Orbis M&A contains information on the target, acquiror and vendor firms. About 700,000 deals represent either changes in majority ownership – such as a firm increasing from 10% to 51% equity ownership – or a majority owner further increasing its stake – such as a firm increasing from 51% to 60% ownership. Both types of deal allow to identify the immediate owner of each target firm at the time of the deal. Furthermore, for changes in majority ownership – when the target firm switches hands – the vendor firm represents the previous immediate owner.

A second step is to use the information available from the table "current" Orbis linkages, which provides direct and indirect ownership linkages. These are used to retrieve the identity of the ultimate owner in cases where the latter is missing but it is possible to identify a shareholder with indirect share higher that 50.01%.

The third step is to translate the changes in immediate owners (from the first two steps above) to changes in the ultimate owner. The immediate owner who acquired the target firm may not be the ultimate owner. To find the ultimate owner, the same procedure used by Orbis is followed. Orbis M&A

[64] Whilst ultimate ownership data starts in 2007, for some firms it is not available until later years. Common approaches to correcting for this in the literature are either to assume that firms without an Orbis ultimate owner are independent or to take data from a recent year - assuming ownership has not changed over time. Both approaches are problematic. With increasing coverage of ownership over time in Orbis, the former approach will falsely equate missing data with independence and lead to an overstatement of ownership changes over time. The latter approach will lead to an understatement of ownership changes over time and will typically overstate the number of markets and countries in which a firm operates.

immediate owner and available information on ownership linkages are combined to find the shareholders of the immediate owners, and the shareholders of their shareholders, and so on. The 50.01% criterium is used until the procedure arrives to a shareholder that is either independent or controlled by an individual. This final shareholder is deemed the ultimate owner.

The fourth step is to impute missing years of ownership information and information and roll the owner backwards and forwards until there was an M&A or change in ownership (from the steps above). The additional information on ownership changes allows to roll the ownership information forwards and/or backwards until there was a change in owner, rather than simply assuming that a missing ultimate owner implies independence between firms. For example, if firm C is the ultimate owner of firm A in 2010, and from Orbis M&A data it is know that firm A was acquired in 2008, then the ultimate ownership information is rolled backwards until 2008. Moreover, in about half of the acquisition cases in the M&A sample, it is also known that firm A was acquired from vendor firm B in 2008, so it is possible to infer that firm B was the (immediate) owner of firm A and roll back further until an earlier M&A transaction.

Data cleaning

Numerous steps are undertaken to identify and correct potential issues in the ownership data, specially to identify missing linkages among the largest firms. Spot checking revealed that some large firm groups are missing ownership linkages between the parent firm and their subsidiaries for some years. This can be problematic because it can lead to a double counting of group activity, with both the parent's consolidated financials and their subsidiary information included as separate groups.

Accordingly, the following checks are undertaken to mitigate this risk. First, ultimate owners that are themselves majority owned by another firm cannot be true ultimate owners and are therefore adjusted in the data. Second, temporary (one or two year) deviations in ultimate owner relationships, whereby a firm's ultimate owner changes for just one or two years and then reverts to its previous owner, are removed, as this is an unusual phenomenon in ownership and is most likely to be measurement error. These two steps affect approximately 10,000 firms per year.

Third, to detect missing linkages, large firms that change from having no subsidiaries to a large number of subsidiaries from one year to the next are examined and manually updated, if necessary. Spot checking revealed cases of intermediate holding companies (that often have no financial information) being temporarily incorrectly identified as the ultimate owner. To address this issue, large

groups of subsidiaries (in terms of sales) that have a parent with no financials but switch to a new parent in the following period that does have financials are examined manually. Cases of M&As identified by Orbis M&A have been excluded, and only cases where more than 90% of subsidiaries transfer to the new parent have been considered. The 1500 largest groups identified, corresponding to groups with sales larger than 10 million of Euro, have been corrected in the following way. For the 150 largest groups, each group has been manually inspected against their financial statements, while for other groups a name-matching algorithm has been used to semi-automate the identification of whether the prior owner was in fact a holding company of the new parent. Those with very similar names have been considered as part of the same group, correcting 147 groups.

