Master Thesis

Economics and Business: Financial Economics

The Impact of Industry Concentration and Digitization on the Financial Performance of European Firms

Emilio Jalil Lohi, BSc.

ID: 649243

Supervisor: Dr. Sjoerd van den Hauwe

Second Assessor: Dr. Philippe Versijp

ERASMUS UNIVERSITY ROTTERDAM

Erasmus School of Economics

August 9, 2023

Abstract

This paper analyzes industry concentration and digitization as potential explanatory variables for financial performance and stock returns. In contrast to previous findings, the log of industry concentration is negatively related to profitability in Europe for nine out of ten concentration metrics. Using a quantile regression approach further shows that the mostly negative impact of industry concentration on ROA is less in magnitude for highly profitable firms than for those with poor profitability, which indicates that mostly firms with weak profitability suffer from increases in their industries' concentration. Post-2000, a high degree of industry-level digitization consistently positively affects firm profitability, and stronger so for top-ROA-percentile firms, potentially providing evidence for the winner-takes-all characteristic of digitization. Average yearly stock returns of firms with the highest concentration change are significantly higher than those of firms with the lowest concentration change for two of four concentration metrics for a sample of 5251 firms obtained from *Orbis*. The same analysis with *Compustat* data yields opposite results. Industry concentration has been falling in Europe for the last two decades but at a decreasing rate. The implementation of the European single market seems to have broadly improved competition, but analysis of industry dynamism shows that incumbents are increasingly becoming insulated.

Key Words: ROA, Digitization, Stock Returns, Policy, Concentration-Performance Relationship, Product Market Competition, Market Power, Concentration Levels

JEL Classification: G12, G34, G38, L11, L13, L16

Acknowledgement

I would like to express my gratitude to the Erasmus School of Economics for providing the flexibility in course selection I was looking for. Thank you, Kim, for keeping me on track with your thesis coaching – you gave me structure and a positive mindset. Furthermore, I am immensely grateful for the freedom to explore my interests and for the insightful guidance and support of my supervisor, Sjoerd van den Hauwe. Lastly, many thanks to my parents, Barbara and Massoud – without whom, none of my accomplishments would have been possible.

Contents

1 Introduction

Does industry concentration affect financial performance? For American firms, Grullon et al. (2019) found a significant positive impact during the last two decades. But no previous study analyzed the impact of industry concentration on financial performance in Europe. It is the key contribution of this paper to fill this gap. Adopting a holistic view on performance, competition, and digitization, I analyze determinants of return on assets, excess stock market returns and the general trend of industry concentration.

Study of companies' financial performance is a typical subject of industrial economics, accounting, and finance (Pattitoni et al., 2014). At the micro-level, determinants of firm performance are easily identified. Management can improve performance either operationally, strategically, or financially. Operational improvements are typically efficiency gains, which could be obtained by merging with a rival, for instance. Strategic improvements are usually good investment choices. Examples of performance enhancing financing choices include adjusting debt-levels and ensuring liquidity. Ultimately, however, the direction and magnitude of performance-improving measures depends on a firm's specific situation, which makes aggregate analysis not clear-cut.

For instance, higher industry concentration might lead to greater market power and higher profits because the firm has less competition. However, it might also lead to less market power and lower profits because the drop-out of competition was due to the industry's unattractiveness. Losing a rival may not be enough to compensate for lower demand (Bain, 1951). Apart from the potentially ambiguous nature of explanatory variables, the structure of the economy itself might have changed. Digitization transformed many business models and shook-up entire industries. Especially network effects seem to be a driving force behind the benefits of digitization (Stallkamp & Schotter, 2021).

Most papers about European industry concentration analyze its time-series trend and that of related economic measures, such as the labor share and capital investment. Gutierrez and Philippon (2018) found a generally decreasing industry concentration trend in Europe. They developed a simple theory that predicts stronger antitrust institutions of the EU than of the US. The authors assert that a joint competition authority is stronger and more independent than that of any individual country. Individual countries are more concerned with preventing others from capturing the institution, than influencing it themselves. One of their testable predictions is that tougher and more independent regulators, as set up by the EU, lead to more competition in product markets. Using price data, Gutierrez and Philippon (2018) show decreasing European profit margins, mostly driven by deregulation. I use return on assets as a proxy for profitability and show that higher concentration is associated with lower profitability, strengthening

Gutierrez and Philippon's insight that deregulation made Europe more competitive. But my research goes beyond linear panel regression and consists of three parts.

In part one, I use historical balance sheet data to conduct panel regressions of ROA on industry concentration and common control variables, such as size, age, and capital expenditure using four different industry concentration metrics and two datasets, *Compustat* and *Orbis*. Grullon et al. (2019) use *Compustat* in their analysis of North America, while Kalemli-Ozcan et al. (2015) advocate the usage of *Orbis* because it is more comprehensive. *Orbis* does indeed cover more firms, but its pre-2000 coverage is low, just as with *Compustat*. Post-2000, however, the coverage is stable and analysis therefore valid. Following Bajgar et al. (2023), I use national- and supranational concentration measures and market-share-based as well as HHI concentration metrics. My *Compustat* market-share based measures use national accounts data to mitigate potential coverage bias by the commercial databases (Bajgar et al., 2023). I test one model with the natural logarithm of industry concentration as an explanatory variable, and two models with the natural logarithm of industry concentration and an interaction term (size and market share). Furthermore, I estimate the parameters for regression equations with the absolute value of concentration and its squared value to check for potential non-linear relationships.

However, because I suspect that increasing industry concentration is more beneficial for already highly profitable firms, I conduct cross-sectional quantile regressions in part one as well. Specifically, I estimate the industry concentration parameters for the 10^{th}, 25^{th}, 50^{th} 75^{th} and 90^{th} percentiles, using the same econometric models as with the panel regressions.

Moreover, I assess the impact of the degree of industry-level digitization in this series of cross-sectional quantile regressions. Gaspar et al. (2022) argue that there are endogeneity concerns. Firms with high stock-market valuations might have access to the necessary capital to invest in (profitable) digitization. Therefore, they conducted quasi natural experiments with the global financial crisis and Covid-19 pandemic and found that digital firms are more resilient to external shocks. Given the nature of the Covid-19 pandemic, this is not surprising. But higher resilience of digital firms during the global financial crisis might indicate generally higher profitability of digital firms. Hence, I conduct cross-sectional regressions of ROA on industry-level digitization. While a true causal relationship remains to be proven, significant correlation is already an interesting insight, especially because a comprehensive meta-analysis of the digitization-performance relationship is outstanding. Kohtamäki et al. (2020) analyzed the relationship between digitization and servitization, Broccardo et al. (2023) between digitization, sustainability and performance in an Italian context, and Theiri and Hadoussa (2023) examined the effect of digitzation on the financial performance of African banks.

In part two, I test the hypothesis that stocks of firms in industries with the highest concentration gains outperform those with the lowest. My enquiry is motivated by Grullon et al. (2019)'s finding of 6.6

percent excess stock returns of American top concentration change stocks as compared to bottom concentration change stocks (risk-corrected by the CAPM). Because their results do not change qualitatively using more refined asset pricing models, I employ portfolio sorted analysis of CAPM excess stock returns as well.

Part three analyzes the time-series trend of industry concentration metrics along criteria such as digitization and macro-industry affiliation. I follow the common practice to split between manufacturing and services and to exclude financial industries (Affeldt et al., 2021; Bajgar et al., 2023; Gutiérrez & Philippon, 2017). My main contribution is to compare the *Compustat* and *Orbis* datasets, to analyze geographical differences (Northern, Eastern, and Southern Europe), and to examine differences between the time-series of digitized and non-digitized industries.

My main results are as follows. The natural logarithm of industry concentration is significantly negatively associated with ROA for three of four *Compustat* concentration metrics, with the magnitude ranging from -2.7% to -10.1%, depending on the model. That is, a one percent increase in HHI is associated with a maximum decline in ROA of 0.11 percentage points. Given a mean ROA of 5.9%, industry concentration is economically impactful. Analysis with *Orbis* qualitatively confirms the insight that industry concentration is negatively associated with ROA in Europe.

Digitization consistently positively impacts profitability and increasingly so for the *Compustat* sample. While the parameter estimate of the *Compustat* industry-level digitization score is mostly negative in pre-2000 cross-sectional regressions and around zero from 2000 to 2010, it becomes positive post-2010, ranging from 0.01 to 0.06. An increase of one in the digitization score (which ranges from 1 to 4) is associated with a one to six percentage point increase in ROA. In *Orbis*, the digitization score parameters are already pre-2000 positive and constant around 1% across time, all models and measures.

The sorted portfolio analysis of the potential outperformance of firms that experienced high concentration gains yields mixed results. *Compustat* data shows that European stocks with the top-10% annual concentration change do not significantly outperform those with the bottom-10% concentration change. Furthermore, in *Compustat* the average excess return of the firms with the lowest concentration changes is higher than the one of the firms with the highest, which is counterintuitive. However, based on the country-level market-share based *Orbis* concentration measure, the mean excess annual return of the top-change portfolio (3.7%) is significantly higher at the 1%-level than that of the bottom-change portfolio (1.6%). Still, using different concentration measures makes the difference insignificant in *Orbis* as well, and even yields a higher average bottom-change excess return using EU-level HHI.

The analysis of the general trend of industry concentration shows that European industry concentration has been falling at a decreasing rate during the last two decades. In *Compustat*, only average concentration of post-2000 non-digital firms tends to increase over time. With *Orbis*, however, average industry concentration is falling or constant across all measures and subsamples.

The rest of this paper is organized as follows. In Section 2, I review the relevant literature and present my hypothesis development. Section 3 explains my empirical methods. In Section 4, I describe the data. Then I present the main results (section 5), conduct robustness checks (section 6) discuss my findings (section 7). Section 8 concludes.

2 Literature Review and Hypothesis Development

First, I review the determinants of financial performance, with a focus on industry concentration and digitization. Then, I summarize the relevant literature on the impact of industry concentration on stock returns. Finally, I outline the findings on the time-series trend of industry concentration trends for Europe and the USA.

2.1 Determinants of Financial Performance

The following subsections first summarize traditional determinants of firm profitability before motivating my focus, and reviewing related papers on digitization and industry concentration as drivers of financial performance.

2.1.1 Traditional Determinants of Firm Profitability

Financial performance is commonly explained by the structure-conduct-performance paradigm (Pattitoni et al., 2014). In table 2.1 you can see an overview of traditional determinants of profitability.

Strategy	**Environment**	**Organization**
*Growth (+)	*Industry Concentration (+)	Capacity Utilization (+)
*Capital Investment (-)	*Industry Growth (+)	
*Firm Advertising (+)	*Industry Capital Investment (+)	
*Market Share (+)	*Industry Size (+)	
*R&D Expenditure (+)	*Industry Advertising (+)	
Debt (-)	Industry Imports (-)	
Diversification (-)	Industry Minimum Efficient Scale (+)	
Quality of Product & Service (+)	Industry Geographic Dispersion (+)	
Vertical Integration (+)	Industry Barriers to Entry (+)	
Corporate Social Responsibility (+)	Industry Exports (-)	
	Industry Economies of Scale (+)	

Table 2.1: Determinants of financial performance: The structure-conduct-performance paradigm. The signs in brackets indicate literature's consensus on the hypothesized relationship. Explanatory variables marked with a star were confirmed as significant in meta-analysis. Source: Capon et al., 1990, p. 1156.

At the firm-level, growth, advertising, market share, and R&D expenditure have a significant positive impact on profitability, while capital investment has a negative one. Capon et al. (1990) found that the structural factors of industry concentration, industry growth, capital investment, size, and advertising, have significant positive effects on firm performance. Furthermore, Capon et al. (1990) stress that interaction terms might add explanatory power.

Concerning micro-level variables, Pattitoni et al. (2014) found that liquidity has a consistently positive effect on profitability for panel data of approximately 30.000 EU-15 firms. Moreover, they found that leverage has a negative impact on profitability, highlighting the detrimental effect of high debt service during an economic downturn. Although the authors find nonlinearities, the mean firm is far below the threshold after which debt holdings enhance profitability, possibly due to Jensen's control hypothesis (Jensen, 1986). Furthermore, growth has a significant positive effect on ROA in all their models, probably because of a motivating effect on employees. Size is consistently negatively related to ROA. The higher a firm's opportunity cost of capital, the lower the profitability and because higher opportunity costs lower shareholders' commitment levels, the negative effect is additionally amplified (Pattitoni et al., 2014). Regarding the macro-level determinants of profitability, their study of non-linear effects showed that firm profitability is pro-cyclical, i.e., that it is higher in bull-markets and in times of high GDP growth. Therefore, I include time-fixed effects in my ROA regressions. Finally, Pattitoni et al. (2014) find a negative time-series trend of average EU-15 profitability between 2004 and 2011, based on *Amadeus* data, which is relevant for the third part of my paper.

Goddard et al. (2005) conducted dynamic panel regressions of ROA of Belgian, French, Italian, Spanish, and British firms from 1993 to 2001. Their covariates include two lags of profitability and firm-fixed effects. They found a negative impact of size but a positive one of market share. Furthermore, debt levels are negatively, and liquidity positively related to ROA in their models.

Another approach to modeling firm profitability is the random-walk hypothesis of profits, which dates back to the 1960s and finds some support (Chan et al., 2003; Little, 1962; Rayner & Little, 1966). However, my focus is on the first strand of performance modeling literature. Therefore, I use seven of the above listed traditional determinants of profitability as control variables in my regressions of ROA on industry concentration and digitization: market share, age, log of assets, inverted assets, R&D intensity (expenditure over assets), and capital investment intensity (CAPEX/Assets). My choice is mainly motivated by data availability of *Compustat* and *Orbis*.

2.1.2 Digitization as a Driver of Financial Performance

Kohtamäki et al. (2020) assert that there is still little empirical research on the impact of digitization on firm performance. In the context of manufacturing firms, they further argue that digitization alone may not be enough to impact financial performance. They state that many manufacturers struggle with appropriating the value of their investments in digitization, due to operational difficulties. The direct impact of digitization on financial performance is therefore complex and possibly non-linear. However, combined with servitization, i.e., the provision of goods as services, they found a U-shaped relationship between firm performance and the interaction term of digitization and servitization. Given a high level of servitization, the profitability initially decreases with digitization before increasing again.

Broccardo et al. (2023) conducted partial-least squares regressions of profitability for a sample of publicly listed Italian companies and found that both sustainability and digitization significantly positively affect returns.

Gaspar et al. (2022) found that digitization improves firms' resilience to economic shocks. They created a portfolio consisting of the highest quartile of firms in terms of digital exposure and one of the lowest quartile. During the period from February 24 to March 23, 2020, the high digitization portfolio earned a cumulative abnormal return (CAR) of 4%, while the low digitization portfolio underperformed with -8% CAR. During health crises digitized firms benefit inter alia from contactless transactions. For the global financial crisis, they found that a one standard deviation increase in digitization yields an increase in abnormal stock returns of 82 basis points per month (significant at the 1%-level). They only analyzed stock returns and not return on assets, a gap I intend to fill. The advantage of their approach to measuring digitization is that you can construct it for every publicly listed firm. Put simply, they computed the share of digitization-related OECD vocabulary in firms' 10k reports and conference calls. The downside, however, is that you need text mining skills to construct the measure and that pure usage of vocabulary might not truly reflect implementation of digitization but rather promises of CEOs. Still, their paper shows that stocks of highly digitized firms tend to outperform poorly digitized ones during shocks.

Hua (2022) constructs a firm-level digitization score by first attributing a score to 800 different jobs, computing their share of employment for 300 industries (thereby obtaining an industry-level digitization score), and using segment data to extrapolate a time-varying firm-level score for approximately 5000 firms. She finds that "a strategy which is long (short) digital (non-digital) firms earns 6.5% per annum beyond common risk factors over the sample period July 2000 to June 2019" (Hua, 2022, p. 3) Therefore, stock market outperformance of digital firms is well documented. Hence, I focus on potential ROA outperformance. The construction of my digitization scores is outlined in the methodology section.

2.1.3 Industry Concentration as a Driver of Financial Performance

Apart from industry concentration as a proxy for competition, profitability is influenced by factors such as demand, technical progress, and entry and exit conditions (Bain, 1951). While industry concentration moderates the other factors, it is not the sole determinant of profitability. Hence, the relationship between industry concentration and profitability presumably changes over time, region, and analyzed industries. However, there is a general tendency of monopolies to outperform competitive oligopolies or atomistic structures, especially in the long-run and on average. Differences in technological progress, demand, and other such factors usually average out (Bain, 1951).

Empirical research on industry concentration as a determinant of firm profitability dates back to the 1950s. Bain (1951) found that average return on equity of firms with higher concentration was significantly higher than that of firms with lower concentration. His sample covered US manufacturers from 1936 to 1940. He formed their portfolios of companies along the threshold of 70% market share of the eight largest firms.

Based on data of US manufacturing companies from 1958 to 1981, Domowitz et al. (1986) found profit margins of highly concentrated industries were more pro-cyclical than those of less concentrated industries. That is, profits of highly concentrated industries rise stronger during an economic upswing but also fall stronger during a downturn.

Recently, Grullon et al. (2019) found, in line with most previous empirical papers, that industry concentration in the US positively impacted ROA. For their main analysis, the authors used the *CRSP-Compustat* merged dataset of NYSE, AMEX and NASDAQ-listed companies over the time period of 1972 to 2014. For the subsample from 1987 to 2000 their coefficient of the log of HHI was slightly positive (0.0007) but insignificant. However, for the subsample from 2001 to 2014 the parameter estimate was 0.0168 and significant at the 1%-level. Over the entire sample period the coefficient of the log of HHI was 0.0027 and significant at the 10%-level.

In short, studies on industry concentration's impact on ROA were mostly conducted for the US and found a positive relationship. However, also a negative impact is theoretically possible, especially due to demand conditions, technical progress and regulation (Bain, 1951).

2.2 Industry Concentration and the Cross-Section of Stock Returns

Hou and Robinson (2006, p. 1927) found that "firms in more concentrated industries earn lower returns". They conducted sorted portfolio analysis of average monthly stock returns of American companies from 1963 to 2001 and from 1973 to 2001. In June of each year, they sorted the firms into quintiles. Based on this classification, they compared the average monthly stock returns of the top concentration quintile with the bottom quintile and found significantly lower returns of the top. Part of their explanation is that companies in highly concentrated industries are less risky and therefore have lower expected returns. Furthermore, they feel less competitive pressure and have therefore not as much need to innovate (Döttling et al., 2017), which leads to lower profitability and stock returns (Arrow, 1972).

Gallagher et al. (2015) come to opposing findings for stocks of large Australian firms. They found a positive association between concentration and innovation expenditure, which has been linked to positive excess stock returns by extant literature. That large companies in highly concentrated industries earn positive excess stock returns, was confirmed by their Australian data set. They argue that the opposing findings are resolved when examining the structural differences between American and Australian industries.

Bustamante and Donangelo (2017, p. 4216), however, postulate a negative relationship between concentration and returns, in line with Hou and Robinson (2006) arguing that "competition erodes markups such that firms are more exposed to systematic risk [and that] the threat of entry by new firms lowers exposure to systematic risk of incumbents".

Using three different asset pricing models (CAPM, FF3, FF5), Grullon et al. (2019) do not find significant abnormal returns of high-concentration change portfolios compared to low-concentration change portfolios over the period of 1972 to 2014. However, analyzing the subsample from 2001 to 2014, a period during which American industry concentration substantially increased, Grullon et al. (2019) found positive and significant alphas from 6.6% to 8.2% per year. In this case, the authors corrected for systematic risk by using the CAPM, but the results did not change with the other two models. Therefore, the relationship between concentration and stock returns seems to have inverted from a significant negative relationship pre-2001 (Hou & Robinson, 2006) to a significant positive relationship post-2001, at least in the US.

