OECD publishing

INDUSTRY CONCENTRATION IN EUROPE AND NORTH AMERICA

MATEJ BAJGAR
GIUSEPPE BERLINGIERI
SARA CALLIGARIS
CHIARA CRISCUOLO
JONATHAN TIMMIS

OECD PRODUCTIVITY WORKING PAPERS
January 2019 **No. 18**

GLOBAL FORUM ON PRODUCTIVITY

OECD
BETTER POLICIES FOR BETTER LIVES

OECD PRODUCTIVITY WORKING PAPERS

The OECD Productivity Working Papers are associated with the Global Forum on Productivity that provides a forum for mutual exchange of information and fosters international co-operation between public bodies with responsibility for promoting productivity-enhancing policies, including in undertaking joint policy analysis. It offers a platform for exchanging views, experiences and information, institutional and governance arrangements and government structures, with a view towards developing better policies. The Forum extends existing work in the OECD through a well-prioritised and coherent stream of analytical work serving the policy research needs of participants on the drivers of productivity growth.

This paper has been authorised for publication by Luiz de Mello, Director, Policy Studies Branch, Economics Department, Álvaro Santos Pereira, Director, Country Studies Branch, Economics Department, and Andrew Wyckoff, Director, Directorate for Science, Technology and Innovation.

Comments on OECD Productivity Working Papers are welcome and may be sent to: productivity@oecd.org

OECD Productivity Working Papers are published on: oe.cd/productivity

Series: OECD Productivity Working Paper Series

ISSN 2413-9424

Please cite this paper as:

Bajgar, M., Berlingieri, G., Calligaris, S., Criscuolo, C. and Timmis, J., "Industry Concentration in Europe and North America", OECD Productivity Working Papers, 2019-18, OECD Publishing, Paris.

ABSTRACT/RÉSUMÉ

Industry Concentration in Europe and North America

This report presents new evidence on industry concentration trends in Europe and in North America. It uses two novel data sources: representative firm-level concentration measures from the OECD MultiProd project, and business-group-level concentration measures using matched Orbis-Worldscope-Zephyr data. Based on the MultiProd data, it finds that between 2001 and 2012 the average industry across 10 European economies saw a 2-3-percentage-point increase in the share of the 10% largest companies in industry sales. Using the Orbis-Worldscope-Zephyr data, it documents a clear increase in industry concentration in Europe as well as in North America between 2000 and 2014 of the order of 4-8 percentage points for the average industry. Over the period, about 3 out of 4 (2-digit) industries in each region saw their concentration increase. The increase is observed for both manufacturing and non-financial services and is not driven by digital-intensive sectors.

JEL classification: D4, L11, L25

Keywords: Industry concentration, business dynamics, measurement.

****

Concentration Industrielle en Europe et en Amérique du Nord

Ce rapport présente des éléments nouveaux concernant les tendances de la concentration industrielle en Europe et en Amérique du Nord. Il utilise deux nouvelles sources de données : des mesures de concentration représentatives au niveau de la firme provenant du projet MultiProd de l'OCDE, et des mesures de concentration au niveau du groupe d'entreprises utilisant des données appariées Orbis-Worldscope-Zephyr. Sur la base des données MultiProd, le rapport établit qu'entre 2001 et 2012 un secteur industriel moyen de 10 économies européennes a vu croître de 2 à 3 points de pourcentage la part des 10 % plus grandes entreprises en termes de ventes de cette industrie. A partir des données Orbis-Worldscope-Zephyr, le rapport montre clairement une augmentation de la concentration industrielle en Europe ainsi qu'en Amérique du Nord entre 2000 et 2014 de l'ordre de 4 à 8 points de pourcentage pour un secteur industriel moyen. Sur la période, environ 3 industries (au niveau sectoriel à 2 chiffres) sur 4 de chaque région ont vu leur concentration s'accroître. On observe cette augmentation aussi bien dans le secteur manufacturier que dans les services et elle n'est pas induite par les secteurs à haute intensité numérique.

Classification JEL : D4, L11, L25.

Mots-clés : Concentration industrielle, dynamique des entreprises, mesure.

Table of contents

Tables

Figures

Boxes

Industry Concentration in Europe and North America

By Matej Bajgar, Giuseppe Berlingieri, Sara Calligaris, Chiara Criscuolo and Jonathan Timmis [1]

1. Introduction

1. A heated debate is under way about the evolution of industry concentration in OECD economies in recent years. A number of studies suggest that industry concentration has increased over recent years in the United States (e.g. Furman and Orszag, 2015; Grullon et al., 2015; Autor et al., 2017b) and to a lesser extent in Japan (Honjo, Doi and Kudo, 2014). However, the evidence for other parts of the world is to date limited and inconclusive. Initial results for Europe have so far offered little indication of increasing concentration (Valletti et al., 2017; Social Market Foundation, 2017; Gutiérrez and Philippon, 2018). This has led some economists to suggest that European markets have become more competitive than those in the United States (Gutiérrez and Philippon, 2018).

2. Understanding trends in industry concentration is important because they can reflect or have implications for a range of economic phenomena. Increasing industry concentration may indicate technological change or globalisation allowing the most productive firms to expand (Autor et al., 2017b). An increasing scale of a few firms may also mean fewer buyers in input markets and local labour markets – i.e. monopsony – potentially impacting contractual terms for suppliers and workers (OECDa, 2008). In addition, lobbying is more likely to be undertaken by larger firms and by firms in concentrated markets, which may inform policy differentially in concentrated industries (Dellis and Sondermann, 2017). Last but not least, high concentration may impact firm risk-taking behaviour if they can be seen as "too big to fail".

3. Industry concentration is also sometimes seen as a proxy for the degree of competition, which, in turn, influences a variety of economic outcomes. Product market competition has been shown to be positively correlated with productivity (Disney et al., 2003; Aghion et al., 2009), particularly for firms far from or close to the frontier (Andrews et al., 2016). Industry mark-ups, again a proxy of the degree of competition, can also affect innovation (Aghion et al., 2005), industry concentration is correlated with income distribution (Siegenthaler and Stucki, 2015; Autor et al., 2017a,b) and foreign competitors may impact economic volatility (Cravino and Levchenko, 2017).

4. Industry concentration is related to, but distinct from, the concept of *market* concentration[2]. Industry concentration measures the extent to which economic activity is concentrated within a small number of large companies or business groups within an

[1] Matej Bajgar, Giuseppe Berlingieri, Sara Calligaris, Chiara Criscuolo (corresponding author: Chiara.Criscuolo@oecd.org) and Jonathan Timmis from the OECD Science, Technology and Innovation Directorate. The authors are grateful to Antonio Capobianco, Antonio Gomes, James Mancini, Chris Pike from the OECD Directorate for Financial and Enterprise Affairs, Laurence Boone from the OECD Economics Department, and Sarah Box, Dirk Pilat and Andrew Wyckoff from the OECD STI Directorate, as well as members of the Global Forum on Productivity Steering Group, for helpful comments on earlier versions of this paper. We thank Sarah Michelson (OECD Economics Department) and Petia Petrova (OECD STI Directorate) for their excellent editorial support.

[2] Market competition is often the focus of competition authorities.

industry. Market concentration, instead, describes the weight of leading firms in a market for particular products or services that are close substitutes. Accordingly, market concentration is a far narrower definition than what is typically reflected in industry concentration measures. The fact that a large share of industry activity is due to a handful of large firms does not necessarily mean that product markets within that industry are highly concentrated. While industry concentration can be used as an initial indicator to screen for changes in the degree of competition, by itself it can say little about whether or not market competition is changing.[3] Looking at a range of additional metrics – such as mark-ups, profitability, mergers and acquisitions (M&A) activity and business entry and exits – provides a better indication of whether there are changes in the competitive environment.[4] As we discuss later in this paper, these metrics are encompassed within our broader research in this area (see Box 1.2 in particular).

5. This paper contributes to the debate by presenting two new complementary pieces of evidence on industry concentration trends in Europe and in North America.

6. The first piece of evidence consists of representative firm-level concentration measures from the OECD MultiProd project.[5] We document how the share of the 10% largest companies (by sales) has evolved in 10 European economies between 2001 and 2012. We show that the share of industry sales due to these firms has increased on average by 2 percentage points in manufacturing and 3 percentage points in non-financial market services. The *employment* share of the 10% largest firms (again as defined by their sales) has also increased in services, but not in manufacturing. Concentration is measured at the firm-level within each country and 2-digit sector.

7. The second piece of evidence relies on business-group-level concentration measures using matched Orbis-Worldscope data (see Box 1.1 for an overview of datasets used in this paper). Firm-level measures, such as those underlying MultiProd, may understate concentration if many firms are actually part of the same business group. In this paper, we explicitly take account of some firms being part of larger business groups. Specifically, we rely on the rich time-varying ownership information in matched Orbis-Zephyr data to take account of the structure of each business group and apportion group sales to the countries and industries where it operates. We also use information from OECD STAN and OECD SNA data to obtain reliable industry sales denominators for our concentration measures. [6] The resulting database covers 22 countries in Europe and 2 countries in North America, for the period 2000-2014. Concentration is measured at the

[3] The imprecision flows from the fact that to get from industry concentration to market competition one needs to be confident that a) there is a reliable relationship between industry concentration and market concentration (see OECD, 2018, and Werden & Froeb, 2018), b) that market concentration is a good indicator of market power (often not the case in differentiated product or geographic markets, platform markets and innovative markets); and c) that market power reflects a lack of competitive intensity rather than being a sign of competition in action.

[4] See OECD (2018) for more detail on the use of industry concentration as a proxy measure of competition intensity. The instances of consistency between the changes in industry concentration and the changes in other, more reliable (but still individually imperfect) indicators should not be taken to suggest that industry concentration is as reliable as they are, at least not without evidence on the systematic correlation of such measures.

[5] For more information, see http://www.oecd.org/sti/ind/multiprod.htm.

[6] For more information, see http://oe.cd/stan and https://doi.org/10.1787/na-data-en.

business group-level, treating each global region as a single market, and calculated within 2-digit industries.[7]

Box 1. Overview of MultiProd and Orbis-Worldscope-Zephyr Data

We construct two complementary measures of concentration: using representative firm-level data from the OECD MultiProd project and business group-level data from Orbis-Worldscope-Zephyr data.

The MultiProd data has the unique advantage of being representative firm-level data that has been constructed comparably across countries. The MultiProd project is based on a standardised STATA® code that aggregates micro-data from national statistical offices. In cases where the micro-data is limited to a sample of firms, rather than the entire population, the data are re-weighted using business registers so they are representative of the whole population of firms. For the purposes of this analysis we restrict our MultiProd sample to the manufacturing and non-financial market services sector, as well as to European countries. We keep only European countries for which data are fully representative of the population of firms. The list includes Austria, Belgium, Denmark, Finland, France, Germany, Hungary, Norway, Portugal and Sweden. For most countries, the time period spans from early 2000s to 2012. For Austria the time horizon is shorter (starting in 2008), whereas for Finland, France and Norway data are available at least since 1995.

The Orbis-Worldscope-Zephyr data has the major advantage that we are able to construct business-group level measures of concentration across countries. Enterprise-level measures, such as those underlying MultiProd, may understate concentration if many enterprises are actually part of the same group.[8] To measure business group-level activity requires both group and subsidiary financial information, as well as detailed ownership information detailing parent-subsidiary linkages.

The business group and subsidiary financial information is primarily sourced from Orbis, which we supplement with Worldscope data for listed firms (Orbis includes both listed and non-listed). Orbis is the largest cross-country firm-level database that is available for economic research and has been widely used in OECD research (e.g. Andrews et al., 2016; Calligaris et al., 2018). It is a commercial database provided to the OECD by Bureau Van Dijk. Worldscope is a cross-country firm-level commercial database of listed firms provided to the OECD by Thomson Reuters. The firms in Worldscope represent about 95% of global stock market capitalisation. To capture the activity of all major subsidiaries of the group, a broad range of countries and industries is required. Our resulting subsidiary data contains information across 100 economies and all business sectors, from 2000-2014.

[7] Two-digit industries are the most disaggregate industry level at which we can obtain reliable STAN and SNA data to serve as denominator in our concentration measures.

[8] In this paper, we interchangeably refer to enterprises and to firms, treating them as synonyms. In both cases we have in mind registered legal entities (e.g. a UK retailer). Enterprises/firms can operate at multiple locations and therefore by definition consist of multiple establishments (e.g. branches of a UK retailer). We use business groups to mean corporate groups and consist of all subsidiary firms that share a common parent firm (e.g. a multinational retailer may control several subsidiary firms, such as domestic and foreign retailers, banks, logistics firms etc.). Note that for many smaller entities in our sample there is no such distinction, the enterprise and the parent firm are equivalent (i.e. they are single-firm groups).

However, we report concentration trends only for better-covered economies and industries – focusing on manufacturing and non-financial market services for 21 economies.

Our primary source of parent-subsidiary ownership information is Orbis, which we supplement with data from the Zephyr database of Mergers and Acquisitions (M&As), both provided by Bureau Van Dijk (BvD). Orbis contains comprehensive information on ownership linkages between firms, however the data primarily start in 2007 and sometimes much later for some firms. Our primary use of the Zephyr M&A database is to capture missing earlier changes in ownership, enabling the construction of an ownership series starting as early as 2000 for some firms, and until 2014 in most cases. We undertake extensive cleaning that includes a wide range of automatic checks together with manual checks of the largest business groups in each industry and world region against the subsidiaries listed in their financial statements.

Further details on the data are available in Annex A.

8. Calculating business-group level metrics has required extensive data work and computation time. The "apportioning" approach requires both good financial data for each business group and its subsidiaries and detailed information on ownership linkages between each subsidiary and their parents over time. Accordingly, we have combined Orbis data on business group owners (global ultimate owners), immediate ownership linkages and Zephyr data on Mergers and Acquisitions to generate an ownership series for 2.8 million firms that extends from 2000 to 2016. We have also supplemented Orbis financial data with data on listed firms from Worldscope. In both financial and ownership data, the cleaning included a wide range of automatic checks together with manual checks of the largest business groups in each industry and world region against their financial statements. See Annex A for more information.

9. We document a clear increase in industry concentration in Europe as well as in North America between 2000 and 2014 of the order of 4-8 percentage points for the mean industry. Over the period, about 75% of 2-digit industries in each region saw their concentration increase. The measured absolute increase is somewhat greater in North America, but this could be a result of differences in data coverage between the two regions. The results hold independently of whether larger industries are given a greater weight in the calculations, whether the increase is calculated in absolute or proportional terms and for both manufacturing and non-financial market services. Additionally, the overall increase in concentration does not seem to be driven in particular by digital-intensive industries.

Box 2. Broader Ongoing Work on Competition, Innovation and Productivity

This paper forms one part of a broader theme of work on competition, innovation and productivity being conducted within the context of the Global Forum on Productivity, Inclusive Growth, Going Digital horizontal project and the CIIE.

The ongoing research assesses changes in the competitive environment by presenting a broad range of metrics at the level of firms, business-groups and industries. Existing ongoing projects in collaboration with STI presents evidence across a range of measures that include firm and business-group industry concentration, common ownership of firms, firm profitability, firm mark-ups, business dynamism, entry into innovation, innovation concentration.

The research also outlines several possible drivers of changes in the competitive environments, including the digital transformation, mergers and acquisitions (M&As), venture capital and start-up accelerators, the growing importance of intangibles including patents, as well as competition, tax, intellectual property, innovation and entrepreneurship policies.

Continuing research examines the impact of changes in the competitive environment across a range of firm and industry outcomes. These include firm productivity and productivity diffusion, laggard firms, productivity-wage differentials across firms, profitability, tax revenues and the nature and quality of innovation.