Fourth, large firms that never have any subsidiary and, vice-versa, large groups of subsidiaries that never have a parent with financials are examined to identify missing links. This builds on the previous step, identifying large groups of subsidiaries that never have a parent with financials, and large parents that never have subsidiaries. In total, 1,031 parents with sales of more than 1 billion euros that never have subsidiaries have been found, and 251 groups of subsidiaries with more than 1 billion euros of sales that never have a parent with financials. Again, a name-matching algorithm has been used to semi-automate the identification of whether the prior owner was in fact a holding company of the new parent. This process applies to cases of groups (large groups of subsidiaries or large parents) with sales larger than 40 million and treats those with very similar names part of the same group. In total, based on visual inspection of the name-matching string-matching similarity, 287 groups per year have been corrected.

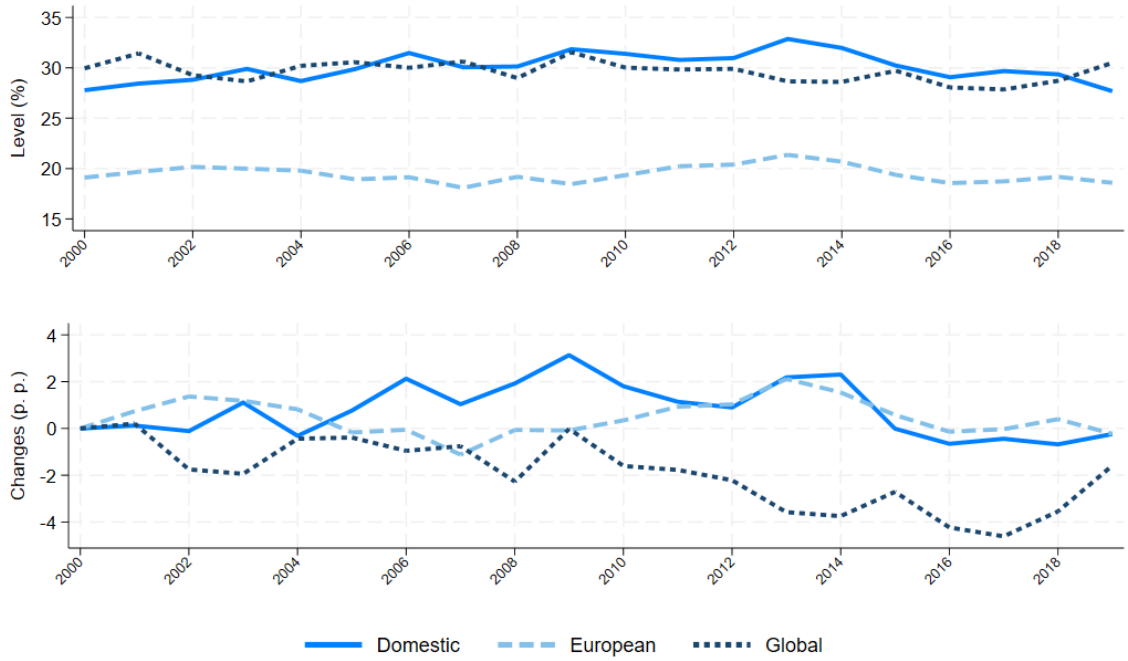
Fifth, missing links where there are ownership changes among firms with very similar names – and are so very likely part of the same group (e.g., ABC Motors acquired by ABC Motors Thailand) – are identified and corrected. This considers any ownership change where the owners have a similar root to their name (e.g., “XYZ Inc” and “XYZ Plc”). These remaining firms are not large, or do not have completely missing subsidiaries; if they had, they would have been encompassed in the earlier cleaning steps. These firms are therefore somewhat less problematic for the resulting concentration metrics. Given this reduced risk and the fact that all firms in the data are considered as part of this step, an automated check using name-matching is carried out, which requires an identical match of the cleaned name. Common company type abbreviations (e.g., Plc, Ltd, SA, Gmbh etc.), country names (e.g., ABC (Viet Nam) Ltd) and punctuation are removed, and the resulting root of the name is required to be identical. The global ultimate owner is modified only when the ownership change involves two

companies with almost exact name and the ownership change happens between one ultimate owner that has financials and the other one that has not financials. In total, approximately 5,000 cases are corrected.