2.3 General Trend of Industry Concentration and Economic Developments

For the United States, there is a general consensus that industry concentration has risen over the past decades (Furman & Orszag, 2015; Philippon, 2019).

2.3.1 USA

For instance, Grullon et al. (2019) show that North American industry concentration has steadily been increasing since 1997. They find that since 1997 the Herfindahl-Hirschman Index (HHI) has "systematically increased in more than 75% of US industries, and the average increase in concentration levels has reached 90%" (Grullon et al., 2019, p. 698). The authors found that the main source of abnormal profits is market power, proxied by the Lerner index, and that higher markups are accompanied by higher stock returns.

Based on data from the *US Census Bureau*, Furman and Orszag (2015) found a significant increase in the market share of the 50 largest companies. Furthermore, wage differences across firms have increased over the last decades, driven by a rise in returns of top-percentile firms. The gap between the profitability of the 90th percentile and the one of the median, became markedly wider since the early 1990s. Lastly, they stress the importance of further firm-level research to uncover the underlying reasons for the increase in abnormal returns of top-percentile firms and its impact on inequality, something I do for European data with quantile regressions.

2.3.2 Europe

Several recent developments suggest that industries in Europe are becoming increasingly concentrated too. First, rather than completing traditional stand-alone deals, private equity (PE) firms progressively engage in strategic serial acquisitions to consolidate industries and reap profits from higher efficiency and greater market power (Bansraj et al., 2020; Bansraj & Smit, 2017; Hammer et al., 2022). Second, anecdotal evidence suggests that family-run small businesses are in decline. Often a walk in your neighborhood is enough to see this. Finally, several academic papers conclude that industry concentration has risen. In 2018, the OECD published a report which suggests that markets around the globe are becoming more concentrated and less competitive (Pike et al., 2018). Bajgar et al. (2023) and Affeldt et al. (2021) assert as well that industry concentration has been rising, using European data.

In contrast, Döttling et al. (2017) find that EU industry concentration has remained stable or even declined in some industries. They showed that both in Europe and the US, investment has been weak for the last 20 years, especially since the global financial crisis. Profits, however, are higher in the US than in Europe, and also stock market valuations (Tobin's Q) are higher in the US. Furthermore, profits have been stable or declining in the EU, while they have been rising in the US. Döttling et al. (2017) hypothesize that weaker American antitrust and competition policy is driving the differences. Q theory holds for Europe, but American investment has been below the predicated values. They identified several causes that might explain the deviations: financial frictions, measurement error of intangible investment, and competition. In short, the presumably temporary European investment gap is mainly explained by depressed asset values due to "financial constraints, high risk premia, low expected demand and low expected cash flows" (Döttling et al., 2017, p. 41). In the US, however, "investment is depressed because industries have become more concentrated over time and competitive pressures to investment are lacking" (Döttling et al., 2017, p. 41).

Most recently, Bajgar et al. (2023) added to the literature on industry concentration in Europe and North America by analyzing country-level and world-region concentration trends for two-digit industries. Their paper highlights the importance of addressing time-varying database coverage of firms, which may cause biased concentration trends. They confirm the hypothesis that globalization and technological change are the likely drivers of industry concentration since they find an increase in Europe as well as America. For the country-level analysis they use representative firm-level data from the OECD STAN database. For the world-region analysis they use Orbis data complemented by certain other databases. In particular, they use OECD STAN data for the denominator of industry sales to address the coverage bias. Furthermore, they address subsidiary linkages by analyzing business groups rather than just individual firms. Moreover, they controlled for imports for the concentration metric of manufacturing industries but could not control for exports of individual firms due to lack of data. However, they argue that the effect might be negligible. While they find an increase in European industry concentration that is in contrast to Döttling et al. (2017)'s decreasing concentration trend, they find the increase to be less in magnitude than in the US. These conflicting findings motivate my comparison of four different industry concentration measures using two different datasets.

3 Methodology

The methodology section proceeds as follows. First, I present my empirical methods. Second, I motivate my choice of classifications and metrics. Third, I briefly discuss pitfalls and limitations before summarizing the testable hypotheses.

3.1 Methods

To examine the impact of industry concentration on ROA, I first conduct several panel regressions. I further conduct a series of cross-sectional quantile regressions, to investigate a potential non-linear relationship between industry concentration and ROA. I also use the cross-sectional quantile regression approach to assess the impact of digitization on ROA. For the analysis of potential abnormal stock returns of highly concentrated industries, I conduct sorted portfolio analysis. To inspect the time-series trend of European industry concentration, I apply a linear trend model and conduct sub-sample comparisons of average concentration along the digitized/non-digitized and manufacturing/services splits.

3.1.1 Panel Regression

I estimate the parameters of the following five panel regression models. All models use continuous de-trended variables as regressors. In all the subsequent models, subscript *i* denotes firms, *t* time, *ROA* operating income before depreciation, *α* firm-fixed effects, *Marketshare* firm sales over industry sales, *age* firm age, *Assets* total firm assets, *R&D* research and development expenditure over assets, *CapEx* capital expenditure over assets, *C* the respective concentration metric, and *ε* denotes the error terms.

Model 1: Basic model

$$ROA_{it} = \alpha_i + \beta_1 Marketshare_{i,t} + \beta_2 \log(Age_{i,t}) + \beta_3 \log(Assets_{i,t}) + \beta_4 \left(\frac{1}{Assets}\right) + \beta_5 R\&D + \beta_6 CapEx + \beta_7 \log(C_{i,t}) + \varepsilon_{it} \quad (1)$$

The basic model includes the natural logarithm of concentration and several common control variables as regressors.

Model 2: Basic model + interaction term of size and concentration:

$$ROA_{it} = \alpha_i + \beta_1 Marketshare_{i,t} + \beta_2 \log(Age_{i,t}) + \beta_3 \log(Assets_{i,t}) + \beta_4 \left(\frac{1}{Assets}\right) + \beta_5 R\&D + \beta_6 CapEx + \beta_7 \log(C_{i,t}) + \beta_8 \log(Assets_{i,t}) * \log(C_{i,t}) + \varepsilon_{it} \quad (2)$$

To check whether big firms benefit more than small ones from higher concentration, I include the interaction term of the log of size and the log of concentration in the second model.

Model 3: Basic model + interaction term of market share and concentration:

$$ROA_{it} = \alpha_i + \beta_1 Marketshare_{i,t} + \beta_2 \log(Age_{i,t}) + \beta_3 \log(Assets_{i,t}) + \beta_4 \left(\frac{1}{Assets}\right) + \beta_5 R\&D + \beta_6 CapEx + \beta_7 \log(C_{i,t}) + \beta_8 Marketshare_{i,t} * \log(C_{i,t}) + \varepsilon_{it} \quad (3)$$

Similarly, model 3 augments the basic model by adding the interaction term of the log of market share and the log of concentration. This interaction term allows for investigating potential nonlinearity and heterogeneity in the relationship between market share, concentration, and ROA.

Model 4: Basic model but with absolute concentration level instead of its log

$$ROA_{it} = \alpha_i + \beta_1 Marketshare_{i,t} + \beta_2 \log(Age_{i,t}) + \beta_3 \log(Assets_{i,t}) + \beta_4 \left(\frac{1}{Assets}\right) + \beta_5 R\&D + \beta_6 CapEx + \beta_7 C_{i,t} + \varepsilon_{it} \quad (4)$$

In model 4, I replace the log of concentration by the absolute concentration measure. This change allows for investigating the impact of the concentration metric on ROA without assuming a logarithmic relationship.

Model 5: Basic model but with squared absolute concentration level instead of its log

$$ROA_{it} = \alpha_i + \beta_1 Marketshare_{i,t} + \beta_2 \log(Age_{i,t}) + \beta_3 \log(Assets_{i,t}) + \beta_4 \left(\frac{1}{Assets}\right) + \beta_5 R\&D + \beta_6 CapEx + \beta_7 (C_{i,t})^2 + \varepsilon_{it} \quad (5)$$

In Model 5, I use the squared absolute concentration level as the concentration metric. By including the squared term, I investigate the possibility of a curved relationship between concentration and ROA.

The parameters β_1 to β_6 refer to the same control variables in all models and can be interpreted as follows. The coefficient β_1 represents the marginal effect of market share on ROA. A one-unit increase in market share leads to a β_1 increase in ROA. For instance, a 0.1 (10%) increase in market share would lead to a 0.1*$\beta_1$increase in ROA, which is measured in fraction form as well. β_2 is the marginal effect of the log of firm age on ROA. For a one-unit increase in the natural logarithm of firm age, holding all other variables constant, the expected change in ROA is β_2. Parameter β_3 represents the marginal effect of firm assets on ROA. For a one-unit increase in the natural logarithm of firm assets, ceteris paribus, the expected change in ROA is β_3. If assets increased by the absolute amount *x*, ROA would change by β_3/*x*. The coefficient β_4 is the marginal effect of the inverse of firm assets on ROA. For a one-unit increase in the inverse of firm assets, the expected change in ROA is β_4. This implies that as a firm's assets decrease (or the inverse of assets increases), its ROA is expected to change by β_4 units. Economies of scale and scope predict a negative β_4. β_5 represents the marginal effect of R&D expenditure over assets on ROA. For instance, predicted ROA would change by β_5 if R&D expenditure increased by the amount of a firm's assets. β_6 is the effect of capital expenditure over assets on ROA. This implies that if a firm increases its capital expenditure relative to its assets by one unit, i.e. by the amount of its assets, its ROA is expected to change by β_6 units, assuming other factors remain constant. Value and amount of β_6 are ex-ante not clear. It depends on a firm's investment opportunities. Q theory would predict a positive β_6 if Tobin's Q were positive.

β_7 is the parameter of industry concentration. Here, the models differ in their interpretation. In models 1-3, β_7 is the marginal effect of the natural logarithm of the respective concentration metric on ROA. In model 1, for instance, a one percent increase in HHI would result in a $(\beta_7/100)$ change in ROA. In model 4, β_7 refers to the absolute concentration level. Therefore, a one percentage point increase in HHI (i.e., a 1/100 unit increase if HHI is on a scale from 0 to 1) would results in a $\beta_7 * 0.01$ change in ROA. In model 5, β_7 represents the marginal effect of the squared absolute concentration on ROA. Therefore, a one-unit increase in absolute concentration would results in a ROA change of $2 * \beta_7 * C$, where C stands for the initial concentration level. The higher the initial level, the higher the impact of a concentration change. This seems reasonable for the top end of the concentration distribution, because a further increase of ex-ante already high concentration reflects the step toward monopoly (in which ROA is expected to rise significantly). However, at the bottom end of the concentration distribution, this relationship does probably not hold. A concentration increase in a very fragmented industry often results in substantial efficiency gains of a large new player.

β_8 is only included in models 2, and 3 as the coefficient of their respective concentration interaction term. In model 2, β_8 is the marginal effect of the product of the log of concentration and the log of assets. I expect β_8 to be positive, because a one percent increase in concentration would lead to a $(\beta_7 + \beta_8 * \log(Assets_{i,t}))/100$ change in ROA, meaning that larger size should be associated with higher profitability gains when concentration increases. Similarly, a one percent increase in assets would lead to a $(\beta_3 + \beta_8 * \log(C_{i,t}))/100$ change in ROA, meaning that a higher level of concentration increases the positive size impact on profitability. In model 3, β_8 is the marginal effect of the product of the log of concentration and the market share (using the respective concentration metric). A one percent increase in concentration would lead to a $(\beta_7 + \beta_8 * Marketshare_{i,t})/100$ change in ROA. A one percentage point increase in market share would lead to a change in ROA of $(\beta_1 + \beta_8 * \log(C_{i,t}))/100$, with a positive parameter again indicating a stronger marginal effect of market share on ROA in the case of a high concentration level.

To control for potential autocorrelation and endogeneity, I cluster standard errors at the firm level and include firm, and time-fixed effects. Including firm-fixed effects prevents a mechanical relation between ROA and industry concentration. Should profitable firms regularly take over nonprofitable ones, an increase in industry concentration would automatically lead to a decline in profitability (Grullon et al., 2019). Therefore, controlling for time-invariant firm specific characteristics helps disentangle the effect of industry concentration on profitability from such potentially confounding effects.

3.1.2 Quantile Regression

Developed by Koenker and Basset (1978) quantile regressions provide a non-parametric approach to estimate models for the conditional quantile functions. Their main purpose is to examine non-linear relationships. This method also has the advantage of being "more robust to outliers and non-normality than OLS regression" (Brooks, 2019, p. 169). My reason for conducting quantile regressions additional to panel regressions is to examine whether highly profitable firms are more sensitive to changes in industry concentration levels. Autor et al. (2020) characterized "superstar-firms" as the most productive companies in each industry, who capture an increasingly large market share. They are characterized by a "above-average markups and below-average labor shares" (Autor et al., 2020, p. 648). If industry concentration had a more positive impact on returns for these superstar-firms (proxied by the portfolio of firms in the 90th ROA percentile), there would be evidence that revenue maximization is beneficial for profit maximization and shareholder value (holding industry structure of rivals roughly constant).

The estimated regression equations are similar to the panel regression models, with a few differences. The control variables of R&D and CapEx intensity had to be removed because *Stata's* statistical software could not estimate their parameters due to many missing values.

Model 1: Basic model

$$ROA_i = \alpha + \beta_1 Marketshare_i + \beta_2 \log(Age_i) + \beta_3 \log(Assets_i) + \beta_4 \left(\frac{1}{Assets}\right) + \beta_5 Dig_i + \beta_6 \log(C_i) + \varepsilon_i \quad (6)$$

Model 2: Basic model + interaction term of size and concentration

$$ROA_i = \alpha + \beta_1 Marketshare_i + \beta_2 \log(Age_i) + \beta_3 \log(Assets_i) + \beta_4 \left(\frac{1}{Assets}\right) + \beta_5 Dig_i + \beta_6 \log(C_i) + \beta_7 \log(Assets_i) * \log(C_i) + \varepsilon_i \quad (7)$$

Model 3: Basic model + interaction term of market share and concentration

$$ROA_i = \alpha + \beta_1 Marketshare_i + \beta_2 \log(Age_i) + \beta_3 \log(Assets_i) + \beta_4 \left(\frac{1}{Assets}\right) + \beta_5 Dig_i + \beta_6 \log(C_i) + \beta_7 Marketshare_i * \log(C_i) + \varepsilon_i \quad (8)$$

Model 4: Basic model but with absolute concentration level instead of its log

$$ROA_i = \alpha + \beta_1 Marketshare_i + \beta_2 \log(Age_i) + \beta_3 \log(Assets_i) + \beta_4 \left(\frac{1}{Assets}\right) + \beta_5 Dig_i + \beta_6 C_i + \varepsilon_i \quad (9)$$

Model 5: Basic model but with squared absolute concentration level instead of its log

$$ROA_i = \alpha + \beta_1 Marketshare_i + \beta_2 \log(Age_i) + \beta_3 \log(Assets_i) + \beta_4 \left(\frac{1}{Assets}\right) + \beta_5 Dig_i + \beta_6 (C_i)^2 + \varepsilon_i \quad (10)$$

Overview of quantile regression equations. Subscript i *denotes firms,* ROA *operating income before depreciation,* α *firm-fixed effects,* Marketshare *firm sales over industry sales,* age *firm age,* Assets *total firm assets,* Dig *the digitization score,* C *the respective concentration metric, and* ε *denotes the error terms.*

3.1.3 Sorted Portfolio Analysis

In line with standard practice in evaluation of the performance of investment portfolios, I chose sorted portfolio analysis as my methodology to check whether stocks with high concentration changes outperform those with low concentration changes in Europe. Like Hou and Robinson (2006) and Grullon et al. (2019) I use risk-corrected average monthly stock returns as the unit of performance measurement. Furthermore, like Grullon et al. (2019, p. 730), and in contrast to Hou and Robinson (Hou & Robinson, 2006) and Gallagher (2015), I use “changes in, rather than the levels of, concentration to capture the aspect of concentration unanticipated by investors”.

First, I form the “bottom change” portfolio containing all firms below the 10th percentile of a given year’s concentration change distribution. Then, I form portfolios consisting of the top 10th percentile, which is called the “top change” portfolio. Next, I obtain monthly returns from *Compustat* daily for all firm-year observations, compute the market returns by taking the average across all firms in the sample of a given year, and finally compute the stocks’ betas by regressing the stocks’ monthly returns on the market’s returns for every given year. Thereby, I obtain 30 yearly betas for every firm in my sample. Next, I obtain average monthly returns by taking their average for every year. To obtain excess returns according to the CAPM, I subtract the year-average monthly market returns multiplied by the firm’s beta. Finally, I compare the time series of the 30 average monthly excess returns of the “bottom change” and “top change” portfolio.

3.1.4 Time-Series Trend Analysis

To examine the general trend in European industry concentration, I plot the time series of the simple averages of firms’ respective concentrations. After visual inspection, I formalize the analysis by comparing averages of subsamples for statistically significant differences by conducting unpaired t-tests with unequal variances.

3.2 Classifications and Metrics

This subsection starts by outlining the most popular measures and classification systems. Then I explain how I classified and measured industries according to their degree of digitization.

3.2.1 Industry Concentration

When analyzing industry concentration, there are two basic questions that need to be answered. What constitutes an industry? And, how do you measure concentration?

3.2.1.1 Classification

My choice of industry classification system is determined by data availability. While *Compustat Global* provides NAICS, historical NAICS and SIC codes, the OECD STAN database for industry output is classified according to ISIC4. Hence, I convert NAICS to ISIC4 codes using the official conversion table from the US government[1]. Table 3.1 provides an overview of services industry codes. The manufacturing industry comprises all ISIC4 2-digit codes from 10 to 33.

NAICS 2-digit code	**ISIC4 2-digit code**
23 Construction	41-43 Construction
44-45 Retail Trade	45-47 Wholesale and Retail Trade
48-49 Transportation and Warehousing	49-53 Transportation and Storage
51 Information	58-63 Information and Communication
54 Professional, Scientific, and Technical Services	69-75 Professional, Scientific, and Technical Activities
55 Management of Companies and Waste Management and Remediation Services	77-82 Administrative and Support Service Activities
61 Educational Services	85 Education
71 Arts, Entertainment, and Recreation	90-93Arts, Entertainment, Recreation
72 Accommodation and Food Services	55-56 Accommodation and Food Service Activities
81 Other Services (Except Public Administration)	94-96 Other Service Activities

Table 3.1: Overview of included services industries. On the left you can see the NAICS 2-digit industry codes, on the right the correspond ISIC4 2-digit code, on which my classification is based.

3.2.1.2 Direct Metrics

When measuring industry concentration, you need to decide on the type of measure and its aggregation level. There are two common types of measures: the market share of leaders and the Herfindahl-Hirschman Index (HHI). Both have advantages and disadvantages and there is not a single best measure. The choice mainly depends on the research purpose and data availability. Then, there is the question whether the relevant market of competition should be a country or an entire world region. For services, country-level measures are more appropriate because services tend to be local. For manufacturing, EU-wide measures are better because manufacturing markets are more international. To maximize validity, I use HHIs, as well as market shares of leaders on the country- and the EU-level. Table 3.2 provides an overview of the used concentration metrics.

Type\Aggregation	Country Level	EU Level
Market Share of Leaders	Country-Level C4	EU-Level C4
HHI	Country-Level HHI	EU-Level HHI

Table 3.2: Overview of concentration metrics.

To construct the HHI, I calculate the ratio of firm sales to total industry sales for every firm-year. Total industry sales are the sum of all firms' sales of a given industry, which may introduce a downward bias

[1] https://www.census.gov/naics/?68967

if the coverage of firms increases over time. Next, I sum the square of this ratio within every two-digit industry-year.