The specific projects contributing to these different aspects are detailed below:

- **M&As, concentration and productivity** is a project conducted under the Global Forum on Productivity. It empirically examines the link between M&As, measures of industry concentration and firm productivity growth and productivity diffusion. See the scoping paper here.
- **Industry Concentration in Europe and North America** (this paper) presents summary metrics on firm-level and business-group-level concentration across industries, countries and regions, examining the role of digital sectors in particular.
- **M&As and innovation** is a project conducted under the horizontal Going Digital project. It empirically investigates how M&As affect concentration of and entry into innovation, and how they impact the level, the quality and the location of firms' innovative activity. See the scoping paper here.
- **Mark-ups in the digital era** is a study which explores the evolution of firm mark-ups across countries and the role of the digital transformation, innovation and intangibles in driving these trends. An older version of this paper is available here.
- **Firm characteristics and CIT payments: Micro-level evidence** is a forthcoming study conducted under the horizontal Going Digital project. It examines the link between concentration and firm profitability and tax revenues.
- **Business dynamics and digitalisation** is a forthcoming study under the horizontal Going Digital project which analyses the how different facets of the digital transformation affect business dynamics across countries.
- **Declining business dynamism: evidence and causes** is a forthcoming study which points to some key mechanisms and policies that may be associated with the observed trends in business dynamism across countries.
- **Last but not least: laggard firms, technology diffusion and its structural and policy determinants** is a forthcoming study which documents characteristics of the least productive firms, examines their potential for productivity growth and explores the role of structural and policy determinants in their catch-up.
- **The Productivity-Wage Premium: Does Size Still Matter in a Service Economy?** is a study based on data from the MultiProd project presenting evidence that, in services, the most productive firms are not necessarily the largest ones in terms of employment but they do pay the highest wages. The paper is available here.

10. Our results contribute to a broader line of research on how business dynamism in developed economies has changed over time. This research suggests that economies appear

to be less dynamic, with declining entry and exit rates across most OECD economies[9] and increased divergence between top and bottom productivity firms.[10] At the same time, market power appears to have increased, as indicated by increasing mark-ups of top firms[11] and falling labour share of income.[12] Whether these findings are the result of a reduction in competition, or a sign of competition in action – with market power representing a temporary reward for innovative and efficient firms – remains unclear. Our results should, thus, not be interpreted as unambiguous evidence of reduced competition, and much less yet of a need for particular policy interventions.

11. Instead, this paper intends to fill a factual void and present new evidence regarding industry concentration trends in North America and particularly Europe. Understanding concentration trends in Europe is important on its own but also may shed some light on the mechanisms behind the causes of the trends observed elsewhere. If industry concentration has been increasing only in some countries, we should search for country-specific factors as drivers of these trends. If, on the contrary, it has been concurrently increasing in many OECD countries, broader factors, such as technological change or globalisation, are more likely culprits (Van Reenen, 2018). That we find a broad increase in concentration across regions and broad industries, tentatively points to the latter possibility, but more work and data at different levels of aggregation is needed in this area to examine potential nuances across countries.

12. This paper sits within a broader stream of ongoing work which examines different aspects of the competitive environment, their trends over time, potential drivers of these trends and their impacts on various firm and industry outcomes (see Box 1.2).

13. Given the uncertainty about the drivers of the observed trends, this paper intentionally does not encompass potential policy implications. A change in industry concentration can be caused by a myriad of factors, some of them related to policies and others not. While some of the increase in concentration could be related to anti-competitive regulations or the competition policy environment, it could just as well be that technological developments, integration of global markets or sustained innovation allow the most efficient firms to increase their competitive edge over other firms, contributing to welfare gains and productivity growth.

14. The rest of the paper is organised as follows. The next section briefly reviews the existing evidence on concentration trends and the third explains how we measure concentration in MultiProd and Orbis data. The fourth section presents concentration trends using MultiProd firm-level data, and the fifth section presents concentration trends using Orbis business group-level data. The final section briefly discusses future work that could extend the analysis presented in this paper.

[9] The trends were initially documented for the US; see, for example, Decker et al. (2014) and Decker et al. (2016). For cross-country evidence, see Criscuolo et al. (2014).

[10] See Andrews et al. (2016) and Berlingieri et al. (2017b).

[11] De Loecker and Eeckhout (2017) for the US and Calligaris, Criscuolo and Marcolin (2018) for 26 mostly OECD economies.

[12] Karabarbounis and Neiman (2014) and OECD (2015) document declining labour shares in a large number of countries. Autor et al. (2017b) and Barkai (2017) show evidence for the US and discuss some potential drivers of the observed trends. De Loecker and Eeckhout (2017) show how the rise in markups implies a decrease in labour share.

2. Existing evidence on concentration and competition trends

15. This paper builds upon a growing number of studies pointing to a weakening of business dynamism and competition in the United States over recent years. They have used a range of complementary metrics such as business dynamism, profitability, mark-ups and industry concentration ratios and a variety of data sources to develop this narrative.

16. The recent interest in industry concentration in the United States was triggered by a paper by Furman and Orszag (2015), who analyse the sales shares of the 50 largest companies within 2-digit sectors using the official firm-level Census Bureau data and find that about three quarters of these sectors see an increase in concentration between 1997 and 2007. Autor et al. (2017b) similarly rely on the Census Bureau data, but analyse a range of concentration metrics (CR4, CR20, Herfindahl)[13] within 4-digit SIC industries between 1982 and 2012 and find that CR4 increased on average by 4% in services, 5% in manufacturing, 6% in wholesale, 8% in utilities, 11% in finance and 15% in retail. Another study examining Census Bureau data was undertaken by the Economist Magazine in 2016 (Economist, 2016). Other studies have instead used Compustat data, which covers listed firms, again finding an overall increase in industry concentration (Grullon et al., 2015; Gutiérrez and Philippon, 2017).

17. The evidence on concentration has been complemented by studies looking at other proxies of market competition. For example, De Loecker and Eeckhout (2017) find a strong increase in average mark-ups across many US industries, rising from 18% in 1980 to 67% in 2014, and Hall (2018) obtains similar results. Bessen (2016) in turn documents a rise in operating margins, and influential studies by Decker et al. (2014, 2016) offer evidence of declining business dynamism and entrepreneurship in the US over recent years.

18. The primary contribution of this paper is bringing new evidence on concentration trends in Europe, where the literature is so far limited and inconclusive, and, in doing so, also offering a new perspective on trends documented in other parts of the world, using different measures of concentration and different data sources. The paper most closely related to our paper is Gutiérrez and Philippon (2018), who use Orbis data and find that concentration ratios have remained broadly stable in Europe, both when calculated within countries and when treating all of Europe as a single market. Using Euromonitor data for 2010-2015, Valletti et al. (2017) find mixed concentration trends for European economies. Analysing 10 consumer product markets in the UK,[14] the Social Market Foundation (2015) finds varying concentration trends across specific markets.

19. These findings seem at odds with recent findings on other measures of competition. For example, at an aggregate-level, Valletti et al. (2017) find that profits are rising as a share of GDP across the EU5 since the mid-1990s, and this growth has been in line with the United States. At the firm-level, Calligaris et al. (2018) find strong evidence that mark-ups are increasing over the period 2001-2014 using Orbis data for 26 high-income economies, and this result holds even when US firms are excluded from the sample. The increase is particularly pronounced in the top half of the mark-up distribution and for firms in digital sectors. There is also widespread evidence of declining business dynamism across

[13] CR4 and CR20 are respectively defined as the share of the largest 4 and the largest 20 firms in the industry total, as measured, for example, by sales, employment or production capacity. The Herfindahl Index is defined as the sum of squared shares of all firms in an industry.

[14] These include, Cars, Electricity, Groceries, Gas, Broadband, Personal current accounts, Mobile telephony, Credit cards, Landline-only phone contracts and Mortgages.

European economies, as noted in Calvino et al. (2015). Whilst these metrics reflect different aspects of the competitive environment, one would expect *a priori* for these to move broadly in the same direction, as appears to be the case in the United States.

20. In our analysis using Orbis data, we construct measures of concentration using business group-level data. Some financial data in Orbis is at the business group-level and others at the subsidiary-level. Incorrect treatment of these complexities may inadvertently lead to double-counting, omission or incorrectly allocating activity to industries and countries (we discuss this in detail in the next section). To avoid such issues, we draw from Bloom et al. (2013) who similarly use the information on business group structure in Orbis to apportion group sales to individual subsidiaries, although their aim is to measure product market distance between groups. Other Orbis-based studies on which we build are Altomonte et al. (2018), who analyse how the institutional environment influences the development of complex group hierarchies, and Conconi et al. (2018), who study knowledge diffusion of export market destinations within multinational business groups. Our approach of apportioning sales of each group among multiple countries and industries is also related to a recent study which compares the effect of top firms on concentration at the national level to concentration effects of the top firms' presence in local areas such as counties and Zip codes (Rossi-Hansberg et al., 2018).

3. How do we measure concentration?

21. This section outlines how we measure industry concentration in this study. We first use representative MultiProd firm-level data and then move to business-group-level data from Orbis-Worldscope-Zephyr. All concentration measures aim to capture the weight of the largest firms within an industry, but they can differ in several aspects: firstly, in terms of what is an "industry"; secondly, in their definition of "largest"; third, in their choice of denominator measuring activity of the whole industry; and finally, in focusing on plants, firms or business groups. We consider each of these aspects in turn.

3.1. Industry definition

22. Economic analyses of concentration trends typically use an industry definition based on the industry classification recorded by firms (such as SIC, NAICS, NACE, ISIC etc.; see for instance Autor et al., 2017b). Analyses for a single country often focus on a high degree of disaggregation, such as 3- or 4-digit industries, since firms within more narrowly defined industries are more likely to compete with each other. Cross-country studies, however, are rarer and often constrained by data availability to a 2-digit level, as we are in our analysis using Orbis and MultiProd.[15]

23. In this paper, we present MultiProd concentration ratios calculated within individual countries and Orbis concentration metrics for two world regions – North America and Europe. The latter treats each region as a separate market and is motivated by the number of large business groups that operate across economies (see Altomonte et al.,

[15] Orbis data contain industry information at the 4-digit level, but we are limited by the availability of cross-country information on aggregate industry output, which is only reported at the level of 2-digit industries. In order to maximise country and industry coverage, we aggregate multiple NACE Rev. 2 2-digit industries together in several cases. The industry classification which we use in our analysis of Orbis is related to the STAN A64 classification and our MultiProd-based analysis is at the level STAN A38 industries. For information on A64 and A38 classifications, see http://www.oecd.org/sti/ind/3max.pdf.

2018 amongst others). In future work, we are also planning to develop Orbis concentration metrics at the country-level for a set of countries which are well-covered in the data.

24. The industry definitions we use here differ from the definition of a "market" often used by competition authorities, which typically is narrower and relies on expert definitions of markets. An interesting but also challenging extension of our work would involve defining markets base on contested merger cases, which define the set of relevant competing firms (see Haucap and Stiebale, 2016, for an example). However, such data are not comprehensive and only define a market for contested mergers at a point in time in a particular context, and as such they are not easily generalisable.

3.2. Definition of "Largest"

25. The literature has used various definitions of "largest" firms in an industry. The most common definition relies on an absolute number of the largest firms in terms of sales (e.g. top "CR" 4, 8 or 20 largest firms). Alternatives include the largest percentage of firms (e.g. top 10% of firms) or a Herfindahl index (based on the distribution of firm market shares). MultiProd provides the sales of the top 10% of firms (in terms of sales) for each country and 2-digit industry.[16] For Orbis, we use the measures based on absolute numbers of top firms – CR4, CR8 and CR20 – as our preferred measures.

26. The choice between these metrics is important, as the use of Herfindahl indices or the top 10% of firms is not appropriate when the coverage of firms varies across industries or over time, as is the case with Orbis. Since MultiProd is representative data of the population of firms, the set of top 10% of firms is measured consistently over time. In Orbis, however, the sample size typically improves in later years, with smaller firms often added over time (see Bajgar et al., forthcoming). As the coverage expands, more and smaller firms will cross the threshold of the top 10%, with the effect of reducing the measured concentration. Similarly, since a Herfindahl index relies on the distribution of market shares in an industry, changes in the coverage of firms will again lead to artificial changes in the resulting concentration index. These problems remain even when re-weighting Orbis using employment, since Orbis disproportionately covers prominent and successful firms even *within* size classes (see Bajgar et al., forthcoming). Our definition of concentration based on absolute numbers of firms (CR4, CR8 and CR20) mitigates many of these problems in Orbis, since the data typically contain the largest firms in each industry throughout the period, and the small number of groups included in our measures allows us to manually check that important firms do not suddenly appear or disappear during the sample period as a result of coverage changes.

27. It is also important to decide on the variable in terms of which the weight of each firm is measured. Most studies measure industry concentration in terms of sales, but other dimensions, such as employment, value added, capacity, innovation or valuation could also be considered. In this paper, we also primarily focus on sales, and in our analysis of MultiProd data we additionally report concentration of employment.

3.3. Industry Denominator

28. To construct market shares of the largest firms in each industry, their sales need to be scaled by the total sales in each industry. In principle, this information can either be

[16] Other concentration metrics are not available in the current version of the MultiProd data but will be collected in the future.

calculated by summing sales across all firms in the microdata at hand, or it can come from an industry-level database such as OECD STAN. For the MultiProd country-level analysis, we take the former approach, using the total firm sales for each country as calculated from microdata in the MultiProd project. We take the latter approach for the Orbis regional-level analysis, where we obtain a regional measure of sales by aggregating OECD STAN data across countries within the region.

29. The choice of industry denominator is non-trivial – using the total sales of firms in the data is not appropriate when the coverage of firms varies across industries or over time, as is the case with Orbis. Since MultiProd is representative data at the country-level, total firm sales for that country closely track STAN aggregates (Bajgar et al, 2017). However, for Orbis this is not the case. As noted above, larger firms are often well-covered in Orbis from earlier years, but the sample size typically improves in later years with smaller firms often added over time. Therefore, the industry denominator will tend to increase with coverage in Orbis and lead to artificial changes in the resulting concentration index. However, the addition of larger firms in earlier years (in the raw data) and smaller firms in later years, can make the industry denominator grow differentially over time, which is further complicated if large and small firms have differing sales growth rates. Our definition of denominator mitigates many of these problems, since OECD STAN is comparable across industries and countries over time.

3.4. Concentration Metric

3.4.1. Using MultiProd

30. Our analysis using MultiProd data constructs two concentration measures. The first measure is the market share accounted for by the top decile of firms (ranked according to their sales). The higher the market share, the bigger the weight that these firms have in the economy, and the more concentrated is the economic activity. The indicator is computed as

$$\vartheta_j^{10} \equiv \sum_{i \in S_j^{\overline{10}}} \frac{s_{ij}}{s_j},$$

where S_{ij} designates the sales of firm *i* operating in industry *j*, S_j the total sales in industry *j* and $S_j^{\overline{10}}$ the top decile in the 2-digit sector j of firms ranked by sales.

31. The second indicator examines the share of employment that the same firms (i.e. firms with highest sales) represent to understand their influence on labour markets:

$$L_j^{10} \equiv \sum_{i \in S_j^{\overline{10}}} \frac{L_{ij}}{L_j},$$

where *L* corresponds to employment. Therefore, concentration is measured by the share of gross output and/or employment in the top decile of the sales distribution, and it is measured at the 2-digit level.

3.4.2. Using Orbis

32. In the analysis based on Orbis data, we measure concentration as the share of industry sales due to the four largest companies (or groups) in the industry. Formally, we define it as

$$CR_4 \equiv \sum_{i=1}^{4} \frac{S_{ij}^{ORBIS}}{S_j^{STAN}},$$

S_{ij}^{ORBIS} comes from Orbis data and designates the sales of firm i which is among the top 4 firms in industry j. S_j^{STAN} marks the total output of industry j as reported in the OECD STAN database, which is derived from national accounts.[17]

3.5. Firm-level or business group-level analysis

3.5.1. Existing approaches in the literature and MultiProd

33. The final choice is whether one considers sales of enterprises or business groups. Many large firms may not be independent, but rather part of the same business group (Altomonte et al., 2018). Neglecting these firm ownership linkages may lead to an understatement of concentration. In addition, the activities of the large business groups can often spread across several industries and countries, and therefore, focusing on individual subsidiaries may miss these linkages across markets and industries. Therefore, measuring concentration at the level of firms and business groups can reflect distinct aspects of the underlying economic trends.