Finally, a further check for groups with total gross output (considered as the sum of the unconsolidated accounts of all its subsidiaries) in a given country, industry, and year larger than 150 million of Euro is conducted. Within this subset of business groups, firms with similar names in the same country and industry are checked using a string-matching algorithm to select relevant cases. As this algorithm captures situations where the ownership tree is partially missing some links, the spotted cases might be particularly relevant for concentration measures. For example, the company ACCO brands Europe was reported as a GUO, while it is part of the group ACCO Brands Corporation. Both were active in the same country-industry and year, therefore leading to a downward biased measure of concentration. In this situation, there are two groups (or simple firms), both with large revenues, but that do not have complete ownership structure. The subsample of GUOs with similar names and active in the same year has been manual inspected in order to understand when the GUO was indeed the same, correcting approximately 300 cases.

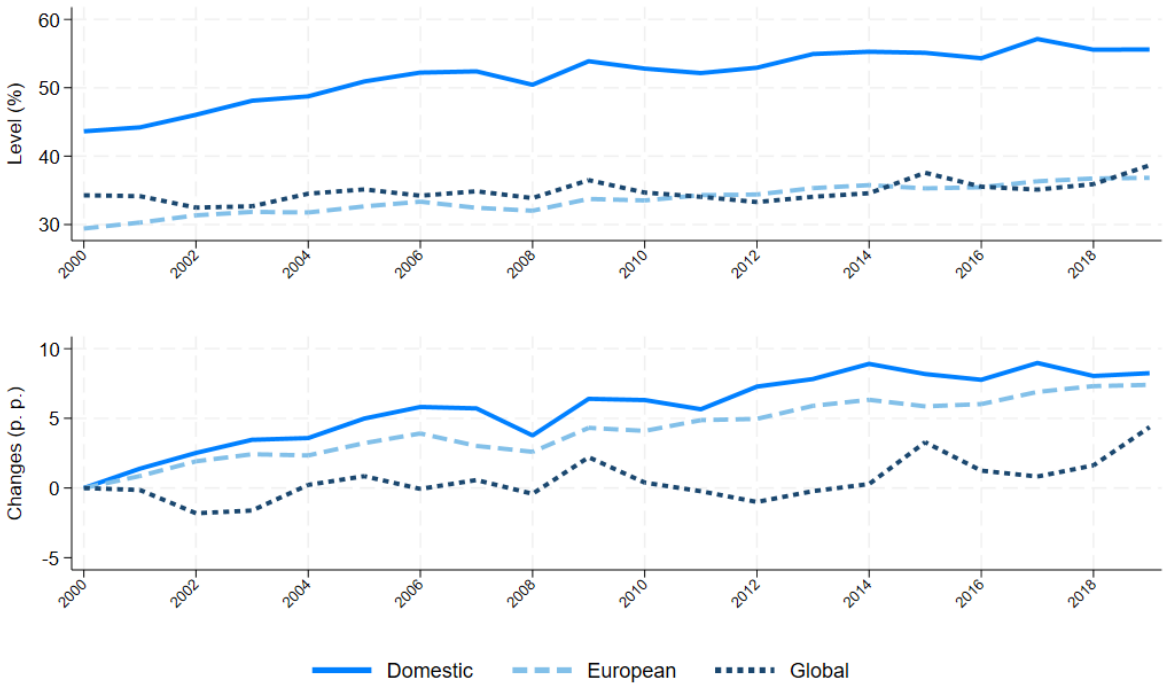
Annex Figures and Tables

Figure A 1.Concentration across geographical buckets (weighted)



Note: The chart shows the weighted average across industries (and countries, for the domestic bucket) of CR4 levels (top panel) and cumulative growth (bottom panel). Weights are given by market size, captured by gross output. As usual, market is the combination of industry and geographic boundaries determined by the taxonomy. Industries included in the analysis are mix of 2 and 3-digit industries belonging to mining, manufacturing, utilities, and non-financial market services and are classified as either domestic, European, or global, depending on the taxonomy developed by Calligaris et al. (2024[17]). The countries included in the sample are BEL, DEU, DNK, ESP, FIN, FRA, GBR, GRC, HUN, ITA, NOR, POL, PRT, SVN, and SWE for the domestic and European bucket, while in the global one also JPN, KOR, and USA are included.
Source: Authors' calculations.