The equation of the HHI metric is:

$$HHI_j = \Sigma_{i \in j} (\frac{s_{ij}}{s_j})^2, \qquad 11$$

where S denotes sales, i the respective firm and j the industry. To obtain the yearly average HHI, I compute the equally weighted average across all firms in the main sample.

The market share of leaders is simply the combined market share of a certain number of leading firms in terms of revenue. When using the market share of leaders, you need to make an additional decision on the level of aggregation. That is, you need to choose whether you want the market share of the four largest, five largest, or sometimes even fifty largest firms in an industry.

For instance, the equation of the C4 measure is:

$$C_j^4 = \Sigma_{i \in s_j^4} \frac{s_{ij}}{s_j}, \qquad 12$$

where S denotes sales, i the respective firm and j the industry. The computation of C5, C8, or C50 metrics works analogously.

3.2.2 Indirect Metrics: Industry Dynamism

Apart from direct industry concentration measures there are two other ways to assess competitive dynamics. First, there is the so-called fallout ratio, i.e., the share of firms who could not keep their position in the top four firms of their industry. Second, I use a reshuffling measure which compares the correlation of the firm rankings within each industry. Both give an indication how the dynamics of competitive forces evolved over time.

3.2.2.1 Fallout Ratio: Measuring the Turnover at the Top

Following Philippon (2019), I compute the likelihood of a top firm to lose its dominating position. Philippon (2019, p. 52) describes the interpretation of the fallout ratio as follows: “given that a firm is at the top of its industry now—among the top four by profits or by market value—how likely is it that it will drop out over the next three years?”

To construct this measure, I rank the firms within each industry-year according to their sales and create a dummy variable that takes on the value of one if the firms’ ranking position in a given year is greater than four and its rank three years prior was within the top four. Then, I compute the average of each firms’ indicator over every year and obtain a time-series of the fallout ratio for the entire period of my sample.

3.2.2.2 Reshuffling
Another way to measure within-industry dynamism is to compare the changes in the market-share ranking of firms. To do this, I compute the correlation between the ranking positions of the top four firms within each industry of two different years. The reshuffling measure is defined as one minus the respective correlation coefficient. For instance, if the correlation between the ranking of 1995 and 2000 was 0.8, then the reshuffling measure would take on the value of 0.2. The higher, the more dynamic the competition within an industry.

3.2.3 Digitization
First, I present characteristics of digitization. Then I outline its metrics.

3.2.3.1 Classification
Calvino et al. (2018) identified three main manifestations of digitization in companies: technology, human capital and output market behavior. The first pillar, technology, relates to investment in information and communications technology (ICT). In other words, the first manifestation of higher digitization is increased capital expenditure to create digital infrastructure within a company. This can encompass investment in goods, services, or even intermediary products, such as software as a service. The second pillar of digitization is the necessary human capital to implement digitization. Calvino et al. (2018) measure this component with the share of ICT specialists employed. The final pillar is firms' output market behavior, which is measured by the ratio of online sales to total sales of a company (Calvino et al., 2018).

3.2.3.2 Metrics
Calvino et al. (2018) classify the digital intensity of ISIC4 2-digit industries into four groups: low, medium-low, medium-high, and high. Based on this, I create a digitization score that takes on the value of one if a firm's industry falls into the category of low digital intensity, two for medium-low industries, three for medium-high and four for high.

For the subsample analysis of industry concentration along the digitized/non-digitized split in part three, I use an industry-level dummy which I construct using a list of digital industries (Nicholson, 2020). The working paper of the US Bureau of Economic Analysis (BEA) lists all the goods and services six-digit NAICS industries included in the estimates of the size of the digital economy. In particular, it provides a comprehensive list of over 250 six-digit NAICS codes of industries that are considered as part of the digital value creation. Importantly, it does not only consist of sub-industries of the traditional information and communications technology sectors but comprises a much wider range of micro-industries.

To construct my dummy, I convert the six-digit NAICS to four-digit ISIC codes and compute the share of them for each ISIC-two-digit industry. The resulting values are between 0 and 1, the mean being equal to 0.12. My final digitization dummy takes on the value of one if the firm's share of four-digit sub-industries is above mean, i.e., greater than 0.12.

3.3 Methodological Pitfalls and Limitations

Bajgar et al. (2023) identified two issues that need to be addressed when computing concentration measures. First, they observed that coverage of firms in *Orbis* varies over time. In *Orbis* coverage of small firms tends to increase. This artificially lowers the concentration metric since the denominator (total industry sales) increases progressively over time. The extent to which this is an issue depends on the size of the additional revenue of the added small firms. Moreover, a decile-based top-market share approach is more susceptible to this pitfall than an absolute concentration metric. For instance, the market share of the eight largest firms is more robust than the one of the top 10 per cent. This is due to the fact that with the percentile-based measure more small firms would pass the threshold of 10 precent, whereas with the absolute threshold the eight largest firms remain most likely unchanged. The solution to this problem is using, for instance, the sum of the sales of the four largest firms of an industry as the numerator and total industry sales from a reliable industry-level database as the denominator. This should effectively address the downward bias caused by increasing coverage (Bajgar et al., 2023). Therefore, I used census-based industry output from *OECD STAN* as the denominator in the country- and EU-level C4 concentration metrics of the *Compustat* data set.

Another distortion of results may come from neglecting subsidiary linkages when analyzing concentration on the world-region level. For concentration within countries, it is unlikely that there is a significant amount of business groups with several subsidiaries active in the same industry and same country. Bajgar et al. (2023) conclude that using business group *Orbis* data results in an upward bias of concentration while using individual firm's data leads to a downward bias. They propose an industry-business-group matching, which is basically a combination of the two approaches. However, because the authors did the matching manually and did not disclose their outcome comprehensively, this approach is beyond the scope of my paper. Hence, I use individual firm's data.

3.4 Main Hypotheses

My research aims to address the conflicting findings on European industry concentration and its understudied impact on profitability and stock returns by constructing and comparing several different concentration measures for numerous European countries. The different antitrust measures and competition policies of the US and the EU have probably led to disparate outcomes. The main hypotheses can be summarized as follows:

1. **Part 1: Financial Performance - ROA Analysis**
    a. General Trend: Gap between 90th percentile ROA and median increased during the last two decades.
    b. Panel Regressions: Industry concentration has a positive impact on ROA.
    c. Quantile Regressions: Industry concentration and digitization have a positive impact on ROA and stronger so for the top percentile of ROA.

2. **Part 2: Stock Performance – Sorted Portfolio Analysis**
    a. Top-concentration-change firms have higher excess returns than bottom-concentration-change firms.

3. **Part 3: General Trend in European Industry Concentration**
    a. **Average Absolute Concentration Levels**
        i. The average absolute level of European industry concentration fell until the late 1990s and has been rising since then.

        ii. There is a significant difference between the average industry concentration of European manufacturing and services firms.

        iii. Mean absolute concentration of digital industries is significantly higher than that of non-digital industries.

    b. Average Changes in Concentration Levels

        i. The relative change in the level of European industry concentration was negative until the late 1990s and has been positive since then.

        ii. There is a significant difference between the average change in industry concentration of European manufacturing and services firms.

        iii. The average change in concentration of firms belonging to digital industries is significantly higher than of those pertaining to non-digital industries.

    c. **Reshuffling Rate:**
        i. The average reshuffling rate of European firms was rising until the late 1990s and has been falling since.

        ii. There is a significant difference between the reshuffling rate of European manufacturing and services firms.

        iii. Mean reshuffling rates of European firms of digital industries are significantly higher than those of non-digital industries.

    d. **Fallout Rate:**
        i. The average fallout rate of European firms fell until the late 1990s and has been rising since.

        ii. There is a significant difference between the mean fallout rates of European manufacturing and services firms.

        iii. Mean fallout rates of European firms belonging to digital industries are significantly higher than those belonging to non-digital industries.

4 Data

In this section, I first explain the rationale behind the sample selection and then present descriptive statistics of the main variables of interest. I use two datasets. One of firm-level data from *Compustat* and one with data from *Orbis*.

4.1 Selection Criteria for Inclusion in Main Sample

One of the aims of my research is to come as close as possible to capturing the entirety of the European economy. Therefore, as many regions and as many years as possible are included in my dataset.

4.1.1 Compustat

The main sample consists of firm- and industry-level data for 20 European countries from 1987 to 2019. The time frame and country selection are mainly driven by data availability. There is a tradeoff between having a balanced panel and as large a sample as possible. Ultimately, I restricted the choice of covered countries and years such that all variables of the main analysis were available for every included country and time point.

Before 1990 many Eastern European countries were part of the Soviet Union and did not have independent economies. After the collapse of the Soviet system, their economies were often not properly developed until the early 2000s, Romania and Bulgaria being a case in point. Hence, I excluded the following countries with a negligible economy prior to 2000 from my sample: Bulgaria, Croatia, Lithuania, and Romania. The fact that the OECD census database also omits these countries confirms my choice. After 1990, most exchange rate and firm-level data is available.

Table 2 below shows the sources of additional data for the construction of the *Compustat* sample. Because *Compustat* reports all values in national currency, I had to convert them all into a common currency, US dollars. I downloaded the respective exchange rates from *fxtop.com* and converted according to *Compustat's* currency codes. Further, I downloaded shares outstanding and stock prices of the entire *Compustat Global Daily* database, to compute market capitalization and stock returns. For the coverage-robust *Compustat* C4 industry concentration metrics, I obtained country-level industry output from the *OECD STAN* database.

Data Type	Source
Balance Sheet Data	Compustat Global Fundamentals Annual
Data for Market Capitalization	Compustat Global Daily
Census-Based Industry Sales	OECD STAN
Exchange Rates	fxtop.com

Table 4.1: Overview of data sources.

4.1.2 Orbis

Construction of the *Orbis* sample was easier than for *Compustat* because all values are already reported in dollars. However, *Orbis* only provides data for nine years prior to the last available date for a given firm.

Inclusion criteria are incorporation in one of the 27 EU countries, and available balance sheet data for at least one of the nine years prior to the latest available year. All industries were included except ISIC4 2-digit codes 64 (Financial service activities, except insurance and pension funding), 65 (insutrance, reinsurcance and pension funding) and 66 (activities auxiliary to financial services and insurance activities).

4.2 Data Cleaning and Construction

First, I outline the procedure for *Compustat*, then for *Orbis*.

4.2.1 Compustat

Compustat dataset construction involves cleaning and unifying the raw dataset obtained from *Compustat Global Fundamentals Annual* for European firms. The initial step involves excluding non-European countries and creating two-digit ISIC4 industry codes by merging the dataset with a conversion table. Next, missing and nonsensical values are dropped, including observations without industry codes or negative sales. The balance sheet items are then converted into dollars using exchange rates from *fxtop.com*. Digitization and zombie lending dummies are added based on industry classifications and financial indicators. Then, I compute HHI industry concentration and the year-on-year HHI concentration changes. To construct the C4 concentration measures I merge the OECD STAN industry output dataset, which reduces the amount of observations substantially. The obtained C4 concentration metrics are winsorized on both ends of the distribution by 0.5% to address outliers. I further add the number of firms per year and drop ancillary and redundant variables to create the main dataset.

4.2.2 Orbis

Cleaning of the *Orbis* sample was restricted to removing duplicates and missing values of key variables, such as ISIN or sales data. Winsorization of ROA was not necessary, because there were no vastly aberrant observations.

4.3 Data Description

First, I will present general summary statistics and then outline the key features of the main variables of interest in more detail.

4.3.1 Compustat

As you can see in the table below, industry codes range from 1 to 96, covering the entirety of industries (except financial industries). The dataset covers the time period 1989 to 2019, average concentration metrics range from 0.029 to 0.483 (in fraction form), and firm age ranges from 0 to 149 years. The average firm has $2.6b in assets, $2.2b in sales, and an ROA of 5.9%. The minimum number of firms per year in the dataset is 60 and the maximum 2311, indicating a time-varying coverage. The average ratio of R&D expenditure to assets is 7.7% and the average ratio of capital expenditure to assets is 4.7%. Average industry sales are $83b.

Variable	Obs	Mean	Std. Dev.	Min	Max
industry code (ISIC4, 2 digits)	25278	37.898	18.996	1	96
year of observation	25278	2009.182	6.8	1989	2019
c4 country	25278	.315	.439	0	2.189
c4 eu	25278	.029	.105	0	2.734
HHI country	25278	4826.333	2841.768	0	10000
HHI eu	25278	1410.532	1223.468	347.638	10000
dollar at	25278	2602.503	11450.066	0	297871.72
dollar sale	25278	2164.759	8807.392	0	235807.63
ROA	25229	.059	.187	-1.167	.456
RnD	10032	.077	.138	-.084	3.524
CapEx	22602	.047	.098	-.064	10.298
age	25278	12.986	12.82	0	149
dollar industry sales	25278	83333.453	121192.99	58.168	2646001.5
firm count	25278	1874.293	425.948	60	2311

Table 4.2: Compustat dataset: Descriptive statistics of key variables. Monetary values are in million USD. c4 country *is the country-level market share of the four largest companies,* c4 eu *is the EU-level market share of the four larges companies,* HHI country *is the country-level HHI,* HHI eu *is the EU-level HHI,* dollar at *are total assets converted into dollars,* dollar sale *is total sales converted into dollars,* ROA *is return on assets, defined as operating income (Compustat item OIBDP) scaled by total assets,* RnD *is the research and development intensity (R&D expenditure scaled by assets),* CapEx *is the capital investment intensity (capital investment scaled by assets),* age *is the maximum of the time in the sample or time from the indicated IPO date,* dollar industry sales *is the sum of all firm sales reported by Compustat for a given industry-year,* firm count *is the number of unique firms per year.*

Figure 4.1 below shows that average ROA, the dependent variable in my regression models, decreased over time. Decomposing average ROA into its regional components in figure 4.2 reveals that it decreased for Northern, Southern, and Eastern European firms in a similar way. Industry concentration, however, seems to be significantly different, with Southern Europe having the highest average country-level concentration, followed by Northern and Eastern Europe, as you can see in figure 4.3. Figure 4.4 shows that the coverage of *Compustat* improved until 2000 and has remained roughly constant since then.

Figure 4.1: Time-series of average ROA. Time is plotted on the horizontal axis, average ROA (across all firms) is on the vertical axis.

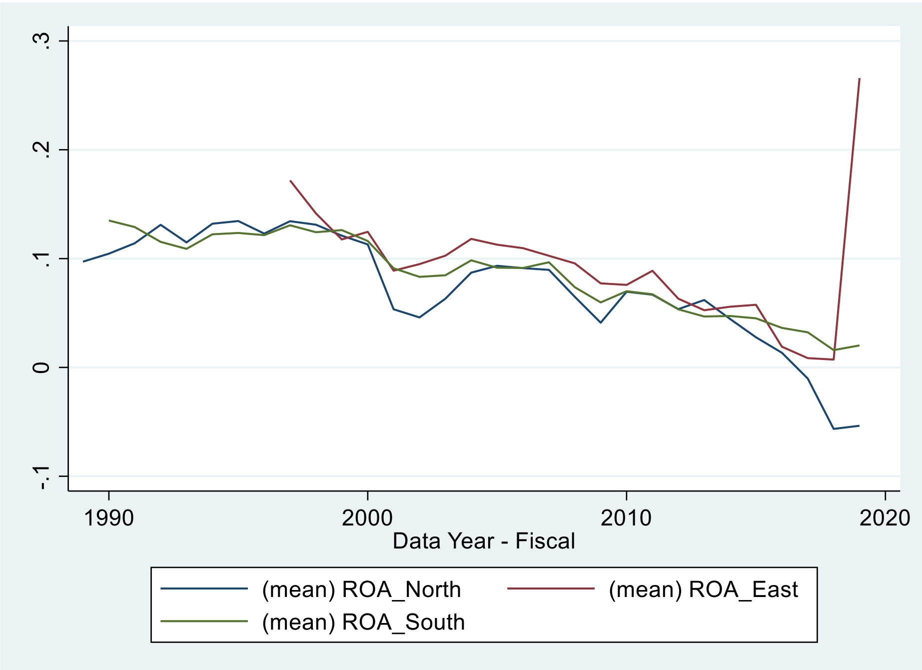


Figure 4.2: Time-series of average ROA for regional subsamples (Northern, Southern, and Eastern Europe). Time is plotted on the horizontal axis, average ROA (across all firms) is on the vertical axis.

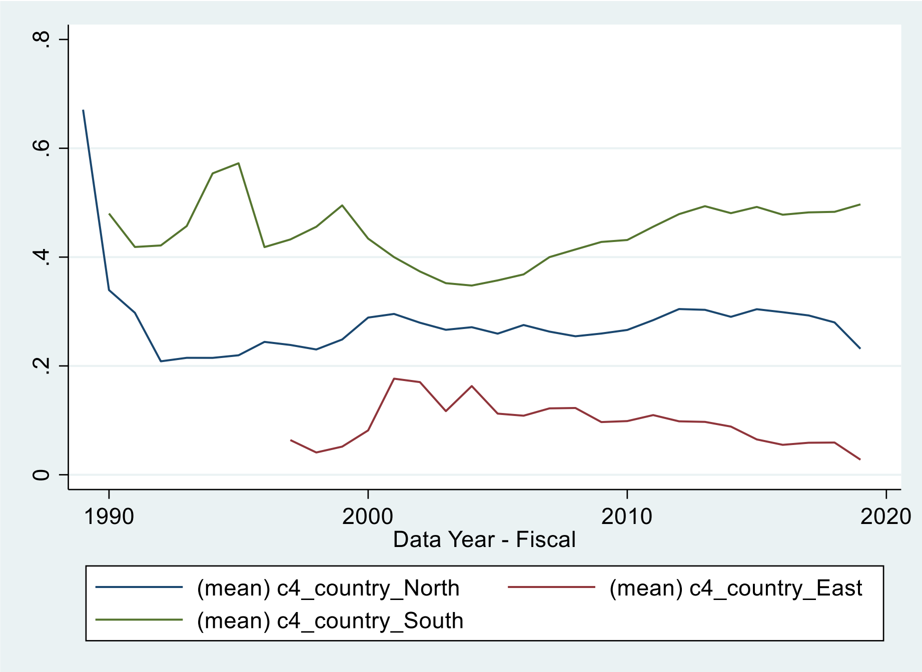


Figure 4.3: Time-series of average country-level C4 industry concentration for regional subsamples (Northern, Southern, and Eastern Europe). Time is plotted on the horizontal axis, average country-level market share of the four largest companies (across all firms) is on the vertical axis.

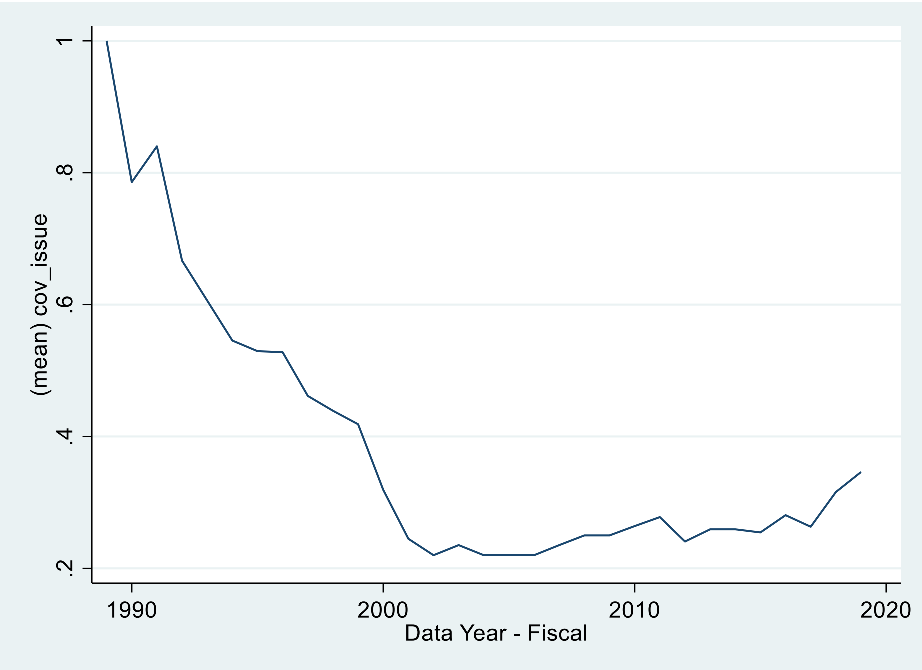


Figure 4.4: Time series of the average share of industries for which there are less than five reported firms in the dataset. Time is plotted on the horizontal axis, average share of industries with less than five reported firms is on the vertical axis.