34. Firm-level analysis often does not contain detailed information on business group ownership, as in the case of MultiProd. Statistical offices can collect information on domestic firm owners, including domestic business groups, but the identity of foreign multinational owners is often unknown, especially to external researchers. In addition, information on those subsidiaries of locally present foreign groups which are located in other countries is never recorded. It is then impossible to measure domestic subsidiaries of the same foreign multinational (for a notable exception, see Conconi et al., 2018). Without detailed information on domestic and foreign firm-to-firm ownership linkages, it is not possible to construct measures at the business group level. For these reasons, MultiProd reports concentration at the level of individual firms.

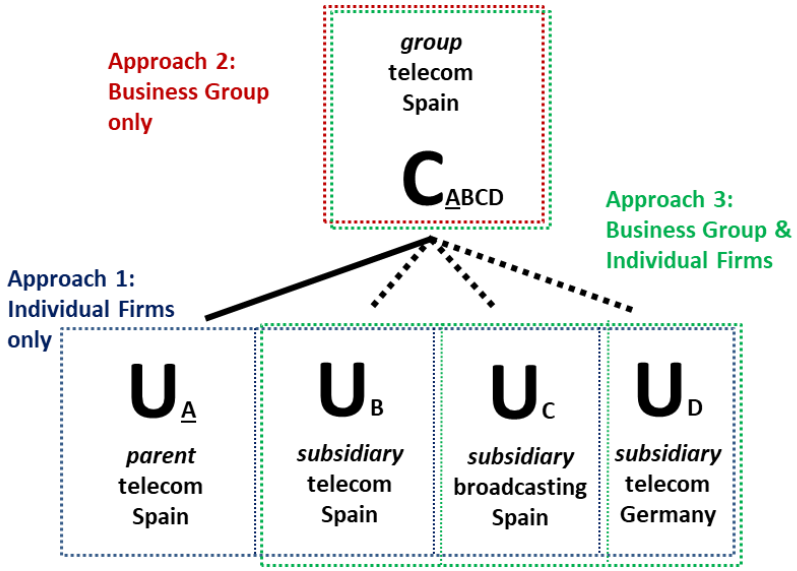
35. Orbis contains comprehensive, enormous firm-to-firm data on ownership linkages, for both domestic and foreign ownership, which allows us to construct measures of business groups. The ownership data has been used extensively in the academic literature on foreign ownership and international shock transmission (e.g. Cravino and Levchenko, 2017).

36. Unfortunately, the data primarily start in 2007. We therefore supplement the Orbis ownership data with the Zephyr M&A database to capture earlier changes in ownership and construct a series starting as early as 2000 for many firms. We then undertake extensive ownership checks and cleaning, including manually inspecting the subsidiaries listed in financial statements for the very largest 300 groups. These steps are detailed in the Appendix.

37. How a researcher treats business groups in Orbis can have substantial impacts on the resulting concentration measures (see Figure 1).

[17] Due to data limitations, we use Orbis *sales* in the numerator but STAN *output* in the denominator. For most industries, sales and output are very similar, although a more pronounced difference might exist in certain industries, such as Wholesale and Retail.

Figure 1. Approaches to selecting between Business Group and Individual Firm information in Orbis



Notes: The figure depicts an example of a hypothetical group consisting of a parent company operating in the telecom sector in Spain, with a Spanish and German subsidiary in the same sector and a Spanish subsidiary operating in broadcasting. Information is available at the business group level ("consolidated) ($C_{\underline{A}BCD}$) alongside information for each individual firm ("unconsolidated"), including the parent company ($U_{\underline{A}}$) and the subsidiaries (U_B, U_C, U_D). Approach 1 uses only individual firm level information ("unconsolidated"). Approach 2 uses only business group level information ("consolidated"). Approach 3 uses business group level information when available and individual firm level information otherwise.

38. The literature has taken several approaches here, each of which has particular limitations:

1. A first approach is to neglect business groups and focus only on individual firms ("unconsolidated" information in Orbis). This *will underestimate concentration* if multiple firms in the same market are part of the same group and exacerbates coverage issues. For instance, most US firms in the database have only business group information.

   Example: If a large firm acquires one of its main competitors and one only considers individual firm-level information and ignores group-level information, measured concentration will counterintuitively remain unchanged.

2. A second approach is to neglect subsidiaries and focus only on the headquarters of business groups ("consolidated" information in Orbis). Accordingly, this attributes the entire activity of the business group to the country of the headquarters and industry of the primary activity. This *will overestimate concentration in the headquarter* industry and country and *underestimate concentration in subsidiaries'* industries and countries, if the group activity is spread across several industries and countries.

Example: In 2015, Telefónica business group recorded global revenues of 47 billion Euros.[18] Yet in the same year, the total output of the "Telecommunications" (ISIC 61) sector in Spain was about 31 billion (OECD STAN). Allocating all of Telefónica business group's revenues to the headquarters would mean measured CR1 concentration in Spanish telecommunications is 150%.

Example: Telefónica is also a major telecommunications player in Germany and Brazil, and the largest subscription television provider in Spain. Neglecting Telefónica's subsidiaries will underestimate the concentration in these subsidiary markets.

3. A third approach is to include both the activity of the business group and other firm subsidiaries. In this approach, researchers drop the firm-level information ("unconsolidated") for headquarter firms – to exclude the most obvious source of double counting – but still include their subsidiaries.[19] This *will overestimate concentration* since it will double-count subsidiary revenues.

   Example: This approach will count the revenues of each Telefónica's subsidiary twice: when calculating concentration in Spanish telecommunications and again when calculating it in the country and the industry where the subsidiary is located.

39. The problem with the approaches introduced above can be seen as undertaking too little (approach 1) or too much (approaches 2 & 3) consolidation. Our approach aims to reduce the biases arising from these approaches by combining information from both types of accounts.

3.5.2. Our business-group apportioning approach using Orbis

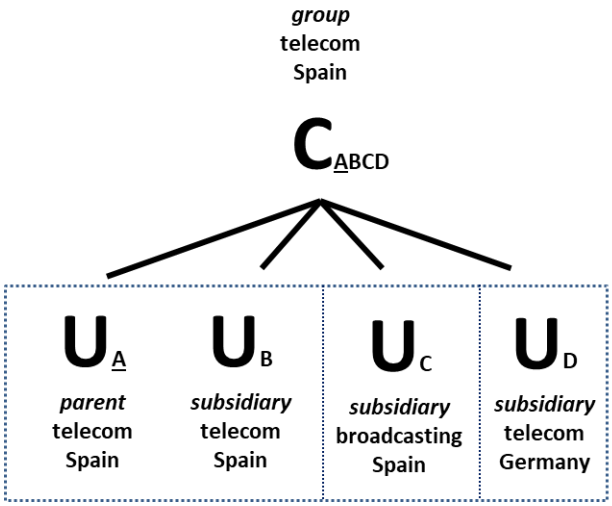
40. Our approach leverages the detailed Orbis-Zephyr ownership database to construct measures of business groups and apportion activity across the countries and industries in which the group operates. This apportioning approach combines the advantages of the three procedures noted above. We specifically draw inspiration from Bloom et al. (2013) who use a similar approach to apportion business group sales to individual subsidiary industries and countries to measure the product market distance between business groups.

41. Our approach uses the business group structure to measure the group sales within each industry and country. In particular, we aggregate the sales of subsidiaries of a business group within the same market (e.g. 2-digit industry and a country), as noted in Figure 2. In contrast to Approach 1 above, this means that if a business group acquires one of its main competitors in the same market, the business group's market share would increase and hence our measured concentration would increase.

[18] https://www.telefonica.com/documents/153952/13347920/informe_integrado_2015_en.pdf/6aee51fd-ed01-430c-a7dd-e8342922f97d

[19] This is often the case because of complexities in using the Orbis ownership data to identify the subsidiaries of parent business groups.

Figure 2. Aggregating within but not across markets



Notes: The figure depicts an example of a hypothetical business group consisting of a parent company operating in the telecom sector in Spain, a Spanish and German subsidiary in the same sector and a Spanish subsidiary operating in broadcasting. Information is available at the business group level ("consolidated) ($C_{\underline{A}BCD}$) alongside information for each individual firm ("unconsolidated"), including the parent company ($U_{\underline{A}}$) and the subsidiaries (U_B, U_C, U_D). Our "apportioning" approach aggregates revenues across firms operating in the same market but not across firms operating in different markets – so aggregates telecom revenues in Spain, but not other countries or industries.

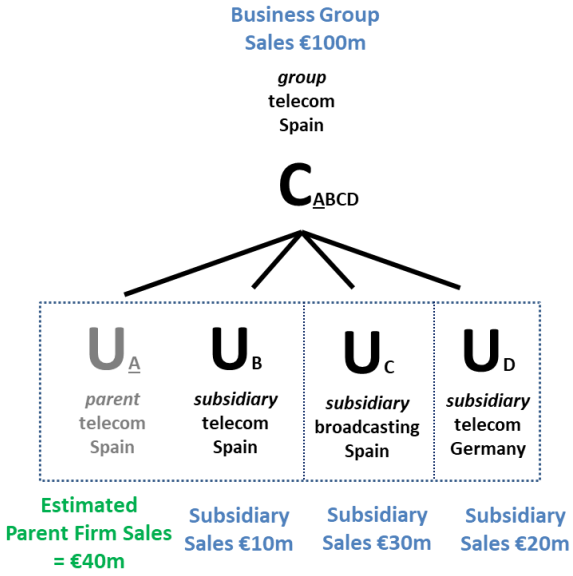
42. Our apportioning approach prevents allocation of the entire business group sales to the headquarters, even in cases where the parent only reports business group-level information. To continue the example of a Spanish telecoms business group, see Figure 3. Suppose the business group has total revenues of €100 million, of which €20 million is accounted for by a German telecoms subsidiary, €30 million reflects sales of a Spanish broadcasting subsidiary and €10 million of a Spanish telecoms subsidiary. In this case, importantly the parent is missing firm-level ("unconsolidated") information available. An approach not using information on the group structure would either allocate the total of €100 million to the Spanish telecoms parent, or double-count both €100 million of the parent and €60 million of the subsidiaries (recall Approach 2 and 3 respectively in the prior section). In contrast, the "apportioning approach" allocates the subsidiary sales to the subsidiary industries and only the rest (€40 million in this case) to the parent. Here, we remove the double counting issue and also apportion the group activity across all the market where the firm operates - telecoms and broadcasting industries in Germany and Spain.[20]

43. Our approach is computationally and data intensive. Firstly, it requires comprehensive ownership information to determine which firms are part of the same business group. To this end, an extensive cleaning of the ownership module in Orbis has been undertaken based on complementary information in the Zephyr mergers &

[20] Note in the cases where we have firm-level information for all subsidiaries including the parent, the aggregate sales may not always sum to the business group total. This is to be expected given intra-group transactions are removed from group-level ("consolidated") financial information. To ensure that changes in intra-group transaction over time do not drive the measured trends, whenever the total sales of all subsidiaries and the parent exceed the business group-level total, we scale them down to match that total. We examine the robustness of this decision later in the paper.

acquisitions database and manual checks of large firms. The corrected ownership data has been constructed for 2.8 million firms each year, further details are available in the Appendix. Secondly, it requires consistent financial information at the business-group and firm-level, across a broad set of firms, countries and industries. We combine Orbis and Worldscope data for 100 countries across all business sector industries from 2000 to 2014, and supplement this with manual checks for some of the largest business groups. One caveat is our data currently has poorer coverage of subsidiary-level information (but good business group-level information) for US firms – which is something we wish to improve in future work. Again, details of the financial data construction are detailed in the Appendix.

Figure 3. Apportioning when Parent Company Firm-Level Information is Not Available



Notes: The figure depicts an example of a hypothetical business group with a Spanish and German subsidiary in the same sector and a Spanish subsidiary operating in broadcasting. Information is available at the business group level ("consolidated) ($C_{\underline{A}BCD}$) alongside information for each individual subsidiary information ("unconsolidated") (U_B, U_C, U_D). In contrast to earlier examples, we are missing parent firm-level ("unconsolidated") information. Our "apportioning" approach aggregates revenues across firms operating in the same market but not across firms operating in different markets and attributes the balance to the parent firm. In this case, the business group has sales of €100m, with total subsidiary sales of €60m (=€10m+€30m+€20m), so we estimate parent firm sales in Spanish Telecom to be €40m (=€100m – €60m).

4. Firm-level concentration measures with MultiProd data

44. In this section, we present analysis for European economies based on the MultiProd project, which calculates statistics from nationally representative microdata. Accordingly, the concentration metrics in this section are measured at the firm-level, contrasting with business-group level metrics in the next section. See more details about the project in the Appendix and in Berlingieri et al. (2017a).

4.1. Sample

45. We use countries for which data are fully representative of the population of firms. Countries included and period covered are as follows: Austria (2008-2012), Belgium

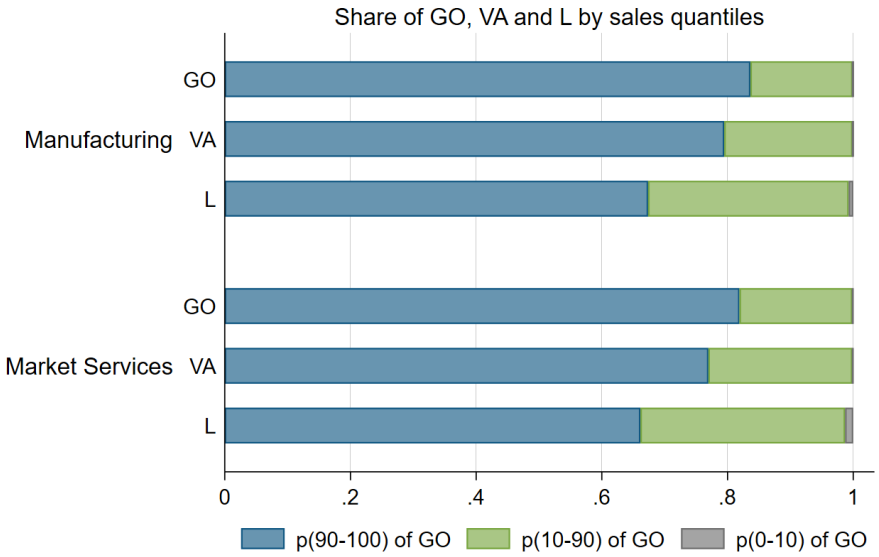
(2003-2011), Germany (2003-2012), Denmark (2001-2012), Finland (2001-2012), France (2001-2012), Hungary (2001-2012), Norway (2001-2012), Portugal (2004-2012), Sweden (2002-2012).

46. The MultiProd database generally covers most sectors of the economy, but this study focuses on manufacturing and non-financial market services to enhance cross-country comparability. Manufacture of Coke and Refined Petroleum, and Real Estate are excluded from the analysis (also for consistency with later Orbis analyses). Macro-sectors are defined according to the STAN A7 classification and detailed industries follow the STAN A38 classification.

4.2. Results

47. Both manufacturing and services look equally concentrated in terms of the share of gross output, value added and total employment accounted for by firms with the largest sales (see Figure 4). In particular, we plot the mean gross output (GO), value added (VA) and employment (L), for the bottom 10%, the top 10% and the 10th to 90th percentile of the sales distribution (averaged over all the available European countries and years). The figure shows that about 83% (82%) of the gross output in manufacturing (services) is produced by the firms in the top decile of sales. The rest of the distribution accounts for less than 20% of the gross output, with the bottom decile producing a negligible share. Very similar results are obtained when looking at valued added: about 79% (77%) of the value added in manufacturing (services) is produced by the firms in the top decile of sales, with the rest of the distribution accounting for slightly more than 20% of the value added. Employment is slightly less concentrated: the firms in the top decile of sales employ 67% (66%) of total employment in manufacturing (services). Both manufacturing and services look equally concentrated.

Figure 4. Gross output, value added and employment shares in European countries by quantiles of sales



*Notes: C*ountries included are AUT, BEL, DEU, DNK, FIN, FRA, HUN, NOR, PRT, SWE. The period considered is 2001-2012. Averages across all countries and years.