Figure A 2. Concentration across geographical buckets (CR8)



Note: The chart shows the unweighted average across industries (and countries, for the domestic bucket) of CR8 levels (top panel) and cumulative growth (bottom panel). Weights are given by market size, captured by gross output. As usual, market is the combination of industry and geographic boundaries determined by the taxonomy developed by Calligaris et al. (2024[17]). Industries included in the analysis are mix of 2 and 3-digit industries belonging to mining, manufacturing, utilities, and non-financial market services and are classified as either domestic, European, or global, depending on the taxonomy. The countries included in the sample are BEL, DEU, DNK, ESP, FIN, FRA, GBR, GRC, HUN, ITA, NOR, POL, PRT, SVN, and SWE for the domestic and European bucket, while in the global one also JPN, KOR, and USA are included.
Source: Authors' calculations.

Table A 1.Taxonomy list of industries and geographical dimension

Industry Code	Industry Description	Taxonomy Geography
081	Quarrying of stone, sand and clay	European
089	Mining and quarrying n.e.c.	Global
091	Support activities for petroleum and natural gas extraction	Domestic
099	Support activities for other mining and quarrying	Domestic
101	Processing and preserving of meat and production of meat products	Domestic
102	Processing and preserving of fish, crustaceans and molluscs	European

Industry Code	Industry Description	Taxonomy Geography
103	Processing and preserving of fruit and vegetables	European
104	Manufacture of vegetable and animal oils and fats	European
105	Manufacture of dairy products	Domestic
106	Manufacture of grain mill products, starches and starch products	European
107	Manufacture of bakery and farinaceous products	Domestic
108	Manufacture of other food products	European
109	Manufacture of prepared animal feeds	Domestic
110	Manufacture of beverages	Domestic
131	Preparation and spinning of textile fibres	European
132	Weaving of textiles	European
133	Finishing of textiles	Domestic
139	Manufacture of other textiles	European
14	Manufacture of apparel	Global
151	Tanning and dressing of leather; manufacture of luggage, handbags, saddlery and harness; dressing and dyeing of fur	Global
152	Manufacture of footwear	European
161	Sawmilling and planning of wood	European
162	Manufacture of products of wood, cork, straw and plaiting materials	European
171	Manufacture of pulp, paper and paperboard	European
172	Manufacture of articles of paper and paperboard	European
181	Printing and service activities related to printing	Domestic
182	Reproduction of recorded media	European
19	Manufacture of coke / petroleum	European
201	Manufacture of basic chemicals, fertilisers and nitrogen compounds, plastics and synthetic rubber in primary forms	European
202	Manufacture of pesticides and other agrochemical products	European

Industry Code	Industry Description	Taxonomy Geography
203	Manufacture of paints, varnishes and similar coatings, printing ink and mastics	European
204	Manufacture of soap and detergents, cleaning and polishing preparations, perfumes and toilet preparations	European
205	Manufacture of other chemical products	European
206	Manufacture of man-made fibres	European
211	Manufacture of basic pharmaceutical products	Global
212	Manufacture of pharmaceutical preparations	Global
221	Manufacture of rubber products	European
222	Manufacture of plastic products	European
231	Manufacture of glass and glass products	European
232	Manufacture of refractory products	European
233	Manufacture of clay building materials	European
234	Manufacture of other porcelain and ceramic products	European
235	Manufacture of cement, lime and plaster	European
236	Manufacture of articles of concrete, cement and plaster	Domestic
237	Cutting, shaping and finishing of stone	Global
239	Manufacture of abrasive products and non-metallic mineral products n.e.c.	European
241	Manufacture of basic iron and steel and of ferro-alloys	European
242	Manufacture of tubes, pipes, hollow profiles and related fittings, of steel	European
243	Manufacture of other products of first processing of steel	European
244	Manufacture of basic precious and other non-ferrous metals	European
245	Casting of metals	European
251	Manufacture of structural metal products	European
252	Manufacture of tanks, reservoirs and containers of metal	European