4.3.2 Orbis

Industry codes range from 1 to 99, covering the entirety of industries (except financial industries). Codes 97 to 99 are very rare, however, since they refer to activities of households as employers and activities of extraterritorial organizations and bodies. The dataset covers the time period 1994 to 2022, average concentration metrics range from 0.23 to 0.935 (in fraction form), and firm age ranges from 0 to 357 years. The average firm has \$3.4b in assets, \$2.4b in sales, and an ROA of 3.3%. The minimum number of firms per year in the dataset is one and the maximum 6611, indicating a time-varying coverage as well.

Variable	Obs	Mean	Std. Dev.	Min	Max
industry code (ISIC 4, 2 digits)	178031	45.368	24.229	1	99
year of observation	178031	2008	8.367	1994	2022
c4 country	145469	.935	.106	.392	1
c4 eu	176274	.688	.182	.296	1
HHI country	145469	.566	.289	.069	1
HHI eu	176274	.23	.199	.035	1
dollar at	87071	3366468.3	17408213	.574	$6.100 \cdot 10^8$
Sales	84799	2394230.9	12405275	-14826.799	$4.702 \cdot 10^8$
ROA	83382	.033	1.536	-379.667	37.553
age	132021	34.791	39.775	0	357
firm count	176059	3322.505	1907.945	1	6611

Table 4.3: Orbis dataset: Descriptive statistics of key variables. Monetary values are in thousand USD. c4 country *is the country-level market share of the four largest companies,* c4 eu *is the EU-level market share of the four largest companies,* HHI country *is the country-level HHI,* HHI eu *is the EU-level HHI,* dollar at *are total assets,* Sales *is the amount of total sales,* ROA *is return on assets, defined as operating income scaled by total assets,* age *is the maximum of the time in the sample or time from the indicated date of foundation,* firm count *is the number of unique firms per year.*

The time-series of average ROA in figure 4.4 reveals a decreasing trend in *Orbis* too, with a shock in 2001, which is probably due to the bursting of the dot-com bubble and the impact of the 9/11 attack. Similar average ROAs of Northern, Southern, and Eastern Europe in figure 4.5 confirm the trend of *Compustat*. Figure 4.6 shows that average industry concentration is also in *Orbis* different for the respective regions, with Southern Europe having again the highest average concentration metrics. Figure 4.7 shows that also *Orbis* seems to have coverage issues, since the share of industries with less than five reported firms fell dramatically until approximately the year 2000.

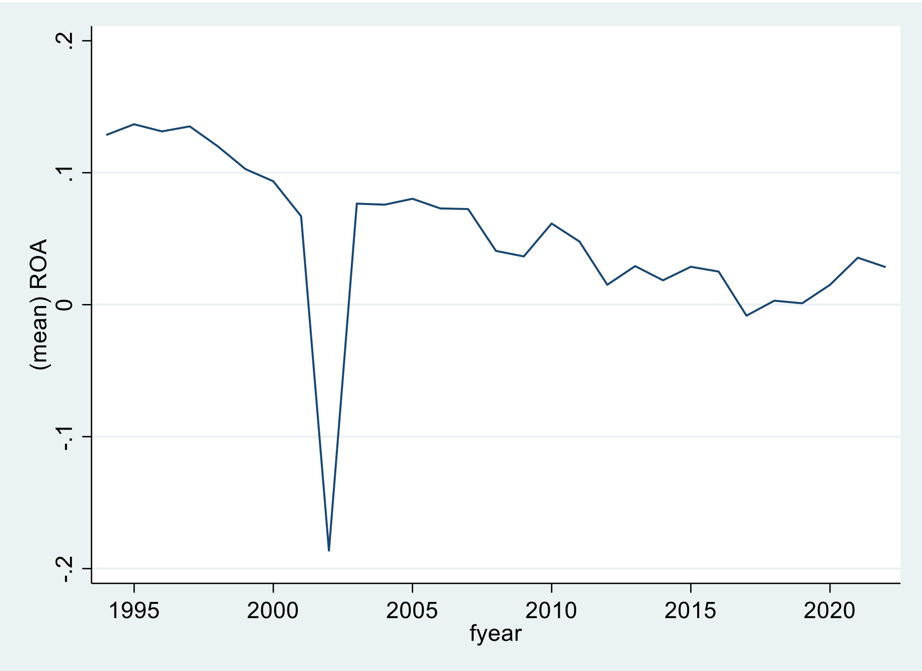


Table 4.4: Time-series of average ROA. Time is on horizontal axis, average ROA (across firms) is on the vertical axis.

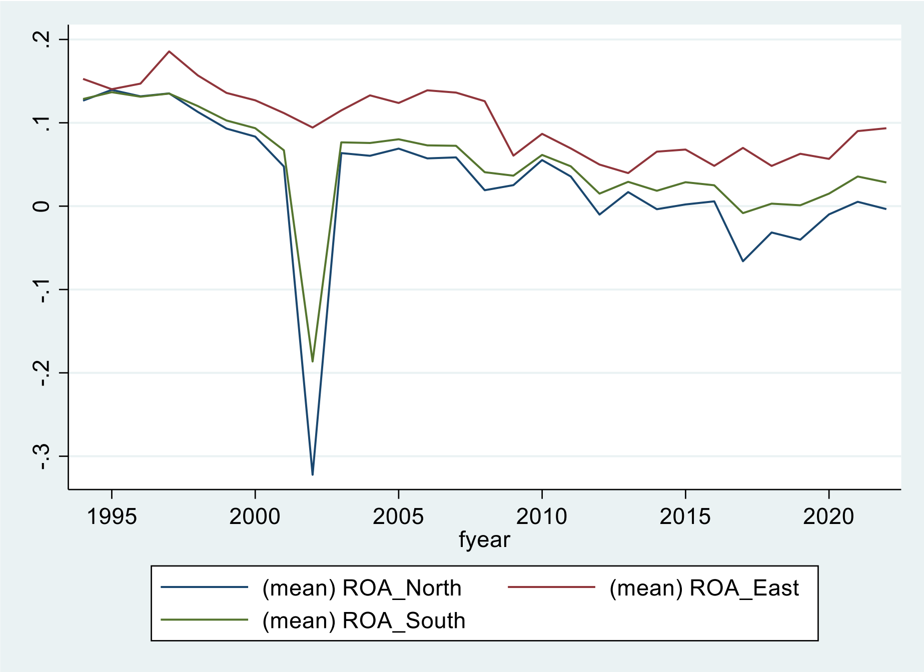


Table 4.5: Time-series of average ROA for regional subsamples (Northern, Southern, and Eastern Europe). Time is on horizontal axis, average ROA (across firms) is on the vertical axis.

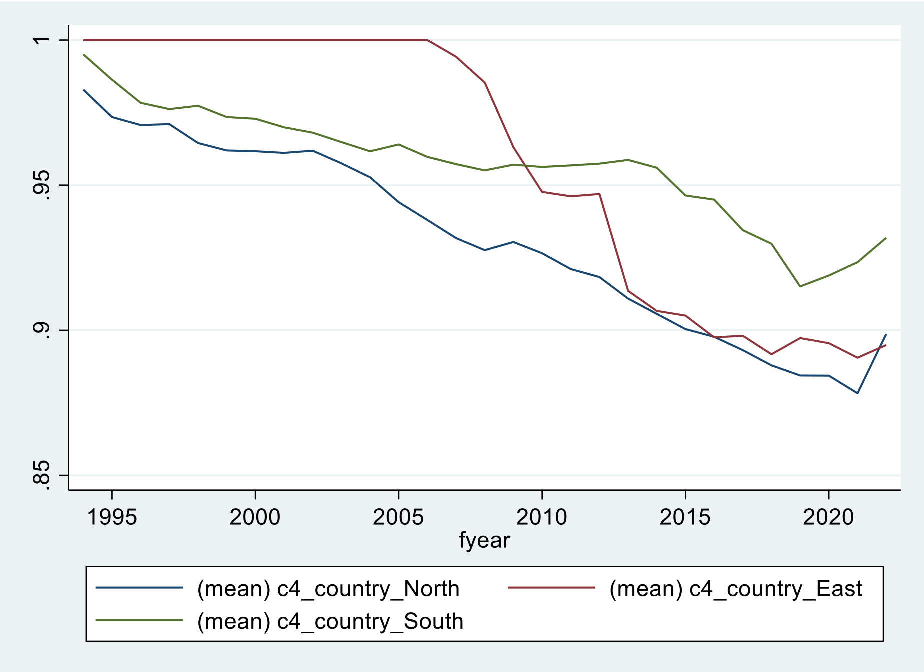


Table 4.6: Time-series of average country-level C4 industry concentration for regional subsamples (Northern, Southern, and Eastern Europe). Time is on horizontal axis, average market share of the four largest companies (across firms) is on the vertical axis.

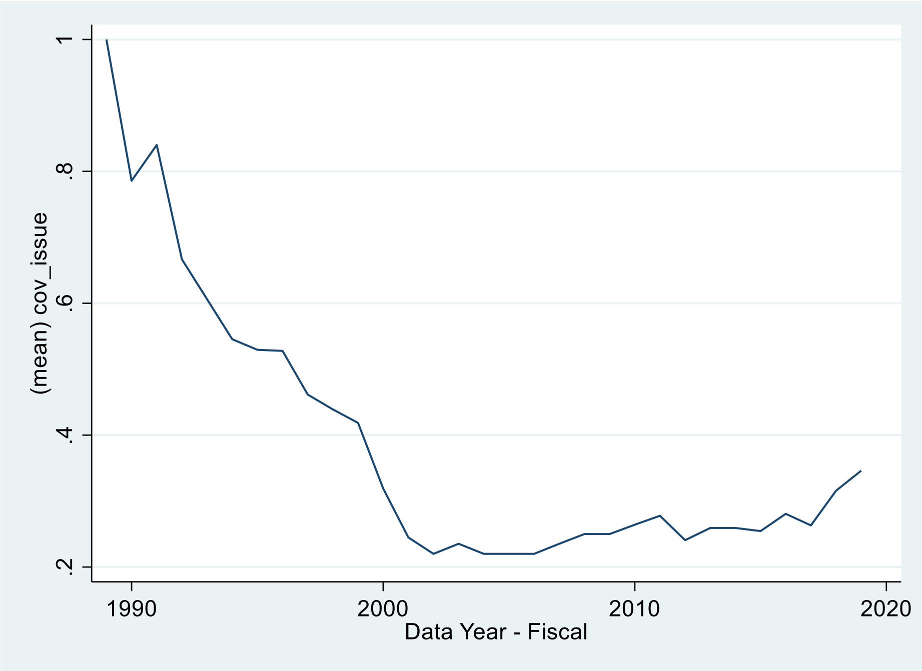


Table 4.7: Time series of the share of industries for which there are less than five reported firms in the dataset. Time is on the horizontal axis, the share of industries with less than five reported firms is on the vertical axis.

5 Main Results

First, I present the results on the impact of industry concentration and digitization on ROA. Second, I analyze the impact of industry concentration on stock returns. Third, I show the general trend of European industry concentration.

5.1 Impact of Concentration and Digitization on ROA

This subsection starts with the panel regression results, before outlining the cross-sectional quantile regression results.

5.1.1 Panel Regressions

In contrast to findings for America (Grullon et al., 2019), the log of European industry concentration is mostly negatively related to profitability in Europe for three out of four concentration metrics of the *Compustat* data set. The interaction term of the log of assets and log of industry concentration is insignificant on the country-level but positive and significant at the EU level. The interaction term of market share and the log of industry concentration is only with the EU-level C4 measure negative and significant.

	C4 Country		C4 EU		HHI Country		HHI EU	
	Range of Estimates	Significance						
Market Share	- 0.022 to - 0.011	ins. - **	0.004 to 0.096	ins. to ***	0.050 to 0.137	ins. - ***	-0.022 to 0.073	ins.
Age	0.015 to 0.016	ins. - *	0.010 to 0.015	ins.	0.013 to 0.015	ins.	0.013 to 0.015	ins.
Size	0.001 to 0.007	ins.	0.007 to 0.043	ins. - ***	-0.082 to 0.004	ins. - *	-0.061 to 0.007	ins. - *
Inverse Size	-0.384 to - 0.381	***	-0.373 to - 0.384	***	-0.388 to -0.374	***	-0.390 to -0.383	***
R&D Exp.	-0.532	***	-0.531 to - 0.533	***	-0.532 to -0.537	***	-0.526 to -0.532	***
CapEx	0.182 to 0.187	***	0.182 to 0.188	***	0.188 to 0.195	***	0.186 to 0.189	***
Log of Industry Concentration	0.020 to 0.033	* - ***	-0.073 to - 0.027	***	-0.101 to -0.043	***	-0.092 to -0.033	***
Concentration x Size	-0.002	ins.	0.007	***	0.010	**	0.010	**
Concentration x Market Share	0.008	ins.	-0.043	**	-0.007	ins.	0.031	ins.
Observations	9.536		9.536		9.536		9.536	
R^2	0.221 to 0.224		0.220 to 0.230		0.226 to 0.229		0.221 to 0.224	

*** p<0.01, ** p<0.05, * p<0.1, ins. insignificant

Table 5.1: Overview of Compustat panel regression results. Regression of ROA on the variables listed in the first column. The variables in the first row indicate the applied concentration metric: C4 Country *is the country-level market share of the four largest companies within an industry,* C4 EU *is the EU-level market share of the four largest companies within an industry,* HHI Country *is the country-level HHI,* HHI EU *is the EU-level HHI. The first column of each concentration metric reports the range of parameter estimates of the regressors across the five models outlined in the methodology section. The second column within each concentration metric indicates the range of significance. Positive coefficients are in green color, negative ones in red. Detailed results can be found in the appendix.*

	C4 Country		C4 EU		HHI Country		HHI EU	
	Range of Estimates	Sig.						
Market Share	0.058 to 0.064	***	0.037 to 0.104	ins.	0.052 to 0.086	** - ***	0.046 to 0.117	ins.
Age	0.024 to 0.030	***	0.021 to 0.030	***	0.024 to 0.029	***	0.024 to 0.030	***
Size	0.034 to 0.042	ins. - *	0.036 to 0.053	ins. - **	0.034 to 0.049	ins. - **	0.035 to 0.060	ins. - ***
Inverse Size	-835.559 to -831.214	***	-834.939 to -827.783	***	-835.924 to -834.435	***	-834.958 to -827.907	***
R&D Exp.	0	**	0	**	0	**	0	**
Log of Industry Concentration	-1.024 to -0.022	ins. - **	-0.415 to 0.016	ins. -***	-0.228 to - 0.014	ins . - **	-0.136 to 0.011	ins.- **
Concentration x Size	0.077	**	0.032	***	0.017	**	0.011	***
Concentration x Market Share	0.199	ins.	0.558	**	-0.002	ins.	0.098	ins.
Observations	19.775		19.775		19.775		19.775	
R^2	0.260 to 0.263		0.260 to 0.263		0.260 to 0.263		0.260 to 0.262	

*** p<0.01, ** p<0.05, * p<0.1, ins. insignificant

Table 5.2: Overview of Orbis panel regression results. Regression of ROA on the variables listed in the first column. The variables in the first row indicate the applied concentration metric: C4 Country *is the country-level market share of the four largest companies within an industry,* C4 EU *is the EU-level market share of the four largest companies within an industry,* HHI Country *is the country-level HHI,* HHI EU *is the EU-level HHI. The first column of each concentration metric reports the range of parameter estimates of the regressors across the five models outlined in the methodology section. The second column within each concentration metric indicates the range of significance. Positive coefficients are in green color, negative ones in red. Detailed results can be found in the appendix.*

The *Orbis* data set confirms the negative relationship between European ROA and the log of industry concentration. The positive impact of the interaction term of size and concentration is strengthened, with consistently positive coefficients at the 5% confidence level. However, for the interaction term of market share and concentration, only two of the six panel regressions with *Orbis* result in significant positive coefficients, strengthening the insight from the *Compustat* data set that this interaction term is insignificant.

Regarding the control variables, inverse firm size is clearly negatively related to ROA, in line with previous empirical findings (Goddard et al., 2005; Pattitoni et al., 2014). *Compustat* data also finds consistently highly significant results for a negative impact of R&D expenditure and a positive impact of CapEx. Unfortunately, *Orbis* has such limited data on both variables that confirmation is not possible. However, *Orbis* data strengthened the previously weakly significant positive impact of market share and age, the latter being almost always strongly significant at the 1% level. Detailed results and all regression tables can be found in the appendix.

5.1.2 Quantile Regressions

To check whether highly profitable firms are more sensitive to industry concentration changes, I run a series of cross-sectional quantile regressions. Furthermore, I include an industry-level digitization score as a regressor to see whether the degree of industry digitization has a systematic impact on profitability. The results are as follows.

5.1.2.1 Industry Concentration

Table 5.3 provides a qualitative overview of the sign and direction of the respective concentration coefficients.

Parameters Pre 2000	C4 Country		C4 EU		HHI Country		HHI EU	
Model	Compustat	Orbis	Compustat	Orbis	Compustat	Orbis	Compustat	Orbis
M1	+	+	+	inc.	-	inc.	inc.	+
M2	+	-	+	inc.	-	-	-	-
M3	+	-	inc.	inc.	-	+	inc.	inc.
M4	+	+	-	+	inc.	inc.	inc.	-
M5	+	+	-	+	inc.	inc.	inc.	-

Parameters Post 2000	C4 Country		C4 EU		HHI Country		HHI EU	
Model	Compustat	Orbis	Compustat	Orbis	Compustat	Orbis	Compustat	Orbis
M1	-	-	-	inc.	-	inc.	inc.	+
M2	+	-	inc.	inc.	-	-	-	+
M3	-	+	inc.	inc.	-	-	inc.	inc.
M4	-	-	+	+	inc.	inc.	inc.	+
M5	-	-	inc.	+	inc.	inc.	inc.	+

Trend	C4 Country		C4 EU		HHI Country		HHI EU	
Model	Compustat	Orbis	Compustat	Orbis	Compustat	Orbis	Compustat	Orbis
M1	↓	↓	↓	inc.	↑	→	inc.	→
M2	↓	↑	↓	inc.	↑	↑	↑	↑
M3	↓	↑	inc.	inc.	↑	↓	inc.	→
M4	↓	↓	↑	→	inc.	inc.	inc.	↑
M5	↓	↓	↑	→	inc.	inc.	inc.	↑

Table 5.3: Qualitative overview of sign and trend of concentration coefficients of Compustat and Orbis quantile regressions. The first two tables report the sign of industry concentration parameters that most cross-sectional quantiles pre-2000 (table 1) and post-2000 (table 2) take on. A green plus indicates that most quantiles reported positive industry concentration parameter estimates. A red minus indicates that most quantiles reported negative industry concentration parameter estimates. Inc *indicates inconclusive results, i.e., some quantile estimates were positive and some negative during the period from 1994 to 1999 (table 1) and 2000 to 2022 (table 2). Results are reported for each model specification and used concentration metric. The second two tables report the trend of the industry concentration parameters that most cross-sectional quantiles pre-2000 (table 1) and post-2000 (table 2) take on. A green arrow indicates that most quantiles reported a positive industry concentration parameter trend. A red arrow indicates that most quantiles reported a negative industry concentration parameter trend.* Inc *indicates inconclusive results, i.e., some quantile estimates exhibit a positive and some a negative trend during the period from 1994 to 1999 (table 1) and 2000 to 2022 (table 2). Results are reported for each model specification, used concentration metric, and dataset.*

Pre-2000, 14 specifications find a positive, 12 a negative, and 13 an inconclusive relationship between industry concentration and ROA. Post-2000, only 9 specifications find predominantly positive coefficients, 15 negative and 16 inconclusive ones. Regarding the time-series trend of coefficients, 12 specifications report an increase, 11 a decrease, 12 an inconclusive and five a constant trend in the magnitude of the significant coefficients. Overall, these qualitative findings do not reveal a systematic pattern.