48. We next explore if concentration has increased over time within each country-2 digit (STAN A38) sector. We adopt an econometric approach that allows presenting trends within countries and disaggregated industries[21]. These trends, computed separately for manufacturing and services, accurately capture the evolution of concentration within each 2-digit sector in each country. The first year is taken as a baseline and normalised to zero, so that the coefficients are interpreted as the average change within country-sector of Gross Output or Employment, relative to the first year. In other words, these reflect cumulative changes since the first year.

49. We present results that account for the industry composition of the economy, to reflect representative trends for manufacturing and market services in each country. We present results of both unweighted and weighted regressions, with weights corresponding to the share of Gross Output of the industry in the total Gross Output of manufacturing or market services (as applicable).

50.

51. Figure 5 and Figure 6 show that concentration seems to have increased in the European countries considered over the period 2001-2012. In particular, in manufacturing, the share of gross output in the top decile of the sales is around 2 percentage points higher in 2012 than in 2001 (see Figure 5). This is true irrespective of whether or not the results are weighted by the total sales of each industry. The increase in non-financial market services is of similar magnitude when all industries are given an equal weight; it is closer to 3 percentage points when a greater weight is assigned to industries with greater aggregate sales, indicating that some large industries are among those which saw a higher increase in concentration. Looking at concentration in terms of employment, we find that in services it has increased by about 1.5 percentage points (over 2 percentage points when weighting for industry size), but, after an initial increase between 2001 and 2004, it has remained essentially constant in manufacturing and even appear slightly declining when we apply weighting (see Figure 6). The two take-away messages therefore are that (1) sales concentration has increased more than employment concentration, and (2) concentration has increased more in non-financial market services than in manufacturing.

52. These results are consistent with what was found using alternative measures of concentration in Andrews et al., (2016) using the commercial database Orbis for 24 countries and Autor et al. (2017a,b) for the US. They find a negative correlation between the labour share and concentration, if the latter is expressed in terms of sales. If, instead, concentration is calculated in terms of employment the correlation becomes positive. This result might be explained by the fact that canonical superstar (service) firms such as Google or Facebook employ relatively few workers compared to their revenues. In other words, superstar firms grow "without mass", that is, mostly in terms of revenues rather than employment. An additional result confirming that firms are experiencing "scale without mass" is found in Berlingieri et al. (2018a,b), where we find a rather flat link between

[21] This econometric estimation of trends relies on the estimation of year dummies, controlling for country-industry fixed effects. More formally, to analyse the evolution over time of a variable, say concentration C, we estimate the following regression:

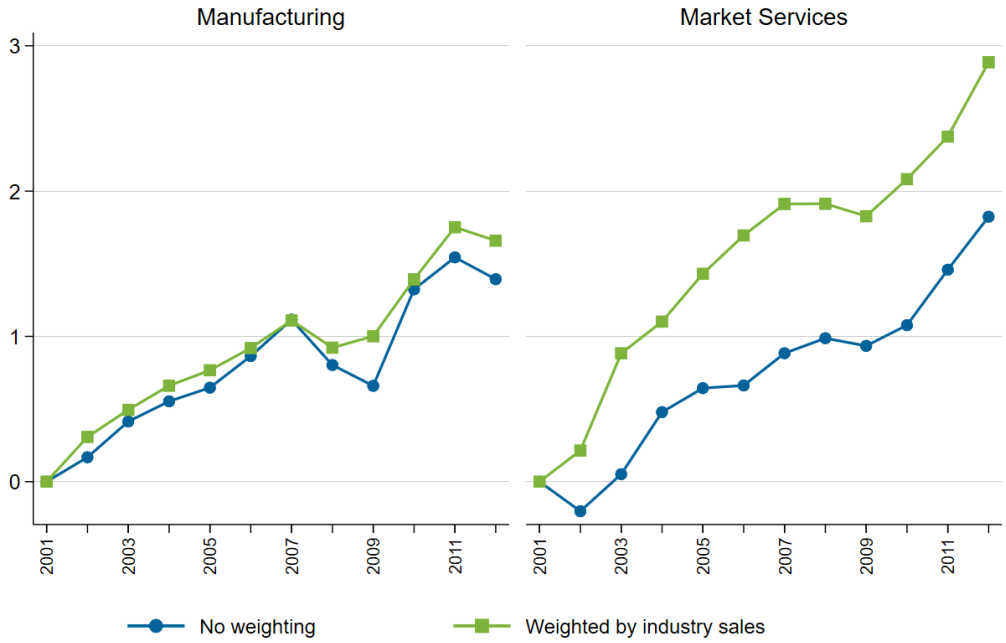
$$C_{cjt} = \alpha + \beta_t y_t + z_{cj} + \varepsilon_{cjt},$$

where z_{cj} correspond to country c-industry j fixed effects. Thus, the coefficients β_t associated with the year dummies y_t capture the average concentration in a given year already cleaned from country-sector specificities.

wages and size expressed in terms of employment, but a tight and positive link of wages with both productivity and size expressed in terms of sales.

Figure 5. Share of gross output accounted for by largest firms in select European countries

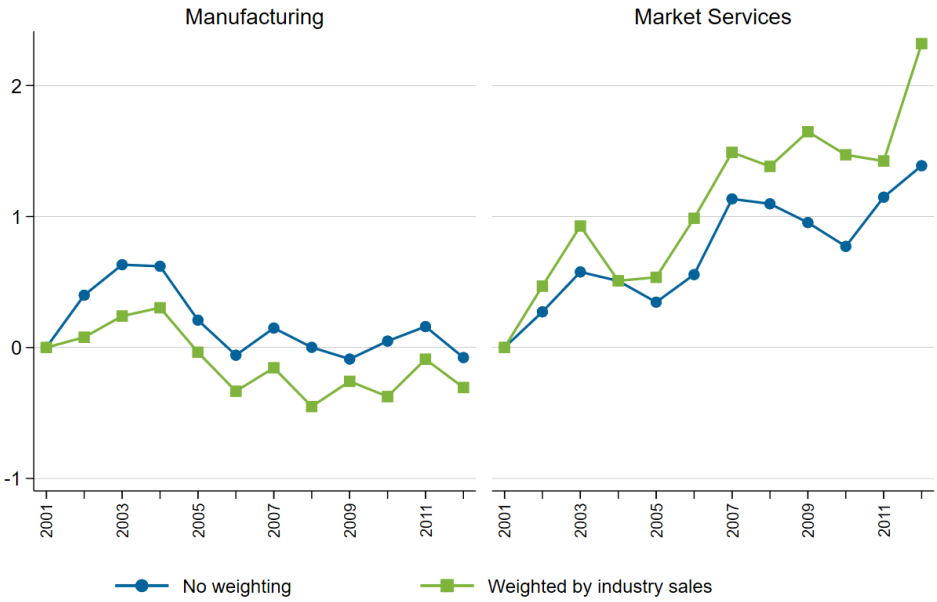
Top-10% firms in terms of sales, year effects from weighted regression



Note: Countries included are AUT, BEL, DEU, DNK, FIN, FRA, HUN, NOR, PRT, SWE. Year effects from regressions including country-industry and year dummies. The graphs can be interpreted as the cumulated growth rates of concentration in gross output within each country and 2-digit sector over the period. For instance, in 2012 non-financial market services the share of gross output accounted by the top decile of the sales is roughly 3 percentage points higher than in 2001. The estimates reported in the graph are those of year dummies in a cross-country regression of share of gross output in the top decile of the sales distribution.
Source: MultiProd 2017.

Figure 6: Share of employment accounted for by largest firms in select European countries

Top-10% firms in terms of sales, year effects from weighted regression



Note: countries included are AUT, BEL, DEU, DNK, FIN, FRA, HUN, NOR, PRT, SWE. Year effects from regressions including country-industry and year dummies. The graphs can be interpreted as the cumulated growth rates of concentration in employment within each country and 2-digit sector over the period. For instance, in 2012 non-financial market services the share of employment accounted by the top decile of the sales is roughly 2 percentage points higher than in 2001. The estimates reported in the graph are those of year dummies in a cross-country regression of share of employment in the top decile of the sales distribution.
Source: MultiProd 2017.

5. Business group-level concentration measures with Orbis data

53. The MultiProd measures described in the previous section reflect concentration within each country and industry. The underlying MultiProd microdata consist of observations at the level of individual firms, but these firms may not be independent and may be part of the same business group. Neglecting these firm ownership linkages may lead to an understatement of concentration. In this section, we complement the previous analysis by using Orbis data, which allows us to measure concentration at the business-group level – explicitly taking into account these firm ownership linkages.

5.1. Orbis-Worldscope-Zephyr data

54. The financial data in this section primarily come from Orbis database complemented with Worldscope data on listed firms. We undertake a number of cleaning steps on the financial data in Orbis and Worldscope and, to improve coverage for listed firms particularly in North America, we combine the two data sources using firm identifiers.[22] Identifying business groups requires comprehensive firm ownership data.

[22] We use the International Securities Identification Number to match firms in Orbis and Worldscope.

Accordingly, we have supplemented Orbis data on business group owners (global ultimate owners) and immediate ownership linkages with Zephyr data on Mergers and Acquisitions and a series of manual and automated checks. In both financial and ownership data, the checks included extensive cleaning steps, through a series of automatic and manual checks of the largest business groups in each industry and world region against their financial statements. See Annex A for more information.

55. We use the combined data to measure concentration at the level of 2-digit NACE rev.2 industries and of world regions.[23] For reasons of data coverage and comparability, we focus on manufacturing and non-financial market services (excluding Manufacture of coke and refined petroleum products, Real estate activities and Activities of head offices; management consultancy activities, with the former two also excluded from earlier MultiProd analysis).[24]

56. We include the years 2000-2014 for 21 economies. The metrics for Europe are based on Belgium, Denmark, Estonia, Finland, France, Germany, Greece, Hungary, Ireland, Italy, Latvia, the Netherlands, Norway, Poland, Portugal, Slovenia, Spain, Sweden and the UK, and those for North America reflect the US and Canada.

57. Despite the effort so far, further data cleaning is still under way, so the results below come with some caveats. One particular caveat is that our data currently have poorer coverage of subsidiary-level information for North American firms, although this is not the case for North American business-group-level information (Orbis and Worldscope together cover close to the universe of listed firms). This means that all domestic activity of these business groups may be allocated to the industry of the parent firm, rather than spread across domestic subsidiaries. In North America, this likely leads to higher measured concentration in levels and may also lead to more pronounced absolute changes in concentration. This limitation is something we wish to address in the future.

5.2. Results

58. We examine if concentration has increased over time within each region-2 digit sector. We adopt an econometric approach that allows presenting trends within regions and disaggregated industries. This is similar to that noted earlier for MultiProd analysis, but defined at the level of a region, rather than a particular economy. Unless otherwise noted, we show below cumulative average *absolute* changes in industry concentration levels, with the average calculated across industries in each region and year. These are normalised to zero in the first period. For some analyses, however, we show cumulative average *proportional* changes in industry concentration, which are again normalised to zero in the first year. Compared to the former approach, the latter approach effectively assigns more weight to concentration trends in industries that start at low levels of concentration initially.

5.2.1. Overall concentration trends in Europe and North America

59. We find a noticeable increase in industry concentration for both Europe and North America over the period (see Figure 7). Over the period 2000-2014, 77% of 2-digit industries in Europe and 74% in North America saw their concentration increase. Over the

[23] We plan to present results at an economy-level in a future work.

[24] We also exclude sectors where, in some years, we observe fewer than 20 business groups in one of the regions. This is the case for 5 sectors in North America: Repair of machinery & equipment; Water transport; Warehousing; Postal and courier activities; and Travel agency. For better comparability, we drop these sectors for both world regions.

period, concentration levels increased by 4 percentage points in the average European industry, compared to around 8 percentage points in the average North American industry. This means that the top 8 firms account for 4% (8%) more industry sales in 2014 than 2000 in Europe (North America). At first glance, this may suggest there is a stronger increase in North America. However, these increases are somewhat stronger than Autor et al. (2017b), who find an increase of around 4% for the top 4 or 20 firms since 2000 (admittedly using firm-level rather than business group-level concentration), which suggests that this could very well be due to the data limitations for North America noted above.[25] In particular, North American firms have poor coverage of subsidiary financial information, which reduces our ability to apportion business group activity when they have large (missing) domestic subsidiaries.[26]

60. We find a similar picture when we weight the results by total industry sales as reported in STAN data, effectively focusing more on concentration trends in larger industries (again see Figure 7). For Europe, weighting strengthens the growth in concentration by about 0.5 percentage points, implying that larger industries are showing somewhat faster increases in concentration. For North America, we find the opposite – a slightly lower increase in the weighted mean concentration.

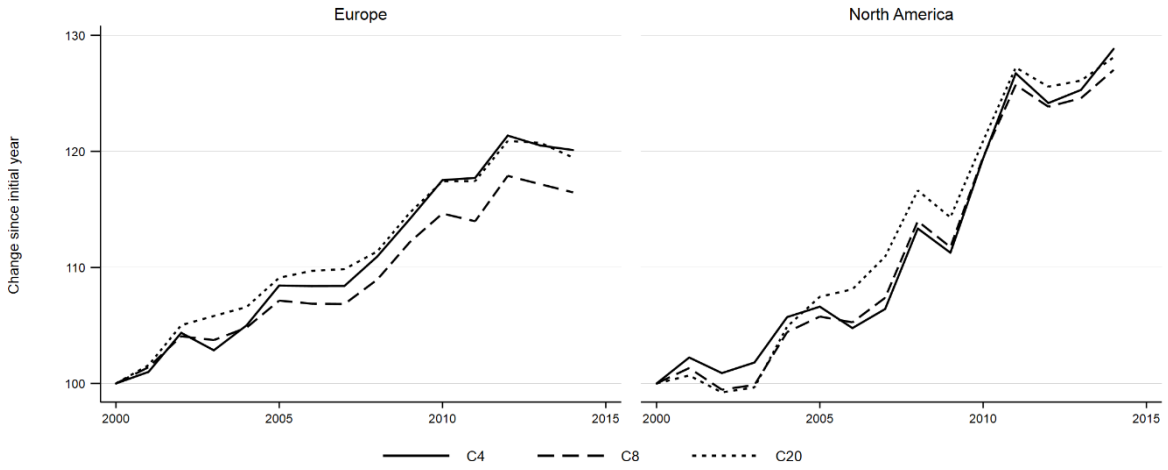
Figure 7. Weighted & Unweighted Industry Concentration (CR8) in Europe & North America

Note: The countries for Europe include BE, DE, DK, EE, ES, FI, FR, GB, GR, HU, IE, IT, LV, NL, NO, PL, PT, SI and SE, and the countries for North America include CA and US. Included industries cover 2-digit manufacturing and non-financial market services. Concentration metrics reflect the share of the top 8 firms in each industry (CR8). The graphs can be interpreted as the cumulated *absolute* changes in levels of sales concentration for the mean 2-digit sector within each region. For instance, in 2014 the mean (unweighted) European industry had 4 percentage point higher sales concentration than in 2000. Weighted figures reflect industry concentration weighted by the industry's share in region total sales, such that larger industries are weighted more heavily.

[25] This difference is not driven by the fact that we analyse North America as a single market while Autor et al. (2017b) focus only on the United States. Excluding Canadian firms does not significantly alter the results.

[26] We generally have better coverage for foreign subsidiaries of North American business groups.

Figure 8. Differing Concentration Metrics (CR4, CR8, CR20) in Europe & North America



Note: The countries for Europe include BE, DE, DK, EE, ES, FI, FR, GB, GR, HU, IE, IT, LV, NL, NO, PL, PT, SI, SE, and for North America include CA and US. Included industries cover 2-digit manufacturing and non-financial market services. Concentration metrics reflect the share of the top 4, top 8 and top 20 firms in each industry – unweighted metrics (CR4, 8 and 20 respectively). To ensure comparability across different metrics, these now reflect *proportional* changes. The graphs can be interpreted as the cumulated *percentage* changes in levels of sales concentration for the mean 2-digit sector within each region. For instance, in 2014 the mean European industry had 20% higher CR4 sales concentration compared to 2000.