Industry Code	Industry Description	Taxonomy Geography
254	Manufacture of weapons and ammunition	Global
255	Forging, pressing, stamping and roll-forming of metal; powder metallurgy	Domestic
256	Treatment and coating of metals; machining	Domestic
257	Manufacture of cutlery, tools and general hardware	European
259	Manufacture of other fabricated metal products	European
261	Manufacture of electronic components and boards	Global
262	Manufacture of computers and peripheral equipment	European
263	Manufacture of communication equipment	Global
264	Manufacture of consumer electronics	European
265	Manufacture of instruments and appliances for measuring, testing and navigation; watches and clocks	Global
266	Manufacture of irradiation, electromedical and electrotherapeutic equipment	Global
267	Manufacture of optical instruments and photographic equipment	Global
271	Manufacture of electric motors, generators, transformers and electricity distribution and control apparatus	European
272	Manufacture of batteries and accumulators	European
273	Manufacture of wiring and wiring devices	European
274	Manufacture of electric lighting equipment	European
275	Manufacture of domestic appliances	European
279	Manufacture of other electrical equipment	European
281	Manufacture of general-purpose machinery	European
282	Manufacture of other general-purpose machinery	European
283	Manufacture of agricultural and forestry machinery	European
284	Manufacture of metal forming machinery and machine tools	European
289	Manufacture of other special-purpose machinery	European

Industry Code	Industry Description	Taxonomy Geography
291	Manufacture of motor vehicles	European
292	Manufacture of bodies (coachwork) for motor vehicles; manufacture of trailers and semi-trailers	European
293	Manufacture of parts and accessories for motor vehicles	European
301	Building of ships and boats	Global
302	Manufacture of railway locomotives and rolling stock	European
303	Manufacture of air and spacecraft and related machinery	Global
309	Manufacture of transport equipment n.e.c.	European
310	Manufacture of furniture	European
321	Manufacture of jewellery, bijouterie and related articles	Global
322	Manufacture of musical instruments	Global
323	Manufacture of sports goods	European
324	Manufacture of games and toys	European
325	Manufacture of medical and dental instruments and supplies	Global
329	Manufacturing n.e.c.	European
331	Repair of fabricated metal products, machinery and equipment	Domestic
332	Installation of industrial machinery and equipment	Domestic
352	Manufacture of gas; distribution of gaseous fuels through mains	Domestic
353	Steam and air conditioning supply	Domestic
360	Water collection, treatment and supply	Domestic
37T39	Sewerage; Waste; Other waste	Domestic
45T47	Motor vehicles; Wholesale; Retail	European
491	Passenger rail transport, interurban	Domestic
492	Freight rail transport	European
493	Other passenger land transport	Domestic
494	Freight transport by road and removal services	European
501	Sea and coastal passenger water transport	Domestic

Industry Code	Industry Description	Taxonomy Geography
502	Sea and coastal freight water transport	Global
503	Inland passenger water transport	Domestic
504	Inland freight water transport	European
51	Air transport	Global
52	Warehousing	European
53	Postal and courier activities	European
55T56	Accommodation & food services	Domestic
581	Publishing of books, periodicals and other publishing activities	Domestic
582	Software publishing	Global
59T60	Motion picture & broadcasting	Domestic
61	Telecommunications	Domestic
62T63	Computer programming & information	European
71	Architectural and engineering	European
72	Scientific R&D	European
73	Advertising and market research	European
77	Rental and leasing	European
781	Activities of employment placement agencies	European
782	Temporary employment agency activities	European
783	Other human resources provision	European
791	Travel agency and tour operator activities	European
799	Other reservation service and related activities	European
801	Private security activities	European
802	Security systems service activities	European
803	Investigation activities	European
812	Cleaning activities	Domestic

Note: The table presents the list of industries used to compute the baseline concentration measure and their associated geographic market.
Source: OECD compilation.