To check whether industry concentration levels have stronger effects for more profitable firms, I conduct paired t-tests of the time series of the concentration coefficients of the 10^{th} and 90^{th} percentiles. Of the 20 *Compustat* model specifications, 15 coefficients of the ninetieth percentile have a significantly higher average estimate. With the *Orbis* data set, I obtained 14 significantly higher average 90^{th} percentile estimates, which indicates that the mostly negative impact of the level of industry concentration on ROA is less in magnitude for highly profitable firms than for those with poor profitability. For the model specifications that result in positive estimates, the interpretation of the positive difference between the 10^{th} and 90^{th} percentiles is that higher industry concentration leads to higher profitability.

5.1.2.2 Digitization Score Coefficients

The previous section showed that the sign and direction of the industry concentration coefficient varies substantially across models and data sets. However, the coefficient of the digitization score reveals a clear pattern. For the *Compustat* sample, it is significantly negative pre-2000, increasing, and consistently positive post-2000. The Orbis dataset's digitization score coefficient is positive and its time-series trend constant across all models and concentration measures. Therefore, industry-level digitization has a positive influence on European ROAs, strengthening the previously fragmented empirical evidence (Broccardo et al., 2023; Kohtamäki et al., 2020). A comprehensive graphical overview of the results can be found in the appendix.

Measure	C4 Country			C4 EU			HHI Country			HHI EU		
Model	Pre 2000	Trend	Post 2000	Pre 2000	Trend	Post 2000	Pre 2000	Trend	Post 2000	Pre 2000	Trend	Post 2000
M1	- (+)	↑ (→)	+ (+)	- (+)	↑ (→)	+ (+)	- (+)	↑ (→)	+ (+)	- (+)	↑ (→)	+ (+)
M2	- (+)	↑ (→)	+ (+)	- (+)	↑ (→)	+ (+)	- (+)	↑ (→)	+ (+)	- (+)	↑ (→)	+ (+)
M3	- (+)	↑ (→)	+ (+)	- (+)	↑ (→)	+ (+)	- (+)	↑ (→)	+ (+)	- (+)	↑ (→)	+ (+)
M4	- (+)	↑ (→)	+ (+)	- (+)	↑ (→)	+ (+)	- (+)	↑ (→)	+ (+)	- (+)	↑ (→)	+ (+)
M5	- (+)	↑ (→)	+ (+)	- (+)	↑ (→)	+ (+)	- (+)	↑ (→)	+ (+)	- (+)	↑ (→)	+ (+)

Table 5.4: Qualitative overview of sign and trend of digitization score coefficients of Compustat (Orbis) quantile regressions. The column Pre 2000 *reports the dominant sign of the digitization dummy parameter estimate in cross-sectional quantile regressions of ROA on digitization, industry concentration, and control variables pre-2000. Compustat estimate signs are reported first, after which Orbis signs are reported in brackets. The second column* Trend *indicates the direction over time of the cross-sectional digitization parameter estimates. A green arrow indicates an increasing estimate over time, a black arrow a rather constant trend. The third column* Post 2000 *reports the dominant sign of the digitization dummy parameter estimate in cross-sectional quantile regressions of ROA on digitization, industry concentration, and control variables post-2000. Green shading indicates matching Orbis and Compustat results. Results are reported for each model specification and used concentration metric.*

To check whether higher industry-level digitization leads to higher ROAs of more profitable firms, I conduct paired t-tests of the time series of the coefficients of the 10th and 90th percentiles. Of the 20 *Compustat* model specifications, 11 coefficients of the ninetieth percentile have a significantly higher average estimate. The mean estimate of the 90th percentile is higher than that of the 10th in all cases. With the *Orbis* data set, all mean coefficient estimates of the 10th percentile range from 0.003 to 0.005, while those of the 90th are between 0.015 and 0.017. All differences are significant at the 1%-level, confirming that being in a highly digitized industry leads to a stronger increase in ROA within the group of top-profitability firms than within that of bottom-profitability ones.

5.2 Portfolio Sorted Excess Stock Returns

The statistical significance of the difference between the average excess return of the top 10th percentile in terms of concentration change and the one of the bottom percentiles varies across concentration metrics and data sets. The *Compustat* data set indicates higher average bottom-percentile excess returns, while the *Orbis* data set results in higher average excess returns of the top-concentration-change portfolio.

The graphical overview below shows that the time series of average bottom-concentration-change percentile excess returns of the *Compustat* data set with its 2251 firms, probably contains outliers. For instance, using the *Compustat* C4 EU metric results in an average excess return of roughly 130% in 2019 for the bottom-percentile portfolio.

The results of the *Orbis* data set seem more reasonable. For one thing, it contains approximately twice as many firms (5215). For another, it appears to be free of outliers and resembles commonly found average excess stock returns. Here, three out of four concentration metrics results in higher excess returns for the top-concentration-change portfolio. Using the country-level C4 concentration metrics delivers a significant difference of the means at the 1%-level. The country-level HHI results in significance at the 10% level. The EU-level measures yield insignificant differences. Overall, the heterogeneous results across market definitions call for more refined analysis, by matching industries with the appropriate geographical market definition.

Compustat	Mean Bottom Return	Mean Top Return	p-value $H_A: mean(bottom) < mean\ (top)$
C4 Country (N=30)	0.023	0.012	0.862
C4 EU (N=28)	0.048	0.008	0.799
HHI Country (N=30)	0.029	0.016	0.787
HHI EU (N=30)	0.036	0.003	0.940

Orbis	Mean Bottom Return	Mean Top Return	p-value $H_A: mean(bottom) < mean\ (top)$
C4 Country (N=25)	0.016	0.037	0.017 ***
C4 EU (N=25)	0.017	0.029	0.174
HHI Country (N=25)	0.025	0.043	0.087 *
HHI EU (N=25)	0.044	0.023	0.876

**** p<0.01, ** p<0.05, * p<0.1*

Table 5.5: Mean monthly returns of concentration-change sorted portfolios. First table: Compustat. Second table: Orbis. The first column indicates the applied concentration metric, the second column reports the mean monthly return of the bottom 10% of firms sorted by concentration change each year, the second column reports the mean monthly return of the top 10% of firms sorted by concentration change each year. The third column indicates the p-value of statistically significant differences in the means.

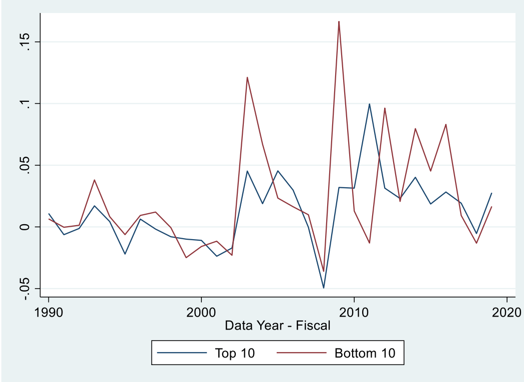
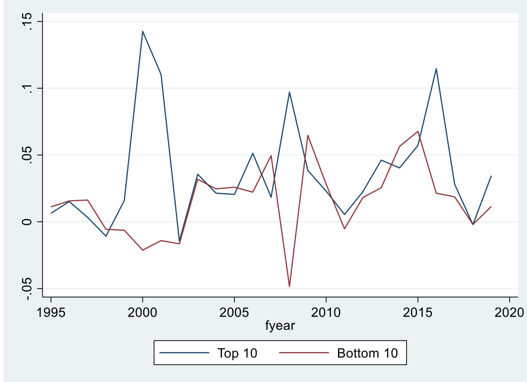
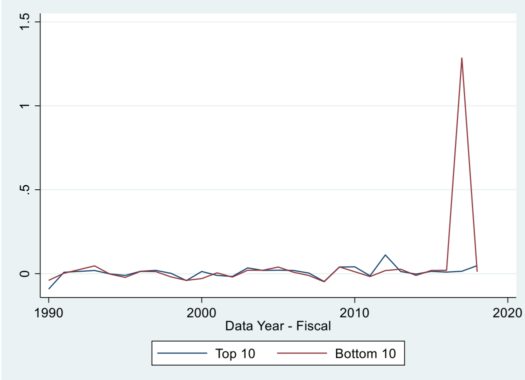
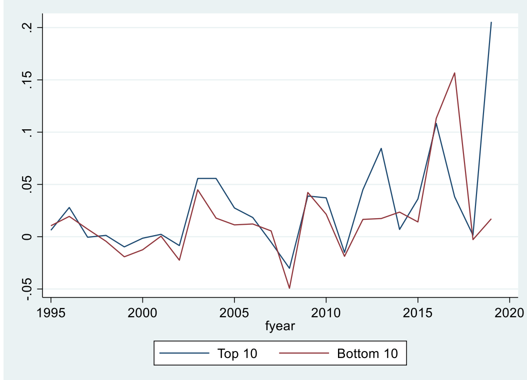
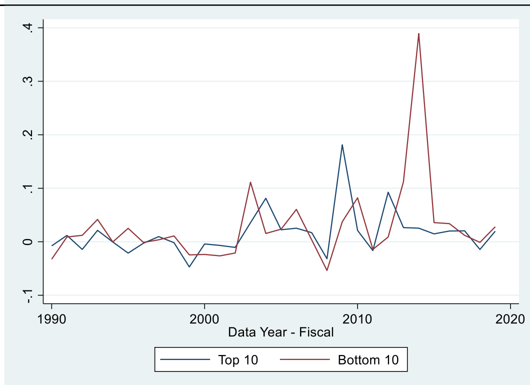
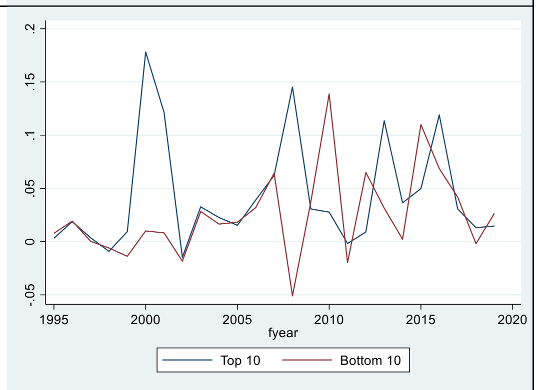
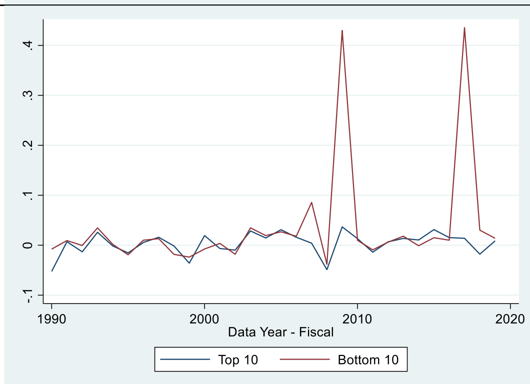
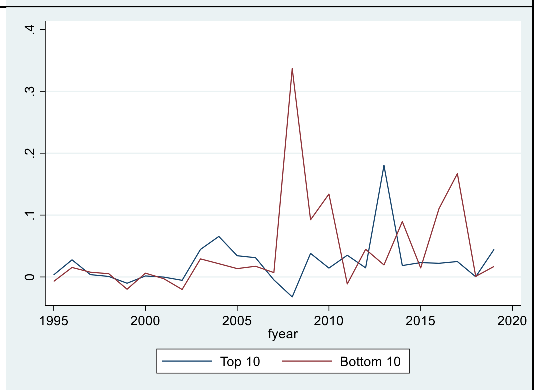
<table>
<tr><th></th><th>Compustat</th><th>Orbis</th></tr>
<tr><td>C4 Country</td><td>
</td><td>
</td></tr>
<tr><td>C4 EU</td><td>
</td><td>
</td></tr>
<tr><td>HHI Country</td><td>
</td><td>
</td></tr>
<tr><td>HHI EU</td><td>
</td><td>
</td></tr>
</table>

Figure 5.1: Average monthly excess stock returns for top and bottom 10^{th} concentration change percentiles. On the y-axis are excess returns computed by taking the cross-sectional average of monthly returns computed by the relative change in stock price $(p_t - p_{t-1})/p_{t-1}$. Time on the x-axis.

5.3 Industry Concentration Trend Results

First, I present the industry concentration trend of the direct competition metrics (average absolute concentration and average concentration change). Second, I outline the findings about the indirect measures of competition.

5.3.1 Direct Measures

Table 5.6 below shows that industry concentration has been falling, and table 5.7 indicates that it is falling at a decreasing rate. All graphs on which this qualitive overview is based can be found in the appendix.

Absolute Concentration Levels	C4 Country		C4 EU		HHI Country		HHI EU	
	Pre 2000	Post 2000	Pre 2000	Post 2000	Pre 2000	Post 2000	Pre 2000	Post 2000
All Industries	↓ (↓)	↑ (↓)	↓ (↓)	→ (↓)	↓ (↓)	↓ (↓)	↓ (↓)	↑ (↓)
Manufacturing	↓ (↓)	↑ (↓)	↓ (↓)	→ (↓)	↓ (↓)	→ (↓)	↓ (↓)	↑ (↓)
Services	↓ (↓)	↓ (↓)	↓ (↓)	→ (↓)	↓ (↓)	↓ (↓)	↓ (↓)	→ (↓)
Digital	↓ (↓)	↓ (↓)	↓ (↓)	→ (↓)	↓ (↓)	↓ (↓)	↓ (↓)	↓ (↓)
Non-Digital	↓ (↓)	↑ (↓)	↓ (↓)	↑ (↓)	↓ (↓)	↑ (↓)	↓ (↓)	↑ (↓)

Table 5.6: Qualitative overview of the trend in absolute concentration levels. ↓ indicates falling, → constant, and ↑ rising concentration. Compustat results are without brackets, Orbis results are in brackets. The first column indicates the analyzed industry subsamples, the first row the applied concentration metrics. Furthermore, analysis has been split into the pre-, and post-2000 trend.

Changes in Concentration Levels	C4 Country				C4 EU				HHI Country				HHI EU			
	Pre 2000		Post 2000		Pre 2000		Post 2000		Pre 2000		Post 2000		Pre 2000		Post 2000	
All Industries	+ (-)	↑ (↑)	+ (-)	↓ (→)	- (-)	↑ (→)	+ (-)	↑ (↑)	- (-)	↑(↑)	+ (-)	↑ (↑)	- (-)	↑ (→)	+ (-)	↓ (↑)
Manufacturing	+ (-)	↑ (→)	+ (-)	↓ (→)	- (-)	↑ (↑)	+ (-)	↑ (→)	- (-)	↑ (↑)	inc. (-)	↑ (↑)	- (-)	↑ (↑)	inc. (-)	↓ (→)
Services	+ (-)	↑ (↑)	+ (-)	↓ (→)	- (-)	↓ (↑)	- (-)	↓ (↑)	- (-)	↓ (↑)	+ (-)	↑ (↑)	- (-)	↑ (↓)	+ (-)	↓ (↑)
Digital	+ (-)	↑ (↑)	+ (-)	↓ (→)	- (-)	↑ (↓)	- (-)	→ (↑)	- (-)	↓ (↓)	inc. (-)	↑ (→)	- (-)	↓ (↓)	inc. (-)	↑ (→)
Non-Digital	+ (-)	↑ (↑)	+ (-)	↓ (→)	- (-)	↓ (→)	+ (-)	↑ (↑)	- (-)	→ (↑)	+ (-)	↑ (↑)	- (-)	↑ (→)	+ (-)	→ (↑)

Table 5.7: Qualitative overview of the trend in average concentration changes. + indicates positive concentration change, - negative concentration change, and inc. stands for inconclusive results.↓ indicates falling, → constant, and ↑ rising concentration changes. Compustat results are without brackets, Orbis results are in brackets. The first column indicates the analyzed industry subsamples, the first row the applied concentration metrics. Furthermore, analysis has been split into the pre-, and post-2000 trend. Within each time-period the left column indicates the dominant sign of concentration the yearly change and the right column the time-series trend.

5.3.2 Indirect Measures

The previous section showed that European Industry Concentration decreased on average but that the decrease is slowing down. Now I inspect whether indirect competition measures support this insight.

5.3.2.1 Fallout Ratio

The results of the time-series trend of the fallout ratio, i.e. the share of firms that could not keep their position in the top four of their industry, contrast the previous findings. Figure 5.2 shows that the average fallout ratio across all industries fell from its peak at 5% around the year 2000 to approximately

1% in 2019, for the *Orbis* as well as the *Compustat* data set. Falling fallout ratios are a sign of incumbents increasingly insulating themselves from competition. While in 2000 still five percent of incumbents lost their position in the top four to a non-incumbent, only 1% of incumbents lost their position in the fop four, in 2019.

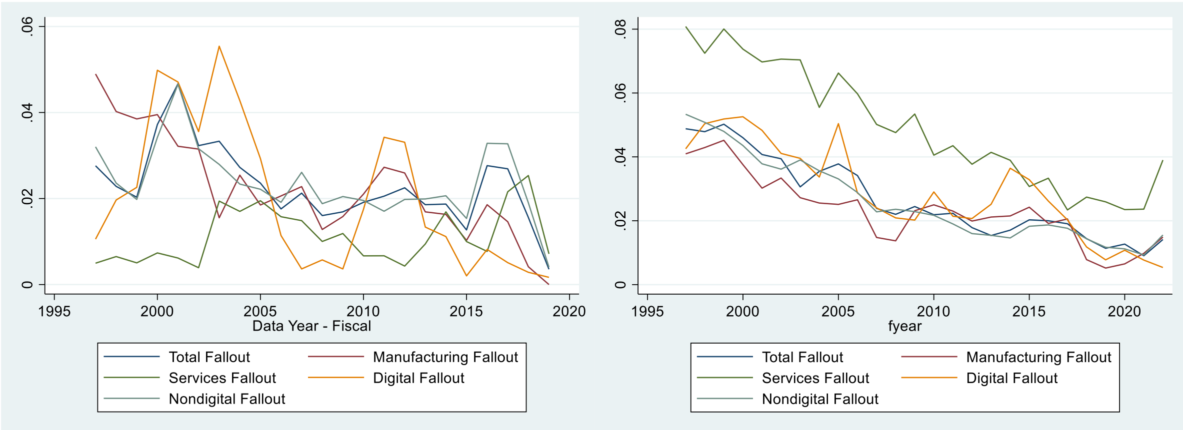


Figure 5.2: Time-series of average fallout ratios across industry classification subsamples (left: Compustat, right: Orbis). Total Fallout *stands for the average fallout ratio across all industries,* Manufacturing Fallout *stands for the average fallout ratio of manufacturing industries only,* Services Fallout *stand for the average fallout ratio of services industries only,* Digital Fallout *stands for the average fallout ratio of digital industries only, and* Nondigital Fallout *stands for the average fallout ratio of nondigital industries only*. Average fallout ratio on the vertical axis, time on the horizontal axis.