61. Our findings are remarkably robust across concentration metrics with varying definition of "top" firms (see Note: The countries for Europe include BE, DE, DK, EE, ES, FI, FR, GB, GR, HU, IE, IT, LV, NL, NO, PL, PT, SI and SE, and the countries for North America include CA and US. Included industries cover 2-digit manufacturing and non-financial market services. Concentration metrics reflect the share of the top 8 firms in each industry (CR8). The graphs can be interpreted as the cumulated *absolute* changes in levels of sales concentration for the mean 2-digit sector within each region. For instance, in 2014 the mean (unweighted) European industry had 4 percentage point higher sales concentration than in 2000. Weighted figures reflect industry concentration weighted by the industry's share in region total sales, such that larger industries are weighted more heavily.

62. Figure 8). We show concentration metrics calculated using the sales of the top 4, top 8 and top 20 firms in each 2-digit region-industry (CR4, CR8 and CR20 respectively). Clearly, the market share of the top 20 firms will always be larger than the top 4. Therefore, to ensure comparability across the different metrics, we show *proportional* changes in concentration, relative to the initial year, rather than the absolute changes. For Europe, the proportional concentration increase when 4 or 20 largest firms are considered is about 20%, slightly more than the 16% proportional increase in the CR8 metric that we use as a baseline. For North America, all measures indicate a similar increase in industry concentration of around 28%. The results paint a similar picture as Autor et al. (2017b), who find similar changes in the CR4 and CR20 measures for many US industries.

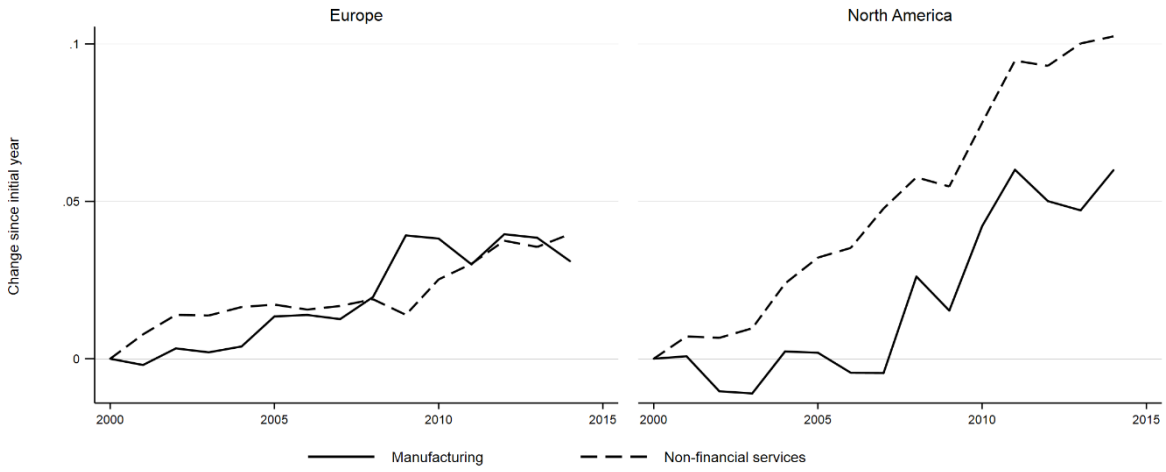
5.2.2. Sectoral concentration trends in Europe and North America

63. We find a similarly strong increase in concentration in manufacturing and services in Europe and a stronger increase in services in North America (see Figure 9). In Europe,

both in the average manufacturing industry and the average services industry the top 8 firms account for about 4% more industry sales in 2014 than the top 8 in 2000. These estimates compare reasonably closely with MultiProd firm-level (sales) concentration measures in the earlier sections, which showed an increase of around 3% for services and 2% for manufacturing. In North America, the increase in concentration is more similar to Europe for manufacturing (6 percentage points), with the higher overall increase in concentration being mainly due to services industries, where the share of sales by the top 8 firms increased by 10 percentage points.

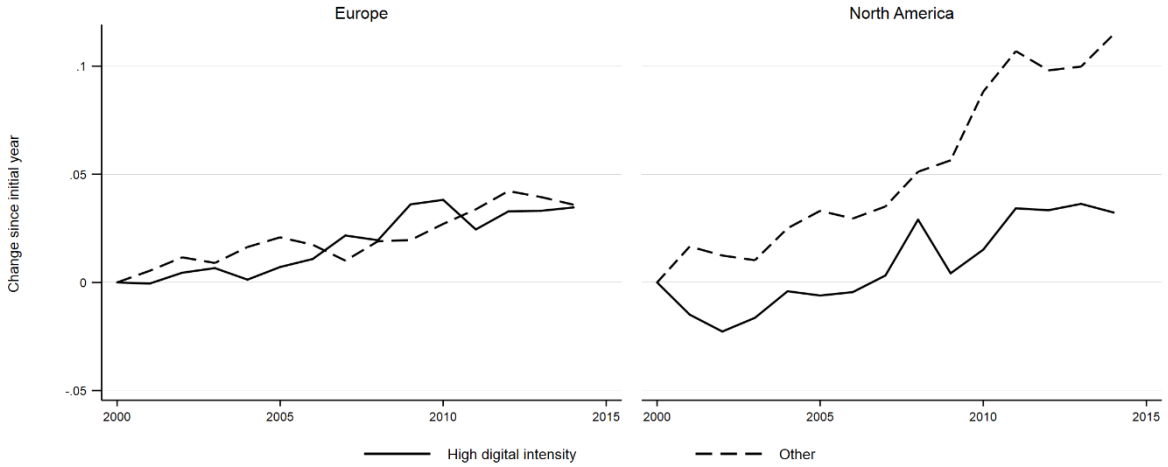
64. The increase in concentration does not seem to be driven by digital-intensive sectors in particular (Figure 10). We define digital sectors as the top quartile of digital intensity as measured by Calvino et al., 2017). In Europe, we see similar trends for both digital-intensive and less-digital sectors. In North America, the increase in concentration seems to actually be somewhat stronger in the less-digital sectors. For digital intensive industries we find initially *flat* concentration in the early 2000s, which could be seen as a period when many new internet-based start-up technologies started to become mainstream. This also mirrors findings for the US where there is a noticeable growth in high-tech business dynamism in the early late 1990s and early 2000s (Decker et al., 2016). From the mid-2000s onwards, we find an upward trend in concentration also for digital industries.

Figure 9. Concentration for Manufacturing vs Services in Europe & North America



Note: The countries for Europe include BE, DE, DK, EE, ES, FI, FR, GB, GR, HU, IE, IT, LV, NL, NO, PL, PT, SI, SE, and for North America include CA and US. Included industries cover 2-digit manufacturing and non-financial market services. Concentration metrics reflect the share of the top 8 firms in each industry (CR8). The graphs can be interpreted as the cumulated *absolute* changes in levels of sales concentration for the mean 2-digit sector within each region. For instance, in 2014 the mean European services industry had 4 percentage point higher sales concentration than in 2000.

Figure 10. Concentration in Digital-Intensive vs Less Digital Industries in Europe & North America



Note: The countries for Europe include BE, DE, DK, EE, ES, FI, FR, GB, GR, HU, IE, IT, LV, NL, NO, PL, PT, SI and SE, and the countries for North America include CA and US. Included industries cover 2-digit manufacturing and non-financial market services. Concentration metrics reflect the share of the top 8 firms in each industry (CR8). The graphs can be interpreted as the cumulated *absolute* changes in levels of sales concentration for the mean 2-digit sector within each region. For instance, in 2014 the mean European high digital intensity industry had 4 percentage point higher sales concentration than in 2000. The digital intensity of sectors is defined using the STAN A38 global digital intensity indicator of 2013-15 constructed by (Calvino et al., 2017).

5.2.3. Impact of key concentration measurement choices

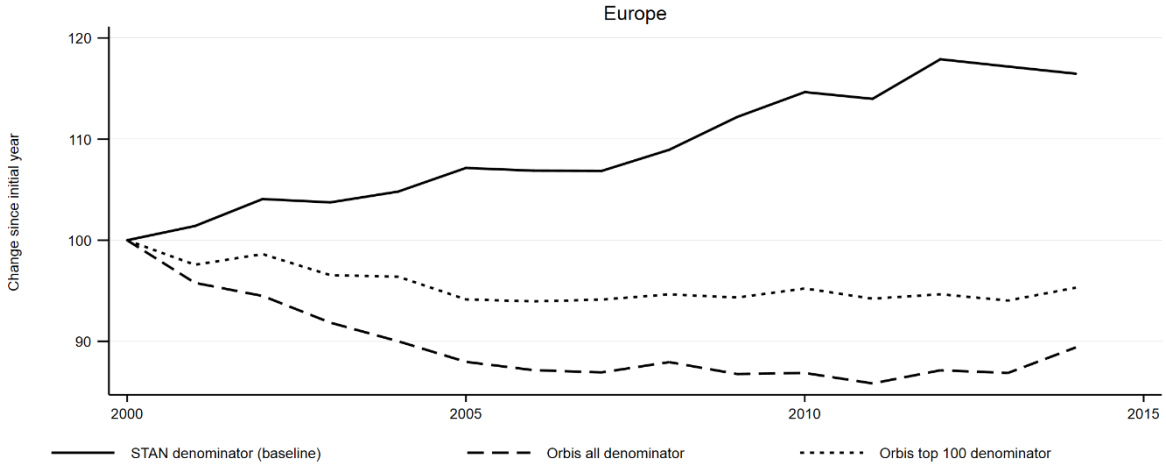
65. This final section briefly shows how different methodological and data cleaning choices influence the measured concentration trends. This section follows closely from the earlier methodological discussion in Section 3.

66. Industry concentration is essentially a ratio. We first focus on key decisions relating to the denominator, before considering the numerator below.

67. The choice of denominator has a striking effect on measured industry concentration trends (Figure 11). Our baseline analysis uses STAN measures of total industry sales for reasons of consistency and completeness. Alternative approaches in the literature have been to use the total sales of all firms in raw Orbis data or the top 100 firms in an industry in the raw Orbis data. Both are vulnerable to well-known Orbis coverage issues, as noted earlier. We compare the concentration trends obtained using the different denominators, looking at proportional changes in concentration, as in the case of comparing CR4, CR8 and CR20 metrics above. We find that scaling our baseline numerator with either denominator from Orbis completely reverses measured industry concentration trends. Concentration is found to fall over time. Our definition of denominator mitigates many of these problems, since OECD STAN is comparable across industries and countries over time.[27]

[27] Another potential concern, mainly in the manufacturing sector, could be that the concentration as measured here does not take consider the fact that exporters based in other regions also sell in Europe and North America, and European and North American firms sell some of their output in other

Figure 11. Impact of Industry Denominator Choice



Note: The countries for Europe include BE, DE, DK, EE, ES, FI, FR, GB, GR, HU, IE, IT, LV, NL, NO, PL, PT, SI, SE. Included industries cover 2-digit manufacturing and non-financial market services. Concentration metrics reflect the share of the top 8 firms in each industry (CR8). The STAN denominator reflects our baseline preferred approach. The Orbis all denominator reflects the total sales of all firms in that industry in Orbis. The Orbis top 100 denominator reflects the total sales of the top 100 firms in that industry in Orbis. To ensure comparability across measures based on different denominators, the graphs reflect *proportional* changes. The graphs can be interpreted as the cumulated *percentage* changes in levels of sales concentration for the mean 2-digit sector within each region. For instance, based on the STAN denominator, the mean European industry had in 2014 16% higher sales concentration compared to 2000.

68. We now turn to methodological choices regarding the numerator, which reflects the sales of the top 8 entities in each industry.

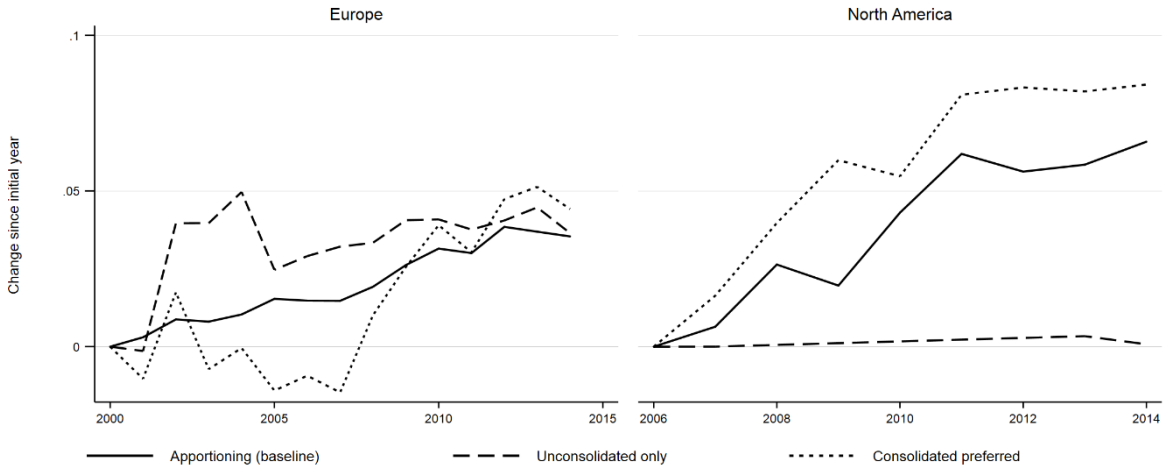
69. The choice between firm-level, group-level and apportioned analysis can lead to significantly different results (Figure 12). The entities (among which the top 8 actors in each industry are identified) can be defined as firms, business groups or apportioned business group segments (baseline). For North American industry concentration, the choice of entity level at which the analysis is undertaken makes a large difference.[28] Using only subsidiary-level data (labelled "unconsolidated only") gives a nearly flat trend, because the coverage of subsidiary-level information for US firms is poor in Orbis. Including both business group and firm subsidiary data ("consolidated preferred") should not lead to substantial double-counting in the case of North America (again due to the poor coverage

regions. Unfortunately, properly accounting for this would require group-level information on imports to as well as exports from Europe and North America, neither of which is available to us. Note also that whether an adjustment for imports and exports is desirable ultimately depends on the reasons for which industry concentration is analysed in any given case. Taking into account imports and exports is more important when treating industry concentration as a proxy for the intensity of competition. In contrast, using industry output alone is more appropriate when considering the implications of concentration related to, for instance, inequality, monopsony or some firms being "too big to fail".

[28] For North America, we only illustrate the difference across methodologies from 2006 onwards. These differences would be much more pronounced if the earlier years were included, as a sharp increase in Orbis coverage for the United States in 2006 leads to an increase in concentration when Orbis is not complemented with information from Worldscope.

at the subsidiary level), but it incorrectly allocates the entire group activity to the single industry and country of the headquarters. Consequently, the resulting concentration measure increases more than the baseline. For Europe, all three approaches lead to a similar overall increase in industry concentration over the sample period, but their trajectory is different, with the increase being concentrated in the early 2000s when we use subsidiary-level information and in the second half of the sample period when we combine business group and firm subsidiary data.

Figure 12. Impact of the Account Selection

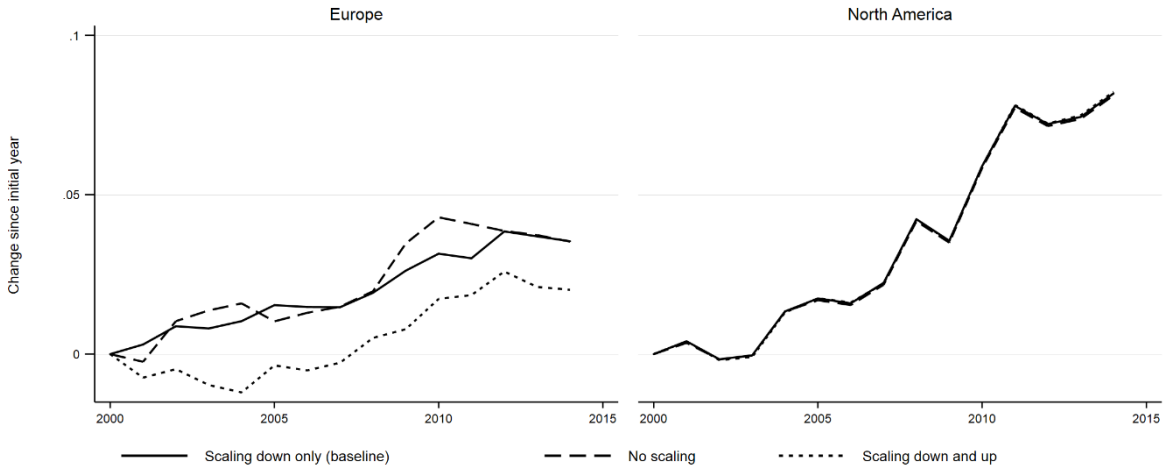


Note: The countries for Europe include BE, DE, DK, EE, ES, FI, FR, GB, GR, HU, IE, IT, LV, NL, NO, PL, PT, SI and SE, and the countries for North America include CA and US. Included industries cover 2-digit manufacturing and non-financial market services. Concentration metrics reflect the share of the top 8 firms in each industry (CR8). The graphs can be interpreted as the cumulated *absolute* changes in levels of sales concentration for the mean 2-digit sector within each region. Using the apportioning (baseline) approach, the mean North American industry had 7 percentage point higher sales concentration in 2014 compared to 2006. The baseline case apportions group output based on corrected Orbis-Worldscope-Zephyr data. "Unconsolidated only" represents the simplest approach, neglect business groups and focus only on individual firms. "Consolidated preferred" includes both the activity of the business group and other firm subsidiaries.