Table A 2. Industries included in the 4-digit analysis

NACE Rev.2 code and description		
2-digit	**3-digit**	**4-digit**
10 Manufacture of food products		
	10.1 Processing and preserving of meat and production of meat products	
		10.11 Processing and preserving of meat
		10.12 Processing and preserving of poultry meat
		10.13 Production of meat and poultry meat products
	10.2 Processing and preserving of fish, crustaceans and molluscs	
		10.20 Processing and preserving of fish, crustaceans and molluscs
	10.3 Processing and preserving of fruit and vegetables	
		10.31 Processing and preserving of potatoes
		10.32 Manufacture of fruit and vegetable juice
		10.39 Other processing and preserving of fruit and vegetables
	10.4 Manufacture of vegetable and animal oils and fats	
		10.41 Manufacture of oils and fats
	10.5 Manufacture of dairy products	
		10.51 Operation of dairies and cheese making
		10.52 Manufacture of ice cream
	10.6 Manufacture of grain mill products, starches and starch products	
		10.61 Manufacture of grain mill products
		10.62 Manufacture of starches and starch products
	10.7 Manufacture of bakery and farinaceous products	
		10.71 Manufacture of bread; manufacture of fresh pastry goods and cakes

		10.72 Manufacture of rusks and biscuits; manufacture of preserved pastry goods and cakes
		10.73 Manufacture of macaroni, noodles, couscous and similar farinaceous products
	10.8 Manufacture of other food products	
		10.81 Manufacture of sugar 1072
		10.82 Manufacture of cocoa, chocolate and sugar confectionery 1073
		10.83 Processing of tea and coffee
		10.84 Manufacture of condiments and seasonings
		10.85 Manufacture of prepared meals and dishes
		10.86 Manufacture of homogenised food preparations and dietetic food
	10.9 Manufacture of prepared animal feeds	
		10.91 Manufacture of prepared feeds for farm animals
		10.92 Manufacture of prepared pet foods
11 Manufacture of beverages		
	11.0 Manufacture of beverages	
		11.01 Distilling, rectifying and blending of spirits
		11.02 Manufacture of wine from grape
		11.03 Manufacture of cider and other fruit wines
		11.04 Manufacture of other non-distilled fermented beverages
		11.05 Manufacture of beer
		11.06 Manufacture of malt
		11.07 Manufacture of soft drinks;production of mineral waters and other bottled waters

Note: This table reports the NACE Rev. 2 code and description of the industries included in the 4-digit level of analysis of concentration
Source: NACE Rev.2 manual.

CENTRE FOR ECONOMIC PERFORMANCE
Recent Discussion Papers

2062	Cong Peng Yao Wang Wenfan Chen	Roads to development? Urbanization without growth in Zambia
2061	Gabriel M. Ahlfeldt Fabian Bald Duncan Roth Tobias Seidel	Measuring quality of life under spatial frictions
2060	Tessa Hall Alan Manning	Only human? Immigration and firm productivity in Britain
2059	Shadi Farahzadi	The integration penalty: Impact of 9/11 on the Muslim marriage market
2058	Italo Colantone Gianmarco I.P. Ottaviano Kohei Takeda	Trade and intergenerational income mobility: Theory and evidence from the US
2057	Maria Guadalupe Veronica Rappoport Bernard Salanié Catherine Thomas	The perfect match: Assortative matching in mergers and acquisitions
2056	Fabrizio Leone	Global robots
2055	Luca Fontanelli Flavio Calvino Chiara Criscuolo Lionel Nesta Elena Verdolini	The role of human capital for AI adoption: Evidence from French firms
2054	Saul Estrin Andrea Herrmann Moren Lévesque Tomasz Mickiewicz Mark Sanders	New venture creation: Innovativeness, speed-to-breakeven and revenue tradeoffs

2053	Stephen J. Redding	Quantitative urban economics
2052	Esteban M. Aucejo Spencer Perry Basit Zafar	Assessing the cost of balancing college and work activities: The gig economy meets online education
2051	Jonathan Colmer Suvy Qin John Voorheis Reed Walker	Income, wealth and environmental inequality in the United States
2050	Debopam Bhattacharya Ekaterina Oparina Qianya Xu	Empirical welfare analysis with hedonic budget constraints
2049	Jonathan Colmer Eleanor Krause Eva Lyubich John Voorheis	Transitional costs and the decline in coal: Worker-level evidence
2048	Ekaterina Oparina Andrew E. Clark Richard Layard	The Easterlin paradox at 50
2047	Stephen J. Redding	Spatial economics
2046	Stephen Machin Matteo Sandi	Crime and education
2045	Hanno Foerster Tim Obermeier Bastian Schulz	Job displacement, remarriage and marital sorting
2044	Randi Hjalmarsson Stephen Machin Paolo Pinotti	Crime and the labor market