5.3.2.2 *Reshuffling*

While fallout ratio results of *Compustat* closely resembled those of *Orbis*, reshuffling results differ a bit for the two data sets. Figure 5.3 below shows the average yearly reshuffling rates, i.e., the complement of the average correlation between the rankings within the top four of each industry. With *Orbis*, you can see a clear decreasing trend, from the peak of 15% in 2001 to the current low level of under 5%. With *Compustat*, the trend is rather flat with a sharp increase in 2019.

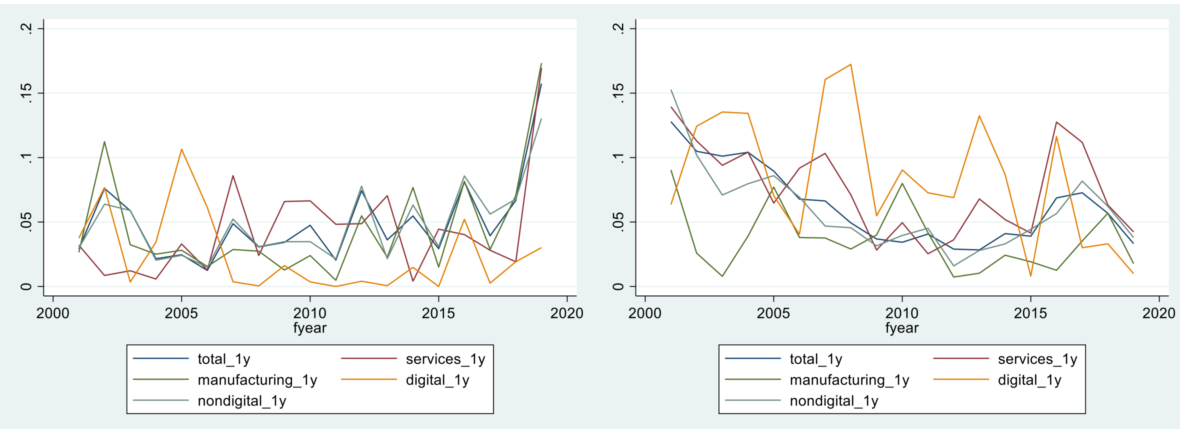


Figure 5.3: Yearly reshuffling rates (left: Compustat, right: Orbis). total_1y *stands for the average annual reshuffling rate across all industries,* services_1y *stands for the average annual reshuffling rate across services industries only,* manufacturing_1y *stands for the average annual reshuffling rate of manufacturing industries only,* digital_1y *stands for the average annual reshuffling rate of digital industries only,* nondigital_1y *stands for the average reshuffling rate of non-digital industries only. Average reshuffling rate on the vertical axis, time on the horizontal axis.*

The time series of average two-year reshuffling rates, depicted in figure 5.4 below, naturally contain fewer data points but with *Orbis* the two-year reshuffling rate tends to be decreasing too (from 13% in 2000 to 7% in 2019), while rates based on *Compustat* fluctuate between 1% and 10%.

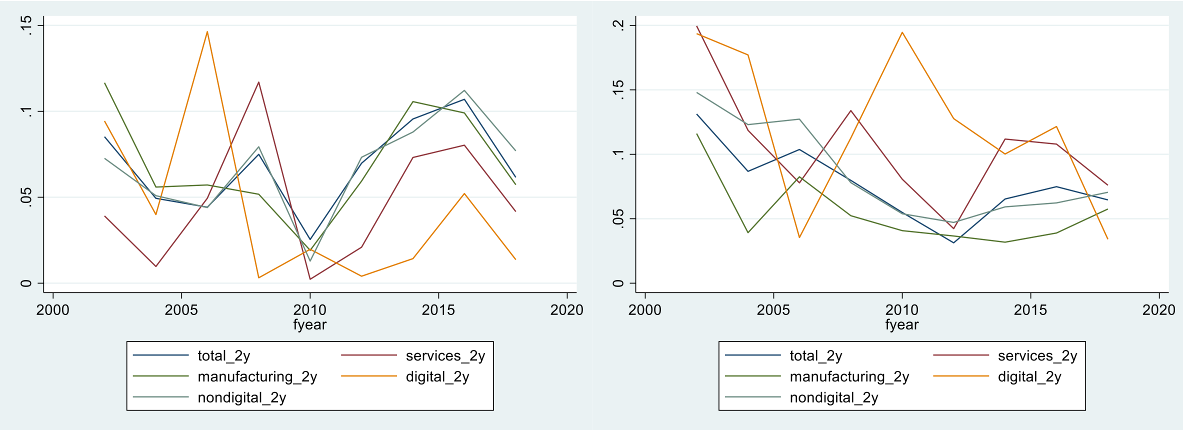


Figure 5.4: 2-year reshuffling rates (left: Compustat, right: Orbis). total_2y *stands for the average 2-year reshuffling rate across all industries,* services_2y *stands for the average 2-year reshuffling rate across services industries only,* manufacturing_1y *stands for the average 2-year reshuffling rate of manufacturing industries only,* digital_1y *stands for the average 2-year reshuffling rate of digital industries only,* nondigital_1y *stands for the average 2-year reshuffling rate of non-digital industries only. Average reshuffling rate on the vertical axis, time on the horizontal axis.*

In summary, average industry concentration is mostly falling at a decreasing rate. In combination with falling profits, this implies that competition has risen during the past two decades but a decelerating rate. Moreover, the mostly decreasing fallout ratios and reshuffling rates point toward less European industry dynamism.

6 Robustness Checks

In this section, I present the results of ROA panel regressions using different subsamples and average industry concentration trends using different weights.

6.1 Panel Regressions

The first and most obvious robustness check is to run my panel regressions for the sub-sample after 2000 because the total number of firms stabilized after this year (see figure 4.4). I use the average of the four concentration metrics as concentration regressors because it allows for parsimonious comparison.

VARIABLES	(1) Model 1	(2) Model 2	(3) Model 3	(4) Model 4	(5) Model 5
market_share	0.018	-0.022	0.007	0.042	0.048
	(0.034)	(0.036)	(0.115)	(0.038)	(0.043)
log_age	0.019*	0.020*	0.019*	0.019*	0.020*
	(0.010)	(0.010)	(0.010)	(0.010)	(0.010)
log_assets	0.009	0.005	0.009	0.008	0.008
	(0.009)	(0.009)	(0.009)	(0.009)	(0.009)
size_inv	-0.373***	-0.366***	-0.372***	-0.375***	-0.379***
	(0.124)	(0.124)	(0.124)	(0.125)	(0.126)
RnD	-0.530***	-0.525***	-0.530***	-0.530***	-0.530***
	(0.076)	(0.075)	(0.076)	(0.076)	(0.076)
CapEx	0.175***	0.186***	0.176***	0.175***	0.174***
	(0.058)	(0.058)	(0.058)	(0.058)	(0.058)
log_concentration	-0.055***	-0.143***	-0.055***		
	(0.014)	(0.044)	(0.015)		
sizexconc		0.016**			
		(0.006)			
msxconc			0.008		
			(0.066)		
concentration				-0.037***	
				(0.010)	
concentration_squ					-0.007***
					(0.002)
Constant	0.004	0.018	0.004	0.048	0.010
	(0.063)	(0.063)	(0.063)	(0.065)	(0.064)
Observations	9,122	9,122	9,122	9,122	9,122
R-squared	0.230	0.233	0.230	0.230	0.228
Number of isin	1,244	1,244	1,244	1,244	1,244

Table 6.1: Panel regression results of ROA on concentration and control variables for the post-2000 Compustat subsample. Models are as described in the methodology section. The applied concentration metric is the average of country-level and EU-level market share of the four largest firms per industry, and the country-level and EU-level HHI. sizexconc *is the interaction term of firm size and industry concentration,* msxconc *is the interaction term of market share and industry concentration.*

Table 6.1 shows that the coefficients of the log of concentration remain negative and strongly significant for the post-2000 *Compustat* sample. In table 6.2, based on post-2000 *Orbis* data only model 2 yields a negative and significant (at the 10%-level) concentration parameter, confirming the main results.

VARIABLES	(1) Model 1	(2) Model 2	(3) Model 3	(4) Model 4	(5) Model 5
market_share	-0.085	-0.124	0.203***	-0.094	-0.103
	(0.120)	(0.140)	(0.062)	(0.128)	(0.134)
log_age	-0.000	-0.020	-0.002	0.000	0.001
	(0.021)	(0.029)	(0.021)	(0.021)	(0.021)
log_assets	0.171**	0.266**	0.174**	0.172**	0.172**
	(0.081)	(0.131)	(0.083)	(0.081)	(0.081)
size_inv	-89.843	-89.716	-89.779	-89.839	-89.836
	(150.332)	(150.214)	(150.293)	(150.330)	(150.328)
log_concentration	0.055	-1.808*	-0.008		
	(0.082)	(1.080)	(0.058)		
sizexconc		0.154			
		(0.094)			
msxconc			0.967**		
			(0.489)		
concentration				0.128	
				(0.158)	
concentration_squ					0.126
					(0.135)
Constant	-1.989**	-3.063**	-2.038**	-2.097**	-2.070**
	(0.868)	(1.425)	(0.891)	(0.987)	(0.950)
Observations	74,970	74,970	74,970	74,970	74,970
R-squared	0.037	0.038	0.037	0.037	0.037
Number of isin	5,996	5,996	5,996	5,996	5,996

Table 6.2: Panel regression results of ROA on concentration and control variables for the post-2000 Orbis subsample. Models are as described in the methodology section. The applied concentration metric is the average of country-level and EU-level market share of the four largest firms per industry, and the country-level and EU-level HHI. sizexconc *is the interaction term of firm size and industry concentration,* msxconc *is the interaction term of market share and industry concentration.*

Figure 6.1 shows that the number of *Compustat*-covered firms in industries 26 (manufacturing of computer, electronic, and optical products) and 58 (publishing activities including software publishing) grew disproportionately from 2000. Hence, I conduct the panel regressions for the post-2000 subsample excluding those two industries. As you can see in figures 9.1 and 9.2 in the appendix, the results stay qualitatively the same, with the exception that in model 2 the parameter estimates become insignificant using *Compustat* as well as *Orbis*.

Using only the post-2000 manufacturing subsamples yields even stronger results. Table 9.3. (Compustat) in the appendix reports negative and significant concentration parameters at the 1%-level. In Orbis, the manufacturing subsample results in only negative concentration coefficients, one of which is significant at the 10%-level (see table 9.4 in the appendix). In contrast, the post-2000 services sector subsample only yields insignificant concentration parameter estimates (tables 9.5 and 9.6 in the appendix). This difference between manufacturing and services indicates that manufacturing has been under stronger competitive pressures because services ROAs are not significantly negatively related to concentration, indicating that higher industry concentration might improve profitability of more services than manufacturing firms.

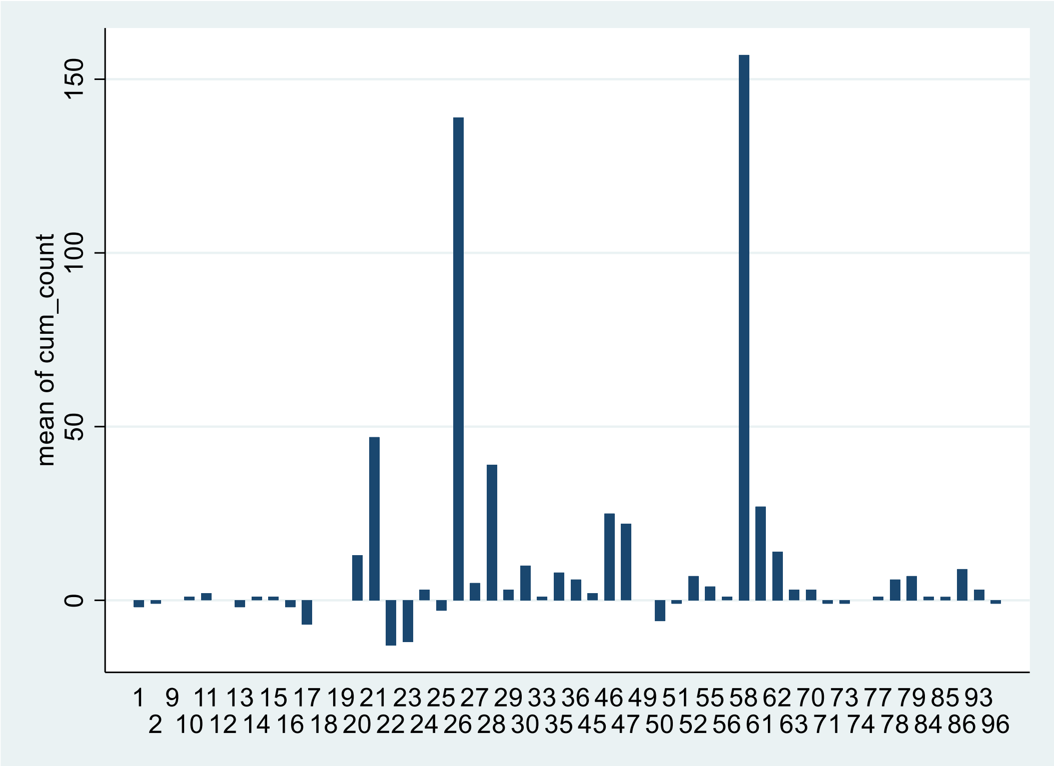


Figure 6.1: Vertical axis: Cumulative change in the number Compustat-covered firms per industry from 2000 to 2019. Horizontal axis: industry code.

Next, I check for geographical differences of the impact of concentration ROA by running the panel regressions for post-2000 Northen, Southern, and Eastern European subsamples. Table 6.3 shows that with *Compustat* data, industry concentration has a negative impact on ROA in all three regions, but while the subsample of Northern European firms has highly significant estimates, the estimates using the Eastern European subsample are insignificant. In table 6.4 you can see that the *Orbis* subsample almost consistently reports negative estimates as well, but only Southern Europe displays some significance. Detailed regression tables can be found in tables 9.7. to 9.12 of the appendix.

Compustat	Model 1	Model 2	Model 3
North	-0.060***	-0.208***	-0.051**
	(0.020)	(0.059)	(0.021)
South	-0.047**	-0.047	-0.054***
	(0.020)	(0.064)	(0.021)
East	-0.062	-0.154	-0.042
	(0.044)	(0.114)	(0.051)

Table 6.3: Compustat post-2000 parameter estimates of the log of industry concentration and standard errors of the log of concentration for the subsample of Northern, Southern, and Eastern European firms.

Orbis	Model 1	Model 2	Model 3
North	-0.047	-0.306	0.014
	(0.062)	(0.759)	(0.064)
South	-0.028*	-0.088	-0.035**
	(0.016)	(0.096)	(0.017)
East	-0.296	-1.963	-0.338
	(0.211)	(1.749)	(0.231)

Table 6.4: Orbis post-2000 parameter estimates of the log of industry concentration and standard errors of the log of concentration for the subsample of Northern, Southern, and Eastern European firms.

Finally, I conduct ROA panel regressions for the subsample of firms pre-2000, i.e. for *Compustat* from 1989 to 1999 and for *Orbis* from 1994 to 1999. Coverage systemically increased during this period, therefore the results need to be interpreted with caution. Tables 6.5 and 6.6 below show that pre-2000 industry concentration seems to be consistently and mostly significantly positively associated with profitability. However, this might be due to sample selection bias.

VARIABLES	(1) Model 1	(2) Model 2	(3) Model 3	(4) Model 4	(5) Model 5
market_share	0.156***	0.167***	0.750***	0.153***	0.157**
	(0.055)	(0.058)	(0.201)	(0.058)	(0.065)
log_age	-0.042***	-0.041***	-0.039***	-0.042***	-0.042***
	(0.012)	(0.013)	(0.012)	(0.012)	(0.013)
log_assets	-0.051*	-0.044	-0.062**	-0.051*	-0.050*
	(0.028)	(0.027)	(0.026)	(0.028)	(0.028)
size_inv	-2.867*	-2.723*	-2.861*	-2.849*	-2.851*
	(1.571)	(1.393)	(1.488)	(1.580)	(1.590)
RnD	-0.114	-0.093	-0.137	-0.112	-0.116
	(0.240)	(0.258)	(0.229)	(0.241)	(0.239)
CapEx	0.030	0.028	0.014	0.030	0.027
	(0.069)	(0.068)	(0.069)	(0.069)	(0.068)
log_concentration	0.018	0.127**	0.057***		
	(0.018)	(0.056)	(0.021)		
sizexconc		-0.018**			
		(0.008)			
msxconc			-0.472***		
			(0.146)		
concentration				0.009	
				(0.011)	
concentration_squ					0.001
					(0.003)
Constant	0.551***	0.510**	0.570***	0.541***	0.551***
	(0.200)	(0.199)	(0.188)	(0.201)	(0.201)
Observations	415	415	415	415	415
R-squared	0.163	0.176	0.200	0.162	0.161
Number of isin	154	154	154	154	154

Table 6.5: Compustat ROA panel regression results using the pre-2000 subsample. Models are as described in the methodology section. The applied concentration metric is the average of country-level and EU-level market share of the four largest firms per industry, and the country-level and EU-level HHI. sizexconc *is the interaction term of firm size and industry concentration,* msxconc *is the interaction term of market share and industry concentration.*

VARIABLES	(1) Model 1	(2) Model 2	(3) Model 3	(4) Model 4	(5) Model 5
market_share	0.094***	0.093***	0.074***	0.093***	0.093***
	(0.029)	(0.028)	(0.028)	(0.029)	(0.029)
log_age	0.006	0.008	0.006	0.006	0.006
	(0.019)	(0.019)	(0.019)	(0.019)	(0.019)
log_assets	-0.015**	-0.023***	-0.015**	-0.015**	-0.015**
	(0.006)	(0.008)	(0.006)	(0.006)	(0.006)
size_inv	-359.454*	-362.320**	-360.068**	-359.393*	-359.248*
	(183.326)	(183.634)	(183.514)	(183.333)	(183.325)
log_concentration	0.025*	0.328	0.040**		
	(0.014)	(0.203)	(0.020)		
sizexconc		-0.024			
		(0.016)			
msxconc			-0.098		
			(0.068)		
concentration				0.040	
				(0.027)	
concentration_squ					0.026
					(0.022)
Constant	0.279**	0.374***	0.287***	0.241**	0.255**
	(0.109)	(0.131)	(0.110)	(0.106)	(0.106)
Observations	4,918	4,918	4,918	4,918	4,918
R-squared	0.071	0.073	0.071	0.071	0.071
Number of isin	1,250	1,250	1,250	1,250	1,250

Table 6.6: Orbis ROA panel regression results using the pre-2000 subsample. Models are as described in the methodology section. The applied concentration metric is the average of country-level and EU-level market share of the four largest firms per industry, and the country-level and EU-level HHI. sizexconc *is the interaction term of firm size and industry concentration,* msxconc *is the interaction term of market share and industry concentration.*

6.2 Trend of Average Concentration and ROA

While the panel regressions include firm- and time-fixed effects that control for unobservable idiosyncratic firm characteristics and time-varying coverage, the time-series trend of average concentration and ROA is more susceptible to increasing coverage and outliers. Therefore, I begin this subsection by comparing the industry sales from *Compustat* and *Orbis* with the OECD STAN gross industry output data. For *Compustat*, industry sales are based on the same countries as the census data. For *Orbis*, total industry sales include a few more small European countries and the UK, inducing a slight upward bias of its coverage ratio, defined as commercial database industry sales over census data industry sales.

6.2.1 Compustat

First, I address the coverage issue and then the impact of different weighting.