70. Another choice surrounds the treatment of cases where the sales for subsidiaries do not sum to the total of the parent (as mentioned in the methodology section). To some extent this is expected, since intra-group sales are a feature of multinationals in general, with one part of the firm purchasing inputs from another part. These intra-company transactions are removed from the group accounts but show up in each subsidiary's sales. To take account of these cases, in our baseline specification we scale down the sales of all subsidiaries when subsidiary sales sum to more than the group. Additionally, it would also be possible to scale up subsidiary sales where they sum to less than the group, which might reflect missing subsidiary data. For North America, the scaling has very little effect (Figure 13). This is again due to the very limited availability of unconsolidated information for US firms in Orbis. For Europe, all three approaches to scaling show an increase in concentration over the analysed period, but with some differences in its magnitude. When we do not scale down the output of business groups with a large intensity of intra-group sales, the average concentration is slightly more volatile but the overall increase over the sample period is very similar as with the baseline approach. When we also scale up subsidiary sales whenever they add up to less than the group sales, the concentration remains flat during

early 2000s but shows a similar trend as with the baseline approach from 2005 onwards. Such scaling up implicitly assumes that the missing data is distributed in the same industries and regions as the sample of subsidiary sales we observe. In contrast, not scaling up implicitly assumes that the missing subsidiary data is located in regions and industries outside the sample. Given that our sample consists of countries with relatively good and consistent coverage in Orbis, we think that the latter assumption will more often hold in our data and consequently opt for not scaling up as our baseline approach.

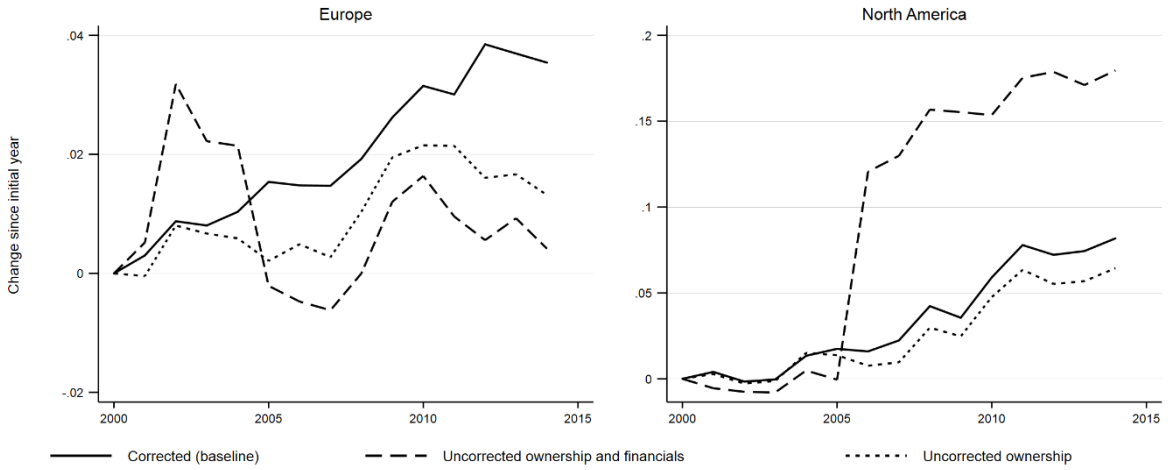
Figure 13. Impact of Scaling Accounts to Match Group Consolidated Accounts



Note: The countries for Europe include BE, DE, DK, EE, ES, FI, FR, GB, GR, HU, IE, IT, LV, NL, NO, PL, PT, SI and SE, and the countries for North America include CA and US. Included industries cover 2-digit manufacturing and non-financial market services. Concentration metrics reflect the share of the top 8 firms in each industry (CR8). The graphs can be interpreted as the cumulated *absolute* changes in levels of sales concentration for the mean 2-digit sector within each region. For instance, using the baseline approach, the mean North American industry had 8 percentage point higher sales concentration in 2014 compared to 2000. The baseline case scales subsidiary sales down if needed to ensure that they do not exceed the group consolidated sales. "No scaling" represents the approach without such correction. "Scaling down and up" involves scaling subsidiary sales in either direction to match group consolidated sales.

71. Finally, the data cleaning undertaken has a large effect on results for both regions (Figure 14). We undertake extensive cleaning of the Orbis financial and ownership data, consisting of a large set of automatic and manual corrections based on additional data and internal consistency checks. In the case of Europe, comparing results using the cleaned and uncleaned data highlights the high demands of the apportioning approach on data quality. Uncorrected data show large fluctuations in concentration, with little overall trend. Cleaning the financial information for the largest firms and complementing it from other sources (e.g. Worldscope, annual reports) leads to less volatility and some upward trend. However, the overall increase in concentration more than doubles when we also correct the ownership information based on Zephyr and manual checks of the largest business groups in our data. For North America, the essential cleaning step involved complementing Orbis financials with Worldscope. In its absence, the data show a spurious increase in industry concentration in 2006, due to a significant increase in Orbis coverage for North America.

Figure 14. Impact of Data Cleaning



Note: The countries for Europe include BE, DE, DK, EE, ES, FI, FR, GB, GR, HU, IE, IT, LV, NL, NO, PL, PT, SI and SE, and the countries for North America include CA and US. Included industries cover 2-digit manufacturing and non-financial market services. Concentration metrics reflect the share of the top 8 firms in each industry (CR8). The graphs can be interpreted as the cumulated *absolute* changes in levels of sales concentration for the mean 2-digit sector within each region. For instance, using the apportioning (baseline) approach, the mean North American industry had 8 percentage point higher sales concentration in 2014 compared to 2000. The baseline case is based on the corrected Orbis ownership and Orbis-Worldscope financial information. "Uncorrected ownership and financials" is based on the original ownership and financial information in Orbis. "Uncorrected ownership" is based on the corrected Orbis-Worldscope financial information but on original Orbis ownership data.

6. Next steps

72. The research outlined in this paper is still in its first phase. The purpose of the paper is twofold: to present evidence filling a gap in the current knowledge and to gain feedback on the methodology employed.

73. This paper presents evidence on an increase in industry concentration in both Europe and in North America over the period 2000-2014. The increase is observed in representative firm-level data from OECD MultiProd as well as in business group-level matched and extensively cleaned Orbis-Worldscope data. It holds for both manufacturing and non-financial market services. For digital sectors, we find initially flat concentration in the early 2000s and a similar growth rate as less-digital sectors thereafter.

74. There are several avenues open for future work building on these findings.

75. The most immediate extension of the work presented here is related to the concentration measures computed with the Orbis-Worldscope data. We are going to compute country-level concentration measures, in addition to the regional ones presented here. Additionally, we also plan to incorporate industry imports into our concentration measures to ensure that they only reflect changes in domestic market size. Along similar lines, we could extend our apportioning approach to take into account exports. Furthermore, we could examine possible differential trends across 2-digit industries and possibly include some of the sectors not presented in this paper, such as financial services or utilities (currently dropped for data quality issues). Going beyond concentration measures, we could explore the churning among the top firms in each industry, and the

extent to which the growth of the top firms is driven by changes in group structure or organic growth.

76. We also plan to examine more details of the mechanisms involved, which has links to work conducted under the Going Digital - Horizontal Project and work chosen by the Global Forum on Productivity, in particular work examining possible links between concentration trends, mergers and acquisition activity and the implications for productivity growth. Another related angle is to examine concentration in patenting activity, which might provide evidence for whether these technologies are diffusing widely throughout the economy or remain held by a few firms. Finally, we can econometrically analyse the relationship between the concentration trends and proxy variables for potential drivers of these trends, such as globalisation, technological changes and product-market regulations.

References

Aghion, P., Bloom, N., Blundell, R., Griffith, R., and P. Howitt (2005), "Competition and Innovation: an Inverted-U Relationship", *Quarterly Journal of Economics*, 120(2), 701-728.

Aghion, P., Blundell, R., Griffith, R., Howitt, P., and S. Prantl (2009), "The effects of entry on incumbent innovation and productivity", *Review of Economics and Statistics*, 91(1), 20-32.

Altomonte, C., Ottaviano, G. and A. Rungi (2018), "Business Groups as Knowledge-Based Hierarchies of Firms", *Mimeo*

Andrews, Dan, Chiara Criscuolo, and Peter N. Gal (2016), "The Best versus the Rest: The Global Productivity Slowdown, Divergence across Firms and the Role of Public Policy", *OECD Productivity Working Papers* 5, Paris: OECD Publishing. https://doi.org/10.1787/63629cc9-en.

Autor, David H., David Dorn, and Gordon H. Hanson (2013), "The China syndrome: Local Labor Market Effects of Import Competition in the United States", *American Economic Review*, Vol. 103, No. 6, pp. 2121–168.

Autor, David, David Dorn, Lawrence F. Katz, Christina Patterson, and John Van Reenen (2017a), "Concentrating on the Fall of the Labor Share", *American Economic Review Papers and Proceedings*, Vol. 107, No. 5, pp. 180–185.

Autor, D., D. Dorn, L. Katz, C. Patterson, and J. van Reenen (2017b), "The Fall of the Labor Share and the Rise of Superstar Firms", *Working Paper No. 23396*, National Bureau of Economic Research.

Bajgar, Matej, Giuseppe Berlingieri, Sara Calligaris, Chiara Criscuolo, and Jonathan Timmis. Forthcoming, "To Use or Not to Use (and How to Use): Coverage and Representativeness of Orbis Data", *OECD Science, Technology and Industry Working Papers*, OECD Publishing.

Bajgar, Matej, Sara Calligaris, and Chiara Criscuolo (2017), "Can Business Micro Data Match Macro Trends? Comparing Multiprod Data with STAN", Unpublished manuscript.

Barkail, S. (2017), *Profits, Redistribution of Income and Dynamic Efficiency*, Mimeo.

Bartelsman, E., S. Scarpetta, and F. Schivardi (2005), "Comparative Analysis of Firm Demographics and Survival: Evidence from Micro-level Sources in OECD Countries", *Industrial and Corporate Change*, Vol. 14, No. 3, pp. 365–391.

Bartelsman, E., J. Haltiwanger, and S. Scarpetta (2009), "Measuring and Analyzing Cross-country Differences in Firm Dynamics", in: *Producer Dynamics: New Evidence from Micro Data*, University of Chicago Press, pp. 15–76.

Berlingieri, G., P. Blanchenay, S. Calligaris, and C. Criscuolo (2017a), "The MultiProd Project: a Comprehensive Overview", *OECD Science, Technology and Industry Working Papers* No. 2017/04, OECD Publishing, Paris, DOI: https://doi.org/10.1787/2069b6a3-en.

Berlingieri, Giuseppe, Patrick Blanchenay, and Chiara Criscuolo, 2017b, "The Great Divergence(S)", 39, *OECD Science, Technology and Industry Policy Papers*, Paris: OECD Publishing.

Berlingieri, G., Calligaris, S., and Criscuolo, C. (2018a), "The productivity-wage premium: Does size still matter in a service economy?" *AEA Papers and Proceedings*, Vol. 108, pp. 328-33.

Berlingieri, G., S. Calligaris and C. Criscuolo (2018b), "The productivity-wage premium: Does size still matter in a service economy?", *OECD Science, Technology and Industry Working Papers*, No. 2018/13, OECD Publishing, Paris, https://doi.org/10.1787/04e36c29-en.

Bessen, J. (2016), "Accounting for Rising Corporate Profits: Intangibles or Regulatory Rents?", *Mimeo.*

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

Calligaris, S., C. Criscuolo and L. Marcolin (2018), "Mark-ups in the digital era", *OECD Science, Technology and Industry Working Papers*, No. 2018/10, OECD Publishing, Paris, https://doi.org/10.1787/4efe2d25-en

Conconi, P., Magerman, G., Plaku, A. and C. Thomas (2018), "Multinational Production Networks and Trade Participation", *Mimeo*.

Cravino, J. and A. Levchenko (2017), "Multinational Firms and International Business Cycle Transmission", *Quarterly Journal of Economics*, 132(2), pp. 921-962.

Criscuolo, C., Gal, P. and C. Menon (2014), "The Dynamics of Employment Growth: New Evidence from 18 Countries", *OECD Science, Technology and Industry Policy Papers*, No. 14, OECD Publishing, Paris. http://dx.doi.org/10.1787/5jz417hj6hg6-en.

Decker, R., Haltiwanger, J., Jarmin, R. and J. Miranda (2014), "The role of entrepreneurship in U.S. job creation and economic dynamism", *Journal of Economic Perspectives*, 28, 3-24.

Decker, R., Haltiwanger, J., Jarmin, R. and J. Miranda (2016), "Where Has All the Skewness Gone? The Decline in High-Growth (Young) Firms in the U.S.?", *European Economic Review*, 86, 4-23.

De Loecker, J. and J. Eeckhout (2017), "The Rise of Market Power and the Macroeconomic Implications", *NBER Working Paper No. 23687*.

Dellis, K. and D. Sondermann (2017), "Lobbying in Europe: new firm-level evidence", *ECB Working Paper 2071*, June 2017. https://www.ecb.europa.eu/pub/pdf/scpwps/ecb.wp.2071.en.pdf

Disney, R., Haskel, J., and Y. Heden (2003), "Restructuring and productivity growth in UK manufacturing", *The Economic Journal*, 113(489), 666-694.

Economist, The, (2016), "Business in America: Too Much of a Good Thing", Published in print on 26th March 2016, https://www.economist.com/briefing/2016/03/26/too-much-of-a-good-thing

Furman, Jason, and Peter Orszag (2015), "A Firm-Level Perspective on the Role of Rents in the Rise in Inequality", Paper presented at Columbia University's "A Just Society", Centennial Event in Honor of Joseph Stiglitz, New York, NY.

Gal, Peter N. (2013), "Measuring Total Factor Productivity at the Firm Level Using OECD-ORBIS", *OECD Economics Department Working Papers No 1049*, OECD Publishing. http://ideas.repec.org/p/oec/ecoaaa/1049-en.html.

Grullon, G., Larkin, Y. and R. Michaely, (2016), "Are US Industries Becoming More Concentrated?", *Mimeo*.

Gutiérrez, Germán, and Thomas Philippon (2017), "Declining Competition and Investment in the US", *NBER Working Paper No. 23583*.

Gutiérrez, Germán, and Thomas Philippon. 2018. "How EU Markets Became More Competitive Than US Markets: A Study of Institutional Drift", *NBER Working Paper No. 24700*.

Hall, R. (2018), "New Evidence on the Markup of Prices over Marginal Costs and the Role of Mega-Firms in the US Economy", *NBER Working Paper No. 24574*.

Haltiwanger, John. (2015), "Top Ten Signs of Declining Business Dynamism and Entrepreneurship in the U.S.", Paper presented at the Kauffman Foundation New Entrepreneurial Growth conference in June 2015. http://econweb.umd.edu/~haltiwan/Haltiwanger_Kauffman_Conference_August_1_2015.pdf

Honjo, Y., Y Doi, & Y Kudo, (2014), "Consideration on Changes of Market Structure of Major Industries in Japan by the Use of Mobility Index, and the Applicability to the Competition Policy: An Analysis Based on the degree of concentration of production and shipment," JFTC, *Competition Policy Research Center Report*, 2014.

Kalemli-Ozcan, S., B. Sorensen, C. Villegas-Sanchez, V. Volosovych and S. Yesiltas (2015), "How to Construct Nationally Representative Firm Level Data from the Orbis Global Database", *NBER Working Paper No. 215*58.

Karabarbounis, L. and N. Brent (2014), "The Global Decline of the Labor Share", *Quarterly Journal of Economics*, 129(1), 61-103.

OECD (2015), "The Labour Share in G20 Economies'. Report prepared for the G20 Employment Working Group", https://www.oecd.org/g20/topics/employment-and-social-policy/The-Labour-Share-in-G20-Economies.pdf.

OECD (2018a), "Monopsony and Buyer Power", proceedings of a Roundtable on Monopsony and Buyer Power held by the OECD Competition Committee in October 2008. https://www.oecd.org/daf/competition/44445750.pdf

OECD (2018b), "Market Concentration", issues paper for the hearing on market concentration taking place at the 129th meeting of the OECD Competition Committee.

Rossi-Hansberg, E., P-D Sarte, and N. Trachter, (2018), "Diverging Trends in National and Local Concentration", *NBER Working Paper No. 25066*.

Siegenthaler, M., and T. Stucki (2015), "Dividing the Pie: Firm-Level Determinants of the Labor Share", *ILR Review* 68 (5): 1157–94. https://doi.org/10.1177/0019793915593952.

Social Market Foundation (2017), "Concentration not competition: the state of UK consumer markets", http://www.smf.co.uk/wp-content/uploads/2017/10/Concentration-not-competition.pdf

Squicciarini, M., and H. Dernis (2013), "A Cross-Country Characterisation of the Patenting Behaviour of Firms Based on Matched Firm and Patent Data", 2013/5. *OECD Science, Technology and Industry Working Papers*, OECD Publishing, http://dx.doi.org/10.1787/18151965 .

Valletti, T., G. Koltay, S. Lorincz & H. Zenger (2017), "Concentration trends in Europe", presentation, December 2017, https://ecp.crai.com/wp-content/uploads/2017/12/VallettiConcentration_Trends_TV_CRA-002.pdf

Van Reenen, J., (2018), "Increasing Differences between firms: Market Power and the Macro-Economy"? *Mimeo*.

Annex A. Data

Overview of MultiProd data

1. The appendix provides a quick overview of the MultiProd dataset. Further detailed information on the MultiProd project and the methodology adopted can be found in Berlingieri et al. (2017a).

2. The analysis conducted in this note relies on the work undertaken in the last few years within the OECD "MultiProd" project. The implementation of the MultiProd project is based on a standardised STATA® routine that micro-aggregates micro-data from production surveys and business registers, via a distributed microdata analysis. This methodology was pioneered in the early 2000s in a series of cross-country projects on firm demographics and productivity (Bartelsman et al., 2005; Bartelsman et al., 2009). The OECD currently follows this approach in three ongoing projects: MultiProd, DynEmp, and MicroBeRD.29 The distributed micro-data analysis involves running a common code in a decentralised manner by representatives in national statistical agencies or experts in public institutions, who have access to the national micro-level data. At this stage, micro-aggregated data are generated by the centrally designed, but locally executed, program codes, which are then sent back for comparative cross-country analysis to the OECD.

3. The advantages of this novel data collection methodology are manifold: it puts a lower burden on national statistical agencies and limits running costs for such endeavours. Importantly, it also overcomes the confidentiality constraints of directly using national micro-level statistical database, while at the same time achieving a high degree of harmonisation and comparability across countries, sectors, and over time.

4. The MultiProd program relies on two main data sources in each country. First, administrative data or production surveys (PS), which contain all the variables needed for the analysis of productivity but may be limited to a sample of firms. Second, a business register (BR), which contains a more limited set of variables but for the entire population of firms. The program works also in the absence of a business register and this is not needed when administrative data on the full population of firms are available. However, when data come from a PS, its availability substantially improves the representativeness of results and, thus, their comparability across countries.

5. Census and administrative data, indeed, normally cover the whole population of businesses with at least one employee. Still, these datasets do not always exist and PS data need to be used. One of the big challenges of working with firm-level production surveys is that the selected sample of firms might yield a partial and biased picture of the economy. Whenever available, BRs, which typically contain the whole population of firms, are therefore used in MultiProd to compute a population structure by year-sector-size classes. This structure is then used to re-weight data contained in the PS in order to construct data

[29] MultiProd, DynEmp, and MicroBeRD are projects carried forward by the Directorate for Science, Technology and Innovation (STI) at the OECD. The DynEmp (Dynamics of Employment) project provides harmonised micro-aggregated data to analyse employment dynamics (find out more: http://www.oecd.org/sti/dynemp.htm) and MicroBeRD provides information on R&D activity in firms from official business R&D surveys (find out more: http://www.oecd.org/sti/rd-tax-stats.htm).

that are as representative as possible of the whole population of firms and comparable across countries.

6. At the time of writing, 20 countries have been successfully included in the MultiProd database (namely, Australia, Austria, Belgium, Canada, Chile, Denmark, Finland, France, Germany, Hungary, Indonesia, Italy, Japan, Luxembourg, Netherlands, Norway, New Zealand, Portugal, Switzerland and Sweden). For most countries the time period spans from early 2000s to 2012. For Chile, Austria and Switzerland the time horizon is shorter (starting in 2005, 2008 and 2009 respectively), whereas for Finland, France, Japan and Norway data are available at least since 1995.

7. MultiProd collects data for all sectors of the entire economy, whenever available. However, for the purposes of this analysis we have restricted our sample to the manufacturing and non-financial market services sector, as well as to European countries. Specifically, in this note we keep into account only European countries for which data are fully representative of the population of firms. The list includes Austria, Belgium, Denmark, Finland, France, Germany, Hungary, Norway, Portugal and Sweden.

Overview of Orbis-Worldscope-Zephyr data

8. We use the firm structure to apportion consolidated sales of the business group across subsidiaries. To do so requires reasonable coverage of both group and subsidiary financial information, as well as detailed ownership information detailing parent-subsidiary linkages. The group and subsidiary financial information is primarily sourced from Orbis, which is discussed in the first section below. We supplement Orbis with financial information from Worldscope for listed firms, which is discussed in the second section. The third section briefly described the combined Orbis-Worldscope financial data. The final section covers the ownership information, which is obtained from Orbis, the Zephyr database on Mergers and Acquisitions and for the largest firms, spot checks of the group structure listed in their financial statements.

Orbis financial data

9. Orbis is the largest cross-country firm-level database that is available and accessible for economic and financial research. It is a commercial database provided to the OECD by the electronic publishing firm Bureau Van Dijk.

10. To maximise the coverage of each firm's subsidiaries, we use information for firms in all business sectors within Orbis and across 100 economies. Large firms may have subsidiaries spread across a variety of industries and economies. Our apportioning approach relies on reasonable coverage of these subsidiaries, so we include a very broad range of industries and economies. Specifically, we include 2-digit NACE rev.2 codes 1-82 – which reflects industries excluding public services. We include firm data from 100 economies, encompassing all economies with at least 500 firm-year observations or 500 merger and acquisition deals.[30] This covers many more economies and industries than are

[30] For completeness, the 100 economies included in the firm subsidiary analysis are: Albania, Algeria, Angola, Argentina, Australia, Austria, Belarus, Belgium, Bermuda, Bosnia Herzegovina, Botswana, British Virgin Islands, Brazil, Bulgaria, Cameroon, Canada, Cayman Islands, Chile, China, Chinese Taipei, Colombia, Costa Rica, Croatia, Cuba, Cyprus, Czechia, Denmark, Ecuador, Egypt, Estonia, Ethiopia, Finland, France, Georgia, Germany, Greece, Hong Kong, Hungary, Iceland, India, Indonesia, Ireland, Israel, Italy, Japan, Jordan, Kazakhstan, Kenya, Kosovo, Kuwait,

often considered in Orbis analyses (see for instance, Andrews et al., 2016). The broad coverage allows a more complete picture of each firm's subsidiaries than focusing on a narrower subset of industries or countries.

11. The financial information primarily derives from company accounts, so some cleaning is required before using the data. We undertake a number of cleaning steps, closely following the suggestions by Kalemli-Ozcan, et al. (2015) and described in more detail in previous OECD analyses (Gal, 2013; Andrews et al., 2016). These include keeping accounts that refer to entire calendar year, dropping duplicate observations and removing outliers identified as implausible changes or ratios.

12. Financial information is available at different levels of aggregation within Orbis. Available information may be at the business group-level ("consolidated" or aggregated financial data consolidated across subsidiaries of the firm) and individual firm-level ("unconsolidated" information referring to an individual firm). As noted earlier, we use the firm structure to apportion consolidated group sales across their subsidiaries. To do so requires reasonable coverage of both business-group and individual subsidiaries' financial information.

13. We expand the coverage of Orbis data by using available unconsolidated information to infer missing consolidated information of the *same* firm and vice versa. Some firms have consolidated and unconsolidated information that is missing at different times. One concern may be that the trends in unconsolidated information is not reflective of trends of the wider consolidated accounts for these missing years. To ensure unconsolidated trends are a reasonable proxy, we require unconsolidated sales to be at least 50% of consolidated firm sales and therefore represent the bulk of firm activity (for the firm on average). Furthermore, we censor the unconsolidated trends at the 5^{th} and 95^{th} percentiles, so we do not use extreme growth of unconsolidated sales. Note we do not apply thresholds to using consolidated trends to fill in unconsolidated sales, since if unconsolidated sales are only a small proportion of consolidated sales any errors are unlikely to impact the group structure substantially.

14. We tested the accuracy of this approach by examining its ability to predict our actual Orbis observations, setting half our actual observations to missing. Specifically, we randomly set half the actual consolidated Orbis observations to missing that also have unconsolidated information. We then use the unconsolidated sales growth rates to fill in the missing numbers and see how close the filled in and actual observations are. We find that our approach closely replicates actual Orbis observations: we find a 95% correlation of between the new series (which is 50% fitted) and actual sales. Therefore, our conservative approach to combining Orbis consolidated and unconsolidated data appears to be appropriate.

Worldscope financial data

15. It is possible to expand the coverage of our data by combining Worldscope and Orbis data using firm International Securities Identification Number (ISIN) numbers,

Latvia, Lebanon, Liechtenstein, Lithuania, Luxembourg, Malaysia, Malta, Mauritius, Mexico, Montenegro, Morocco, Namibia, Netherlands, New Zealand, Nigeria, Norway, Pakistan, Peru, Philippines, Poland, Portugal, Qatar, Rep. of Korea, Rep. of Moldova, Romania, Russia, Saudi Arabia, Senegal, Serbia, Singapore, Slovakia, Slovenia, South Africa, Spain, Sri Lanka, Sweden, Switzerland, TFYR of Macedonia, Thailand, Tunisia, Turkey, Ukraine, United Arab Emirates, UK, Uruguay, USA, Uzbekistan, Viet Nam, Zambia and Zimbabwe.

which uniquely identify listed firms. We find that the majority of listed firms in Orbis are also in Worldscope (and vice versa). However, for some countries, such as the US, and some years, Worldscope is able to improve the coverage of the Orbis.

16. Worldscope is a cross-country firm-level commercial database of listed firms (Orbis also includes non-listed) provided to the OECD by Thomson Reuters. The firms in the data represent about 95% of global stock market capitalisation. Again, the financial information primarily derives from company accounts, so some cleaning is required before using the data. We apply the same cleaning rules to the Worldscope data as noted above for Orbis data. Worldscope reflects consolidated financial data (note Orbis also includes unconsolidated).

17. For the same firm-years that are both in Orbis and Worldscope it is possible to compare the consolidated information in Orbis to the Worldscope data. This gives an idea of how comparable the two data sources are and whether combining the two sources is appropriate.

18. We find that firm consolidated information in Orbis and Worldscope compare closely. Table A1 shows the distribution of differences between Orbis and Worldscope data. Most firms have very similar sales data in both Orbis and Worldscope, with 50% of firms having sales within approximately 5-7% in each data source and sales growth within 9-10% in each data (considering 25th and 75th percentiles in the table below). There are few firms with very different sales in each data: 90% of firms have sales within approximately 29-34% in each data, and sales growth within 31%-50%.

Table A.1. Comparability of Orbis and Worldscope Data

Firm-level ratio of sales in Worldscope to sales in Orbis for a given year

	1st Percentile	5th Percentile	25th Percentile	Median	75th Percentile	95th Percentile	99th Percentile
Sales	0.18	0.66	0.93	1.00	1.05	1.29	5.19
Sales Growth (1 Year)	0.22	0.69	0.91	1.01	1.10	1.50	5.25

Note: Comparison of sales (in current price Euros) for firm-year observations in both Worldscope and Orbis data. Ratios reflect Worldscope sales (growth) as a proportion of Orbis sales (growth).
Source: Thomson Reuters Worldscope and Bureau Van Dijk Orbis databases.

19. We supplement Orbis data with Worldscope in two ways: adding missing years for firms that are already in Orbis and adding missing firms that never have financial data in Orbis. The former involves using two data sources for the same firm at different points in time and so we take a more cautious approach, whereas the latter involves using only one. The first aspect of this conservative approach involves using sales growth rates (rather than levels) from Worldscope, to avoid jumps when the data source changes. The second aspect is to use growth rates only for firms that are sufficiently similar, with sales on average within 35% in the Worldscope and Orbis in the years when we observe them in both. This second aspect is designed to remove possible cases of matching the wrong firms in the two data sets.

20. We tested the accuracy of this approach by examining its ability to predict our actual Orbis observations, setting half our actual observations to missing. Specifically, we randomly set half the actual Orbis observations to missing that are also present in

Worldscope. We then use the sales growth rates in Worldscope to fill in the missing numbers and see how close the filled in and actual observations are. We find that our approach closely replicates actual Orbis observations: we find a 99% correlation of between the new series (which is 50% fitted) and actual sales. Therefore, our conservative approach to combining Worldscope and Orbis data appears to be appropriate.

21. Supplementing Orbis with Worldscope leads to improved coverage, in particular for North America in earlier years. Table A2 below summarises the number of firm-year observations gained by including Worldscope data. For all regions there are non-trivial increases in coverage (of listed firms in Worldscope), however, this is particularly the case for North America. Orbis is known to have poorer coverage for the US, and we find this is the case before 2007 (see Bajgar et al., forthcoming). Since these listed firms are often large, adding Worldscope therefore substantially improves the ability to measure concentration in North America.

Table A.2. Increased Coverage by Including Worldscope Data

	No. Additional Observations		
	North America	Europe	Other
2000	11,770	5,639	6,928
2002	11,206	4,667	4,712
2004	11,051	4,941	5,032
2006	9,859	6,591	5,716
2008	8,969	6,223	6,464
2010	8,800	5,208	5,675
2012	8,284	4,858	5,862
2014	3,312	4,424	4,960

Note: Comparison of the number of additional firm observations from supplementing Orbis with Worldscope data. North America and Europe defined consistent with earlier concentration results. Every second year is shown for parsimony.
Source: Thomson Reuters Worldscope and Bureau Van Dijk Orbis databases.

Orbis-Worldscope financial information

22. The combined Orbis-Worldscope firm financial data allows measures of concentration at the level of 2-digit NACE rev.2 industries and at world regions. We plan to compute country-level metrics in the next steps of this work. One caveat is our data currently has poorer coverage of subsidiary-level information for US firms. Note this is not the case for US business-group level information, since Orbis and Worldscope appear to cover close to the universe of listed firms. However, this limitation is something we wish to improve as a next step.

23. We report trends only for better-covered economies and industries, since Worldscope reflects only listed firms and the coverage of Orbis varies by country, industry and over time (see Bajgar et al., forthcoming). We present concentration metrics for manufacturing and non-financial market services (excluding 19 – Manufacture of coke and refined petroleum products, 68 – Real estate activities and 70 – Activities of head offices;

management consultancy activities).[31] However, recall a much broader set of industries is included in the underlying calculation.