6.2.1.1 Coverage

Figures 6.2, 6.3, and 6.4 show the coverage ratio and total industry sales across industries for 1995, 2005, and 2015. The average coverage ratio is about 6% with high variation across industries. For instance, coverage of industry 51 (air transport) is unusually high with roughly 60%. That is, industry sales using *Compustat* are 60% of those reported in OECD STAN. The figures on the right show total *Compustat* and census industry sales. For better visual comparison, census sales have been scaled down by the coverage ratio, so that coverage outliers in terms of industry sales become more apparent. Industries 10, 28, 46, 58, 68, 84, and 86 particularly stand out due to the high amount of their total sales and low *Compustat* coverage. Hence, I removed them in parts of the subsequent robustness checks to see whether they have an impact on the trends in concentration and profitability. However, not all coverage deviations are cause for concern. For instance, sales of industry 68 (real estate) are naturally higher in census data than in *Compustat* because real estate is a local market with a turnover caused by private activity. While the coverage is rather low, it is pretty homogenously distributed across industries, with the exception of a few outliers.

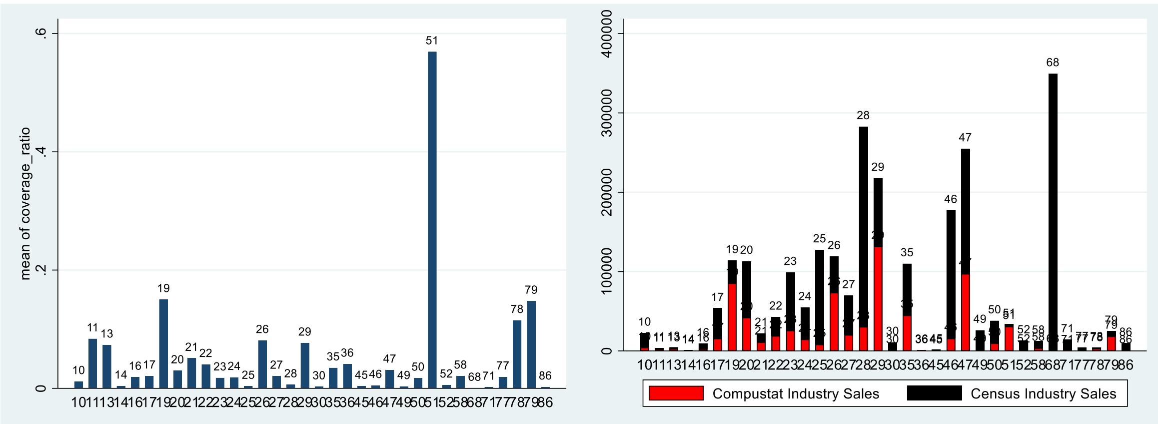


Figure 6.2: Left: Coverage ratio across industries. Right: Industry sales (in millions) – census data multiplied by average coverage ratio (5%). Year: 1995.

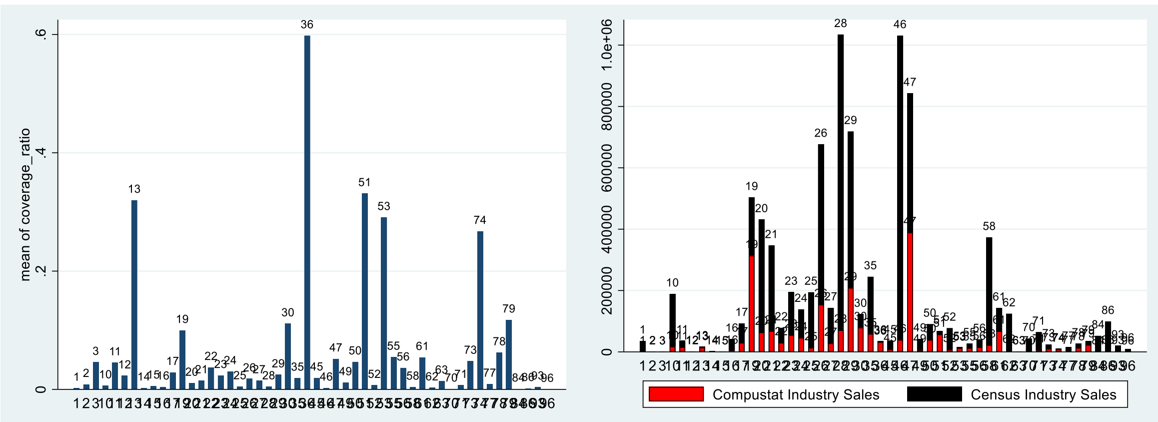


Figure 6.3: Left: Coverage ratio across industries. Right: Industry sales (in millions) – census data multiplied by average coverage ratio (6%). Year: 2005.

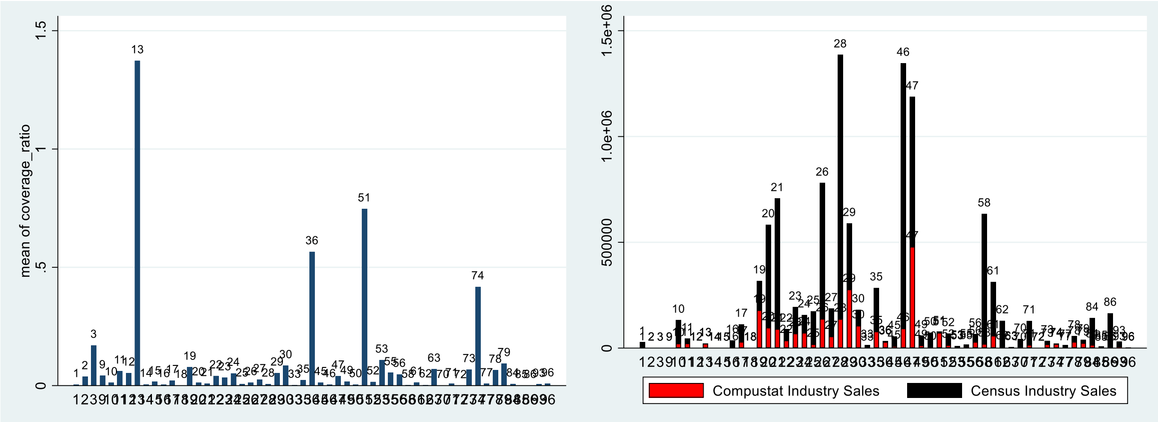


Figure 6.4: Left: Coverage ratio across industries. Right: Industry sales (in millions) – census data multiplied by average coverage ratio (6%). Year: 2015.

6.2.1.2 Weighting

The main results of this paper are based on firm-equally weighted average concentration and ROA. However, the questions arises whether the decreasing trend in concentration and ROA changes when you weigh firms differently. For instance, Grullon et al. (2019) use industry-sales weighted averages in their analysis of times-series trends. Another possible way of aggregation is weighing average industry ROA and concentration equally. Each way of average computation has its merits.

Figure 6.5: Firm-equally weighted average concentration (left) and ROA (right).

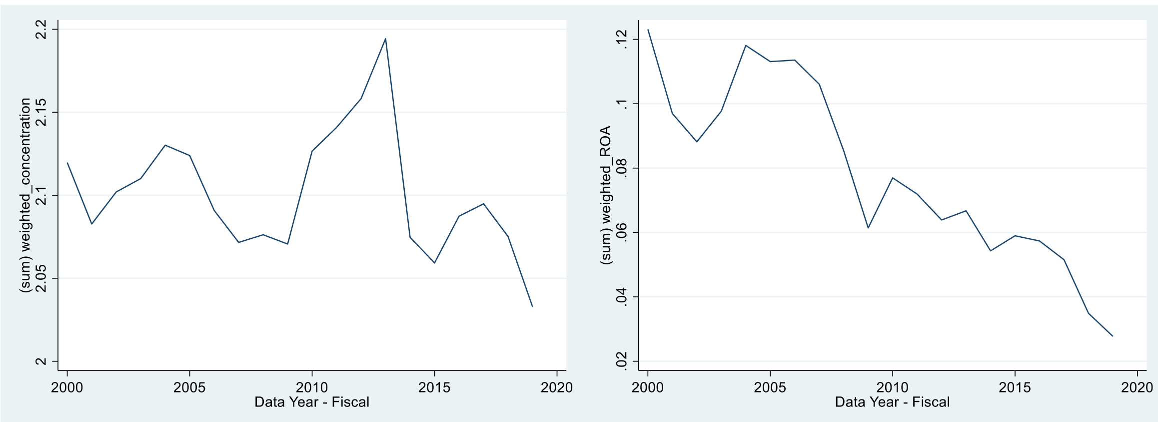


Figure 6.6: Industry sales weighted average concentration (left) and ROA (right).

Figure 6.7: Industry-equally weighted average concentration (left) and ROA (right).

Except the industry-equally weighted concentration, all aggregation methods result in a decreasing time-series trend of concentration and ROA. This seems plausible, considering that some industries are negligibly small but get equal weight in the last method, which might therefore distort results.

Excluding the outlier industries identified in the previous subsection yields the following patterns. Removing industries with low coverage but high industry sales, does not change the decline in profitability across all aggregation methods. However, it leads to an increasing firm-equally and industry-equally weighted average concentration. Industry sales weighted concentration, the dominant method in extant literature (Grullon et al., 2019; Gutierrez & Philippon, 2018), remains falling.

Figure 6.8: Firm-equally weighted average concentration (left) and ROA (right) excluding industries 10, 28, 46, 58, 68, 84, and 86.

Figure 6.9: Industry sales weighted average concentration (left) and ROA (right) excluding industries 10, 28, 46, 58, 68, 84, and 86.

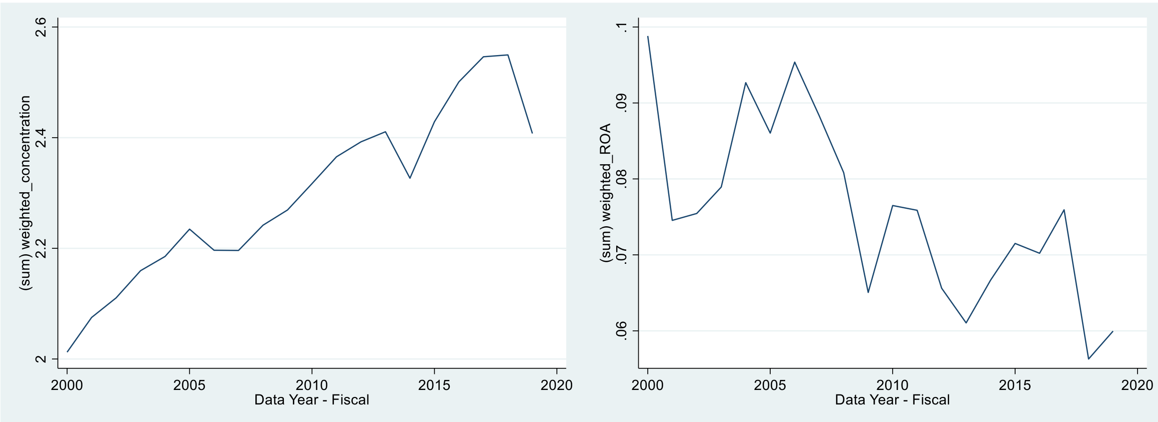


Figure 6.10: Industry-equally weighted average concentration (left) and ROA (right) excluding industries 10, 28, 46, 58, 68, 84, and 86.

Finally, I analyze in figure 6.11 (firm-equally weighted) average concentration of manufacturing only using EU-level concentration, and services only using country-level concentration, thereby aligning the industry-subsamples with more appropriate geographical market definitions. Figure 6.11 shows a hump-shaped concentration trend for manufacturing and decreasing concentration trend for services.

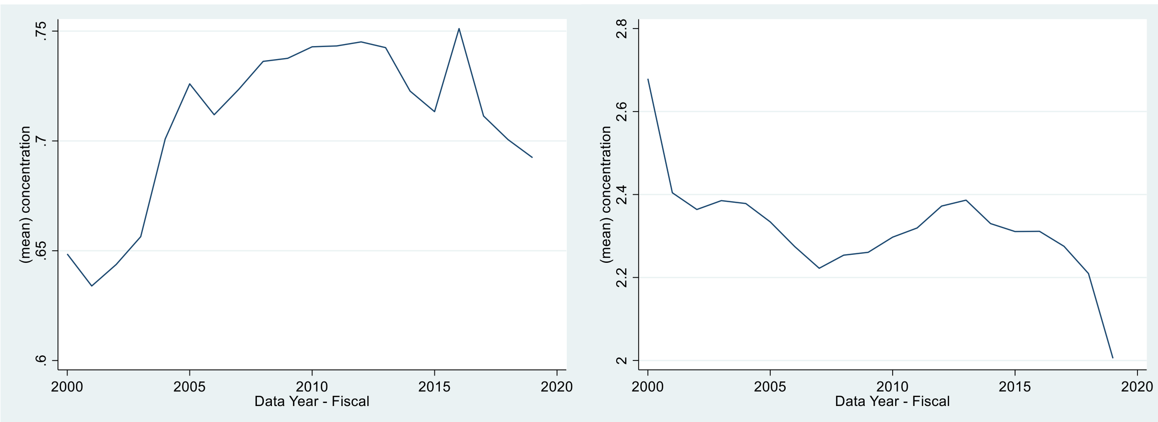


Figure 6.11: Firm-equally weighted average concentration. Left: Manufacturing only (10-33) using the average of C4-EU and HHI-EU concentration metrics. Right: Services only (45-99) using the average of C4-Country and HHI-Country concentration metrics.

6.2.2 Orbis

First, I address the coverage issue and then impact of different weighting.

6.2.2.1 Coverage

Figures 6.12, 6.13, and 6.14 show that the coverage ratios and total industry sales across industries for 1995, 2005, and 2015 vary with *Orbis* data as well. The average coverage ratio is about 13%, twice as high as with *Compustat*. Here, industries 25, 26, 28, 46, 47, 58, 68, 84, 86, and 93 particularly stand out due to the high amount of their total sales and low *Orbis* coverage.

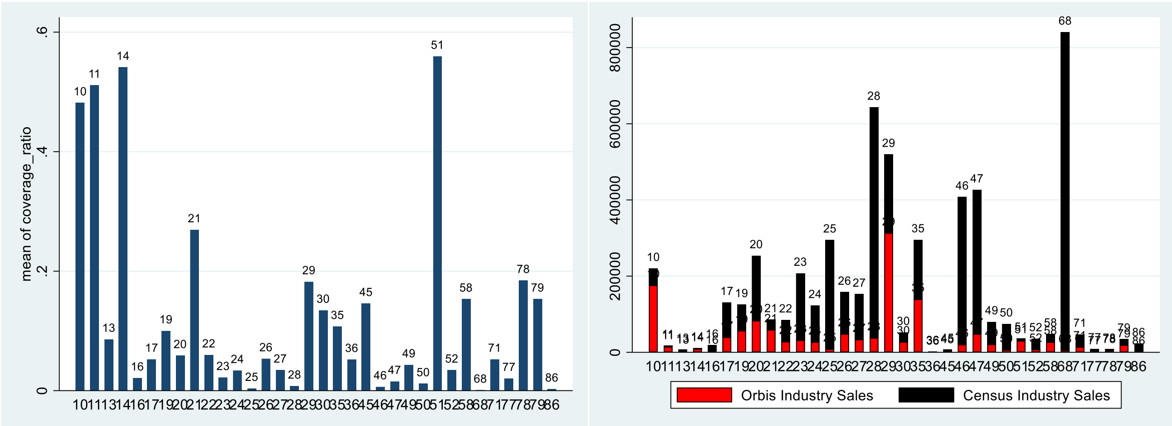


Figure 6.12: Left: Coverage ratio across industries. Right: Industry sales (in millions) – census data multiplied by average coverage ratio (12%). Year: 1995.

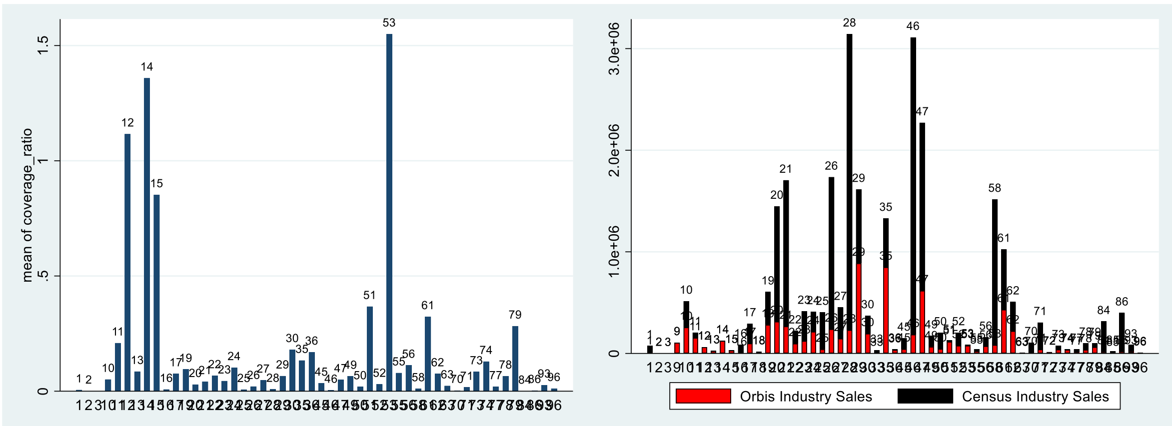


Figure 6.13: Left: Coverage ratio across industries. Right: Industry sales (in millions) – census data multiplied by average coverage ratio (14%). Year: 2005.

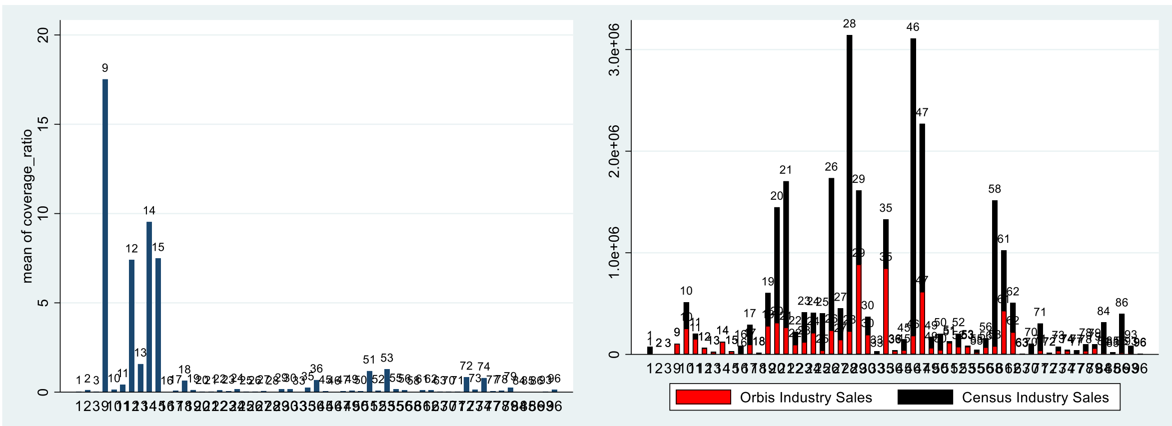


Figure 6.14: Left: Coverage ratio across industries. Right: Industry sales (in millions) – census data multiplied by average coverage ratio (14%). Year: 2015.

6.2.2.2 *Weighting*

Firm-equally and industry sales weighted average concentration and ROA are decreasing in *Orbis*, while industry-equally weighted average concentration is hump-shaped and average ROA rather constant, as you can see in the figures below.