24. For similar data comparability reasons, we present analyses for Europe and North America. We include the years 2000-2014 for 21 economies. We include them and complete the information for the leading firms from other sources (e.g. annual reports). The metrics for Europe are based on Belgium, Denmark, Estonia, Finland, France, Germany, Greece, Hungary, Ireland, Italy, Latvia, the Netherlands, Norway, Poland, Portugal, Slovenia, Spain, Sweden and the UK, and those for North America reflect the US and Canada. However, recall a much broader set of countries is included in the underlying calculation.

Orbis-Zephyr ownership information

25. Our primary source of ownership information is Orbis, which we supplement with data on the Zephyr database of Mergers and Acquisitions (M&As). The data are both provided by Bureau Van Dijk (BvD) and share a common firm identifier, which allows us to merge the two datasets. Orbis contains comprehensive information on ownership linkages between firms, which has been extensively used in the academic literature (e.g. Cravino and Levchenko, 2017), however the data primarily start in 2007 and sometimes much later for some firms. Our primary use of the Zephyr M&A database is to measure earlier changes in ownership, enabling the construction of a series starting as early as 2000 for some firms.

26. Orbis contains information on both ownership linkages between shareholders and subsidiaries and the global ultimate owners of subsidiaries calculated at each calendar year end. There are ownership linkages since the early 1990s, but coverage is better from 2007 onwards – we use all linkages available to us. The global ultimate owner is available at each year end 2007 to 2016 and is calculated by BvD.

27. There are different definitions of global ultimate owner, depending upon the minimum ownership percentage (25.01% or 50.01%) or the type of ultimate owner (reflecting either firms or also individuals). Since our interest is in measuring firm concentration, we consider firm-type ultimate owners.[32] We focus on the 50.01% ownership criteria since this is a commonly used threshold for the definition of control of another firm and hence whether the subsidiary financials are consolidated into the parent accounts[33].

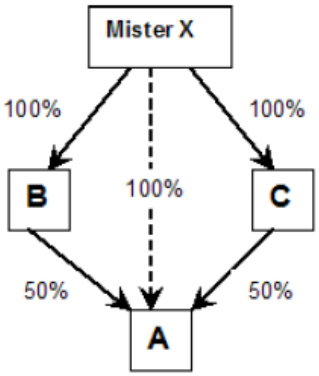
[31] For North America we also exclude 14 – Manufacture of wearing apparel, 15 - Manufacture of leather and related products and 30 - Manufacture of other transport equipment.

[32] Note that the firm-type ultimate owners include industrial, financial and insurance companies and banks. This definition excludes individuals, as well as mutual and pension funds, foundation and research institutes, public / state owners, employees/managers/directors, self-ownership, private equity firms, unnamed shareholders, venture capital or hedge funds.

[33] For example, a majority of voting rights is often a sufficient condition for control under international accounting practices. Whilst we do not observe voting rights, we assume these are reflected in shareholder ownership percentages, such that a majority (50.01%) of shares reflects a majority of voting rights. This is clearly a first-order approximation. For example, some shares can carry more voting rights than others and some types of shares have no voting rights at all. In addition, definitions of control can vary across accounting practices and having a minority of the voting rights can still imply de facto control if the remaining shares are spread across a large number of parties.

28. To calculate ultimate owners, BvD use the tree of ownership linkages for each firm, every year. They identify each firm's immediate shareholders, then their shareholders' shareholders and so on. So for each firm, they start at the bottom and work up their 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 (further details are included in the Orbis Ownership Guide).

29. We use the ultimate owner defined using the 50.01% minimum ownership percentage. Importantly, the minimum ownership criteria applies to both direct ownership linkages and indirect ownership linkages. It is possible to control a firm with more than 50.01% *indirect* shareholding, even though direct shareholdings are beneath this threshold. The example below highlights the difference between indirect and direct linkages. Mister X directly owns 100% of firms B and C, but *indirectly* owns 100% of A (=50% + 50%), through firm B and C.



Source: BvD Orbis User Guide

30. Whilst ultimate ownership data starts in 2007 in theory, for many firms it is not available until later years. The partial and improving coverage over time can generate severe challenges for researchers.

31. Common approaches 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 of these approaches are problematic. With increasing coverage in Orbis ownership over time the former approach will falsely equate missing data with independence and lead to an overstatement of ownership changes over time. The latter approach will clearly lead to an understatement of ownership changes over time and will typically overstate the number of markets and countries in which a firm operates. Whilst other researchers have noted the improving coverage over time, these issues are further complicated by churning in the data, with many living firms leaving the ownership data each year. We find that approximately 50,000-100,000 firms enter the Orbis ownership sample each year, and around 25,000-40,000 leave each year.

32. We complement the rich ultimate ownership data of Orbis with ownership changes identified from the Zephyr M&A database and or Orbis historic ownership linkages. The ownership information in Orbis is particularly large - approximately 116million subsidiaries have some ownership linkage information in the OECD's 2017_1 version of Orbis. Zephyr allows us to identify missing ownership changes, particularly for these earlier years. Zephyr contains deal-level information on M&As from 1997 for European

firms, from 2000 for North America and global deals from 2003. Given our focus on Europe and North America, this constrains our analysis to 2000. Zephyr contains about 1.4million M&A deals from 2000 to 2016.

33. We undertake several steps to expand the coverage:

- A first step is to use Zephyr to identify changes in immediate (rather than global ultimate) owners not available from Orbis.
    - For each deal, Zephyr contains information on the target, acquiror and vendor firms. About 890,000 deals represent either changes in majority ownership – such as a firm increasing from 10% to 51% equity ownership - or a majority owner further increasing their stake – such as a firm increasing from 51% to 60% ownership.
    - Both types of deal allow us to identify the immediate owner of each target firm at the time of the deal.
    - Furthermore, for changes in majority ownership, i.e. when the target firm switches hands, the vendor firm represents the previous immediate owner.
- A second step, is to undertake a similar exercise using Orbis historic ownership linkages to identify changes in immediate owners no available from Orbis.
    - We use the (oddly named) "current" Orbis linkage table. Although there are annual tables (2007-2015) containing direct linkage information, we found this to be a subset of direct linkages in the "current" tables – with older linkages pre-2007 excluded and some linkages in later years too. Given our intention to extend ownership to earlier years, the current table was preferred.
- A third step is to translate the changes in immediate owners (from the first two steps above) to changes in ultimate ownership.
    - The immediate owner who acquired the target firm may not be the ultimate owner. To find the ultimate owner we follow the same procedure as BvD, noted above.
    - We combine both Zephyr immediate ownership (from the preceding paragraph) and available Orbis information on ownership linkages to find the shareholders of the immediate owners, and the shareholders of their shareholders and so on. We use a 50.01% criteria and continue until we find a shareholder that is independent or controlled by an individual. This final shareholder is deemed the ultimate owner.
    - Note this step also corrects for firms in Orbis that are majority owned according to Orbis ownership linkages, but are missing an ultimate owner.
- A fourth step is to impute missing years of ownership 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 one to roll the owner backwards until there was a change in owner, rather than simply assuming a missing ultimate owner implies independence. In particular, we roll-forward and backward known ultimate owners at specific points in time.

- For example, if we know firm A has an ultimate owner of firm C in 2010, and from Zephyr M&A data that target firm A was acquired in 2008, we can roll backwards the ultimate ownership until 2008. Moreover, in about half of cases we know that target firm A was acquired from vendor firm B in 2008, so we infer that firm A was the (immediate) owner of firm B, and roll back further until an earlier M&A transaction.

34. The first three steps above, supplementing the ultimate ownership in Orbis with changes in ownership identified by M&As or changes in immediate ownership (via Orbis historic ownership linkages) adds approximately 400,000 firm observations each year. The fourth step above, rolling backward and forward the ownership adds approximately 1 million firm observations each year.

35. We also undertake an extensive array of cleaning steps to identify and correct potential issues in the ownership data –focused towards identifying missing linkages amongst the largest firms. Spot-checking revealed that some large firm groups are missing ownership linkages for some years between the parent firm and their subsidiaries. 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, we undertake the following checks to mitigate this risk:

- First, we correct ultimate owners that are in fact majority owned by another firm, since by definition they cannot be an ultimate owner.
- Second, we remove temporary (one or two year) changes in ultimate owner that reverse themselves – as this seems highly unlikely to occur in reality.
- Third, to detect missing linkages we examine large firms that change from having no subsidiaries to a large number of subsidiaries from one year to the next.
- Fourth, to identify missing links we examine large firms that never have any subsidiaries and the opposite case, large groups of subsidiaries that never have a parent with financials.
- Fifth, we identify missing links where there are ownership changes amongst firms have very similar names and are so very likely part of the same group all along (e.g. ABC Motors acquired by ABC Motors Thailand).
- Finally, we manually check 300 of the very largest firms, using the subsidiary structure listed in their financial statements each year to cross-check against the resulting ownership data.

36. These ownership checks each correct a differing number of firms. The first two steps correct only around 5,000 firms per year. Almost all firms that are themselves ultimate owners are not majority owned by another firm (although they may be owned by an individual as noted earlier), and there are few changes in ultimate ownership that reverse themselves so quickly[34].

37. The third step uses the panel structure of the data to identify potential missing ownership linkages by using *changes* in subsidiary structure for large firms. Spot checking revealed cases of intermediate holding companies (that often are without financials) being temporarily incorrectly identified as the ultimate owner, rather than the true parent firm. To address this, we examine large groups of subsidiaries (in terms of sales) that have no

[34] Spot checking some of these cases revealed these were all data anomalies.

parent with financials, but switch to a new parent the following period that does have financials[35]. We exclude cases of M&As as identified by Zephyr and consider only cases where more than 90% of subsidiaries transfer to the new parent.

38. We find 7000 cases of this per year, and correct the 2000 largest groups. For 200 groups with sales above 1 billion Euros we manually inspect each group against their financial statements and find 85% need correcting, 10% are correct and 5% are unknown. For the next 1800 largest groups with sales above 100 million Euros, we use a name-matching algorithm to semi-automate identification of whether the prior owner was in fact a holding company of the new parent and then manually inspect each of the close matches[36]. We considered those with very similar names part of the same group and we correct 850 of these groups.

39. The fourth step builds on the step above, identifying large groups of subsidiaries that *never* have a parent with financials, and large parents that *never* have subsidiaries. We found 1500 parents with sales of more than 1billion Euros that never have subsidiaries, and 300 groups of subsidiaries with more than 1billion Euros of sales that never have a parent with financials. Again, we use a name-matching algorithm to semi-automate identification of whether the prior owner was in fact a holding company of the new parent. We treated those with very similar names part of the same group, and consequently, we correct approximately 700 groups per year.

40. The fifth step considers any ownership change where the owners have a similar root to their name, e.g. XYZ Inc, XYZ Plc. By definition these remaining firms are not large, or do not have completely missing subsidiaries, or they would have been encompassed in the earlier cleaning steps. Accordingly these are 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, we undertake an automated check using name matching and require an identical match of the cleaned name. We remove common company type abbreviations (e.g. Plc, Ltd, SA, Gmbh etc.), country names (e.g. ABC (Viet Nam) Ltd) and punctuation, and require the resulting root of the name to be identical. This corrects approximately 400 groups per year, but these are not often large.

41. Finally, we manually check 300 cases of the largest groups against the subsidiaries listed in their financial statements. These groups have been identified through unusual ownership changes, strange trends in group sales or are disproportionately large for their industry. Approximately 50% of these required some form of correction to their data, although sometimes half of these cases are simply timing issues between ownership and financial changes, for instance, resulting from a divestment.

42. With 116 million subsidiaries in Orbis ownership, the potential set of firms for which this exercise could be conducted is enormous. To make the exercise more feasible we restrict the ownership correction to:

- Firms which either have more than 20 employees or more than 2million Euros annual sales in Orbis (on average).
- And/or firms which are ever in Worldscope.

[35] We also considered groups of subsidiaries losing parents with financials, but there were few cases of this due to improving Orbis coverage over time.

[36] All 1800 resulting name matches were visually inspected for additional verification.

- And/or firms which are ever are engaged in an M&A within Zephyr as an acquiror, vendor or target;
- And/or firms which ever hold a patent – identified using the OECD Patstat-Orbis bridge – see Squicciarini and Dernis (2013).
- For the years that the firm is alive (i.e. after their year of birth and up to their year of death)

43. This results in corrected ultimate ownership for 2.8 million firms for the period 2000-2016. For the (small) firms that fall outside this exercise, we use the uncorrected Orbis ultimate owner information. The corrections lead to a substantial gain in the number of firms for which it is possible to identify their ultimate owner. The number of firms with a global ultimate owner in the original uncorrected Orbis data and post-corrections above are shown in Table A3 below:

Table A.3. Number of Firms with Temporary Changes in Ultimate Owner

	Firms with Uncorrected Global Ultimate Owner	Firms with Corrected Global Ultimate Owner
2000	0	1,572,602
2001	0	1,662,406
2002	0	1,756,356
2003	0	1,852,251
2004	0	1,961,766
2005	0	2,062,816
2006	0	2,157,180
2007	853,442	2,232,027
2008	945,696	2,309,021
2009	1,075,344	2,382,528
2010	1,122,678	2,429,363
2011	1,288,558	2,459,739
2012	1,465,139	2,472,003
2013	1,532,972	2,463,395
2014	1,544,480	2,424,373
2015	1,722,394	2,375,192
2016	1,808,716	2,321,353

44. The corrected ownership shows some reduction in sample size in the early 2000s, partly because fewer firms have been born at that point, and partly because of our conservative approach to inferring ownership in earlier years. For example, we roll back the ultimate owner to earlier years up to the point of an M&A change in ownership. For years before that point we use (the ultimate owner of) the vendor, or, if the vendor is unknown or is an individual, we set the ownership to missing for these years.

OECD PRODUCTIVITY WORKING PAPERS

OECD Productivity Working Papers are published on oe.cd/GFP

05. *The Best versus the Rest: The Global Productivity Slowdown, Divergence across Firms and the Role of Public Policy*
(November 2016) by Dan Andrews, Chiara Criscuolo and Peter N. Gal

06. *What makes cities more productive? Agglomeration economies and the role of urban governance: Evidence from 5 OECD countries*
(February 2017) by Rudiger Ahrend, Emily Farchy, Ioannis Kaplanis and Alexander C. Lembcke

07. *Pro-Productivity Institutions: Learning From National Experience*
(April 2017) by Andrea Renda and Sean Dougherty

08. *The impact of structural reforms on productivity: The role of the distance to the technological frontier*
(May 2017) by Gustavo Monteiro, Ana Fountoura Gouveia and Sílvia Santos

09. *Product markets' deregulation: A more productive, more efficient and more resilient economy?*
(September 2017) by Gustavo Monteiro, Ana Fountoura Gouveia and Sílvia Santos

10. *Achieving New Zealand's productivity potential*
(October 2017) by Paul Conway

11. *The Contribution of Multinational Enterprises to Labor Productivity: The Case of Israel*
(February 2018) by Tatiana Slobodnitsky, Lev Drucker and Assaf Geva

12. *GVCs and centrality: mapping key hubs, spokes and the periphery*
(February 2018) by Chiara Criscuolo and Jonathan Timmis

13. *Fear the walking dead: zombie firms, spillovers and exit barriers*
(June 2018) by Ana Fontoura Gouveia and Christian Osterhold

14. *GVC centrality and productivity: Are hubs key to firm performance?*
(June 2018) by Chiara Criscuolo and Jonathan Timmis

15. *Patterns of firm level productivity in Ireland*
(September 2018) by Javier Papa, Luke Rehill and Brendan O'Connor

16. *Productivity spillovers from multinational activity to local firms in Ireland*
(November 2018) Mattia Di Ubaldo, Martina Lawless and Iulia Siedschlag

17. *Unconventional monetary policy and productivity*
(January 2019) Silvia Albrizio, Marina Conesa, Dennis Dlugosch and Christina Timiliotis

BETTER POLICIES FOR BETTER LIVES