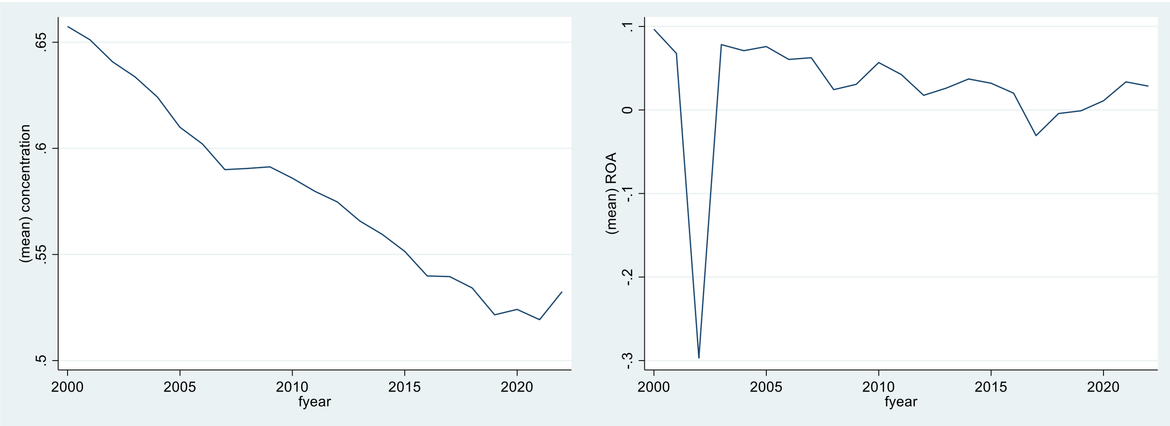


Figure 6.15: Firm-equally weighted average concentration (left) and ROA (right).

Figure 6.16: Industry sales weighted average concentration (left) and ROA (right).

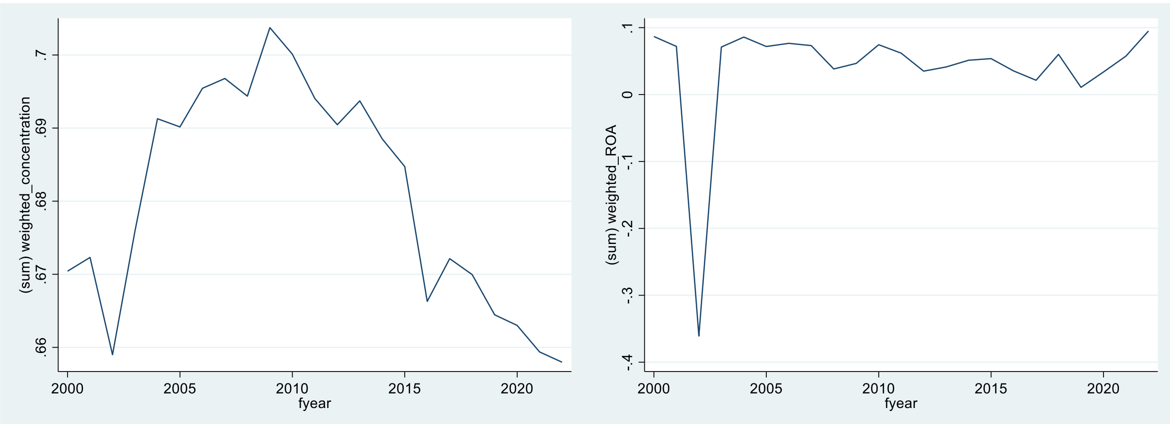


Figure 6.17: Industry-equally weighted average concentration (left) and ROA (right).

Excluding the outlier industries identified in the previous subsection yields the same patterns.

Splitting the *Orbis* dataset into manufacturing only and services only subsamples and plotting the time-series of only EU-level (for manufacturing) and country-level (for services) concentration trends confirms the overall decreasing trend, as can be seen in figure 6.18 below.

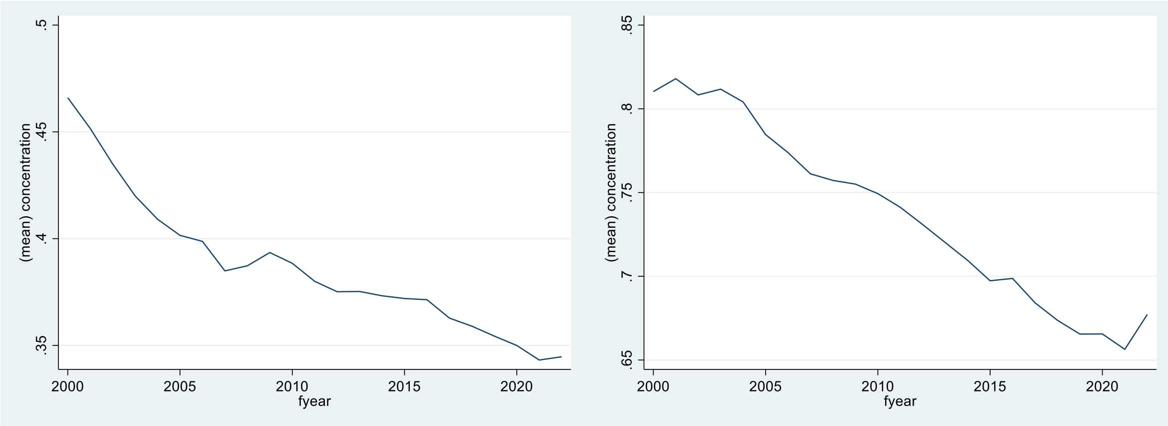


Figure 6.18: Firm-equally weighted average concentration. Left: Manufacturing only (10-33) using the average of C4-EU and HHI-EU concentration metrics. Right: Services only (45-99) using the average of C4-Country and HHI-Country concentration metrics.

7 Limitations, Discussion, and Potential for Future Research

In this chapter, I briefly outline limitations, discuss the findings, and provide suggestions for future research.

7.1 Limitations

The biggest limitation of this research is certainly data availability. Neither for competition nor digitization are the current metrics very accurate. Industry concentration is only an imperfect measure of actual competitive realities. However, in combination with the analysis of industry dynamism, profitability, and stock returns, competition is significantly less elusive than before. In contrast, firm-level digitization remains an underexplored field of research. Although there are efforts by academia and organizations such as the OECD to measure digitization, a public firm-level digitization score is still unavailable. Hence, current analysis is restricted to the industry-level.

Regarding the potential coverage issue of *Compustat* and *Orbis*, reported insights are mostly the same when using two coverage-robust concentration metrics (Country-and EU-Level *Compustat* C4). Both datasets tend to become very similar after 2000, strengthening their validity.

7.2 Discussion

In the following section, I discuss the three parts of this paper in turn.

7.2.1 Industry Concentration and ROA

The mostly negative impact of industry concentration on ROA means that an increase in industry concentration is associated with a decrease in profitability. There are several possible interpretations of this finding. It might be possible that firms only gradually make use of their newly gained market power. Therefore, future research might regress ROA on lags of industry concentration to see whether profitability tends to rise over time after consolidations.

Another possible reason for the negative impact of industry concentration on ROA is foreign competition. Trade liberalizations might have created foreign rivals that were not included in my European datasets. Their competitive advantage might erode European profitability and force consolidations. However, since trade flows of the US and Europe are broadly similar (Philippon, 2019), the trade explanation of decreasing concentration and profits is rather unlikely. Relatively more efficient firms, which tend to be more profitable, often grow at the expense of their rivals (Demsetz, 1973).

Therefore, profitability for these more efficient firms might increase with concentration, something that is not observable when conducting simple linear panel regression. However, my series of cross-sectional quantile regressions shows that also the performance of highly profitable firms is negatively affected by higher concentration. Still, the negative impact of concentration is less in magnitude for highly profitable firms, indicating that mostly firms with poor profitability, i.e., rather inefficient firms, suffer from their industries becoming more concentrated. Future research might conduct panel quantile regressions to additionally account for firm-, and time-fixed effects.

Schmalensee (1989) states that mixed results in the concentration-profitability relationship “may reflect, at least in part, intertemporal changes”. His work points toward “pro-cyclical industry-level changes in the strength of the concentration -profitability relation”. Therefore, conducting panel regressions with further subsamples of time and industries might reveal clearer patterns. However, due to the large number of individual industries and business cycles, conducting this analysis was beyond the scope of this paper.

It seems reasonable to me that higher industry concentration makes incumbents, with an inherently lower pressure to innovate, more dominant. Through this lower innovation rate average European profitability declines (Arrow, 1972), as long as market power is curbed. Döttling et al. (2017, p. 41) found the European economy of the last twenty years characterized by “financial constraints, high risk premia, low expected demand and low expected cash flows”. Beyond the other discussed potential causes, which need to be more closely examined, the negative impact of industry concentration on performance is therfore most likely a manifestation of the investment gap, sluggish demand, and declining industry dynamism of the post-GFC European economy.

7.2.2 Digitization and ROA

The positive impact of the level of industry digitization on the cross-section of ROAs is not surprising, when keeping its many benefits, and positive impact on stock returns (Gaspar et al., 2022; Hua, 2022) in mind. Consumers increasingly demand digital products, services and appreciate digital ways of conducting business. In line with this explanation, the parameter estimates of the digitization score increased over time, reflecting that digitization became more important. What is striking is the significant difference between the average digitization score parameters of the 10th and 90th ROA percentile. Digitization increases profitability more for highly profitable firms than poor-performing ones. The benefit of digitization seems therefore to be higher for high-profitability than low-profitability firms. This might be a manifestation of the notion that poor-performing firms need to improve their core business model before being able to reap the benefits of digitization. Furthermore, it is evidence of the winner-takes-all effect of digitization, induced by network effects (Stallkamp & Schotter, 2021).

7.2.3 Industry Concentration and Stock Returns

The fact that average excess stock returns of firms in industries with the highest concentration change are not significantly higher than those with the lowest is reasonable since higher concentration is associated with lower profitability. An opposite result would have been troublesome.

7.2.4 Time-Series Trend of European Industry Concentration: Outlook and Wider Implications

In terms of its wider implications, the general trend of falling industry concentration is most likely due to the liberalization of European markets and their more independent regulatory institutions. For instance, American antitrust cases are decided in court, whereas in Europe, the directorate-general for competition (DG Comp) first makes its decision which afterwards can be appealed in courts, making the DG Comp more powerful than its American counterpart (Philippon, 2019). Furthermore, European merger control has improved steadily (Duso et al., 2011). A good example for the structural difference between the US and European economy is the airline industry. Aviation in Europe is much more liberal than in the US. For instance, European airlines are prohibited to cater inter-American flights, while many new low-cost carriers captured the European market and made flying in the EU significantly cheaper than a few decades ago. Moreover, "US firms spend substantially more on lobbying and campaign contributions and are far more likely to achieve their lobbying goals than European firms and lobbyists" (Philippon, 2019, p. 148). Quantifying the impact of higher American lobbying efforts has proven to be cumbersome because it is endogenous, but there are plenty of cases that buttress Philippon's contention. Finally, the EU welcomed many new member states through the Eastern expansion over the course of the last two decades. Naturally, this drives inter-European competition and decreases profitability and concentration.

The fact that industry concentration is decreasing at a slower rate might indicate that the trend is in the process of reversing itself. Competition authorities must remain vigilant to ensure continued enforcement of the European single market on all levels.

7.3 Potential for Future Research

As outlined in the limitations section, aggregate analysis of industry concentration alone is imperfect. Overall, the heterogeneity of results across geographical market definitions, as exhibited in the mean stock returns, for instance, highlights the importance to go back to the fundamentals in the analysis of industry concentration. Market definition should be guided by substitutability of goods and services, which is determined by customer's needs, functionality, and transaction costs. Currently, standard industry classifications are not oriented along these criteria (Affeldt et al., 2021). Implementing a comprehensive new classification for all 800 European cities is quite a challenging task but might make macro-level competition analysis substantially more precise. Artificial intelligence, and machine learning, might provide for the first time the tools to do this.

8 Conclusion

In conclusion, my research shows that the exact quantitative impact of industry concentration on ROA and stock returns depends on the used dataset and concentration metric. But the overall qualitative insight of falling industry concentration and its negative impact on ROA holds for almost all specifications. Therefore, this paper strengthens the view that European markets became more competitive than the American one since the turn of the millennium.

The log of industry concentration is negatively related to profitability in Europe for nine out of ten analyzed concentration metrics. This negative impact of industry concentration on performance most likely reflects the investment gap, sluggish demand, and declining industry dynamism in the post-GFC European economy. Using a quantile regression approach further shows that the mostly negative impact of the level of industry concentration on ROA is less in magnitude for highly profitable firms than for those with poor profitability, indicating that mostly firms with poor profitability, i.e., rather inefficient firms, suffer from their industries becoming more concentrated.

Post-2000, a high degree of industry-level digitization consistently positively affects firm profitability, and stronger so for top-ROA-percentile firms. The benefit of digitization seems therefore to be higher for high-profitability than low-profitability firms, providing evidence for the winner-takes-all principle. The parameter estimates of the digitization score increased over time, reflecting that digitization became more important.

Average yearly stock returns of firms with the highest concentration change are only significantly higher than those of firms with the lowest concentration change for two of four concentration metrics for a sample of 5251 firms obtained from *Orbis*. *Compustat* data yields opposite results, rendering the impact of concentration changes on European stock returns insignificant – a finding in line with its negative impact on profitability.

The time-series of industry concentration has been falling in Europe for the last two decades but at a decreasing rate. Due to its similar trade relations and technological prowess with the US, globalization and technology are not likely driving decreasing profits and concentration in Europe, but rather deregulation and stricter antitrust enforcement.

9 References

Affeldt, P., Duso, T., Gugler, K. P., & Piechucka, J. (2021). *Market concentration in Europe: Evidence from antitrust markets*.

Arrow, K. J. (1972). *Economic welfare and the allocation of resources for invention*. Springer.

Autor, D., Dorn, D., Katz, L. F., Patterson, C., & Van Reenen, J. (2020). The fall of the labor share and the rise of superstar firms. *The Quarterly Journal of Economics*, *135*(2), 645–709.

Bain, J. S. (1951). Relation of profit rate to industry concentration: American manufacturing, 1936–1940. *The Quarterly Journal of Economics*, *65*(3), 293–324.

Bajgar, M., Berlingieri, G., Calligaris, S., Criscuolo, C., & Timmis, J. (2023). Industry concentration in Europe and North America. *Industrial and Corporate Change*.

Bansraj, D., & Smit, H. T. J. (2017). Optimal Conditions for Buy-and-Build Acquisitions. *Erasmus School of Economics, Preliminary Version*, *1*, 1–45.

Bansraj, D., Smit, H., & Volosovych, Vadym. (2020). *Can Private Equiy Funds Act as Strategic Buyers?*.

Broccardo, L., Truant, E., & Dana, L.-P. (2023). The interlink between digitalization, sustainability, and performance: An Italian context. *Journal of Business Research*, *158*, 113621.

Brooks, C. (2019). *Introductory econometrics for finance* (Fourth edition., pp. xxxi, 696 Seiten). Cambridge University Press.

Bustamante, M. C., & Donangelo, A. (2017). Product market competition and industry returns. *The Review of Financial Studies*, *30*(12), 4216–4266.

Calvino, F., Criscuolo, C., Marcolin, L., & Squicciarini, M. (2018). *A taxonomy of digital intensive sectors*.

Capon, N., Farley, J. U., & Hoenig, S. (1990). Determinants of financial performance: A meta-analysis. *Management Science*, *36*(10), 1143–1159.

Chan, L. K., Karceski, J., & Lakonishok, J. (2003). The level and persistence of growth rates. *The Journal of Finance*, *58*(2), 643–684.

Demsetz, H. (1973). Industry structure, market rivalry, and public policy. *The Journal of Law and Economics*, *16*(1), 1–9.

Domowitz, I., Hubbard, R. G., & Petersen, B. C. (1986). The intertemporal stability of the concentration-margins relationship. *The Journal of Industrial Economics*, 13–34.

Döttling, R., Gutierrez Gallardo, G., & Philippon, T. (2017). *Is there an investment gap in advanced economies? If so, why?*

Duso, T., Gugler, K., & Yurtoglu, B. B. (2011). How effective is European merger control? *European Economic Review*, *55*(7), 980–1006.

Furman, J., & Orszag, P. (2015). *A firm-level perspective on the role of rents in the rise in inequality. Presentation at 'A Just Society' Centennial Event in Honor of Joseph Stiglitz Columbia University*.

Gallagher, D. R., Ignatieva, K., & McCulloch, J. (2015). Industry concentration, excess returns and innovation in A ustralia. *Accounting & Finance*, *55*(2), 443–466.

Gaspar, J.-M., Wang, S., & Xu, L. (2022). Size and Resilience of the Digital Economy. *SSRN Electronic Journal*.

Goddard, J., Tavakoli, M., & Wilson, J. O. (2005). Determinants of profitability in European manufacturing and services: Evidence from a dynamic panel model. *Applied Financial Economics*, *15*(18), 1269–1282.

Grullon, G., Larkin, Y., & Michaely, R. (2019). Are US Industries Becoming More Concentrated? *Review of Finance*, *23*(4), 697–743.

Gutiérrez, G., & Philippon, T. (2017). *Declining Competition and Investment in the US*. National Bureau of Economic Research.

Gutierrez, G., & Philippon, T. (2018). *How European markets became free: A study of institutional drift*. National Bureau of Economic Research.

Hammer, B., Marcotty-Dehm, N., Schweizer, D., & Schwetzler, B. (2022). Pricing and value creation in private equity-backed buy-and-build strategies. *Journal of Corporate Finance*, *77*.

Hou, K., & Robinson, D. T. (2006). Industry concentration and average stock returns. *The Journal of Finance*, *61*(4), 1927–1956.

Hua, M. (2022). Decrypting the Digital Economy: The Digital Alpha and Its Origins. *Available at SSRN 3886864*.

Jensen, M. C. (1986). Agency costs of free cash flow, corporate finance, and takeovers. *The American Economic Review*, *76*(2), 323–329.

Kalemli-Ozcan, S., Sorensen, B., Villegas-Sanchez, C., Volosovych, V., & Yesiltas, S. (2015). *How to construct nationally representative firm level data from the Orbis global database: New facts and aggregate implications*. National Bureau of Economic Research.

Koenker, R., & Basset, G. (1978). Asymptotic theory of least absolute error regression. *Journal of the American Statistical Association*, *73*(363), 618–622.

Kohtamäki, M., Parida, V., Patel, P. C., & Gebauer, H. (2020). The relationship between digitalization and servitization: The role of servitization in capturing the financial potential of digitalization. *Technological Forecasting and Social Change*, *151*, 119804.

Little, I. M. (1962). Higgledy piggledy growth. *Bulletin of the Oxford University Institute of Economics & Statistics*, *24*(4), 387–412.

Nicholson, J. R. (2020). New digital economy estimates. *Bureau of Economic Analysis*, *2*.

Pattitoni, P., Petracci, B., & Spisni, M. (2014). Determinants of profitability in the EU-15 area. *Applied Financial Economics*, *24*(11), 763–775.

Philippon, T. (2019). *The Great Reversal*. Harvard University Press.

Pike, C., Capobianco, A., & Gomes, A. (2018). *Market Concentration*.

Rayner, A. C., & Little, I. M. D. (1966). Higgledy piggledy growth again: An investigation of the predictability of company earnings and dividends in the UK 1951-1961.

Schmalensee, R. (1989). Intra-industry profitability differences in US manufacturing 1953-1983. *The Journal of Industrial Economics*, 337–357.

Stallkamp, M., & Schotter, A. P. J. (2021). Platforms without borders? The international strategies of digital platform firms. *Global Strategy Journal*, *11*(1), 58–80.

Theiri, S., & Hadoussa, S. (2023). Digitization effects on banks' financial performance: The case of an African country. *Competitiveness Review: An International Business Journal*.

10 Appendix

10.1 Compustat Panel Regression Results

C4 Country	(1) Model 1	(2) Model 2	(3) Model 3	(4) Model 4	(5) Model 5
market_share_c4_country	-0.015*	-0.011	-0.022	-0.020**	-0.016*
	(0.008)	(0.009)	(0.029)	(0.009)	(0.009)
log_age	0.016*	0.015	0.016*	0.016	0.016
	(0.010)	(0.010)	(0.010)	(0.010)	(0.010)
log_assets	0.005	0.001	0.005	0.006	0.007
	(0.008)	(0.010)	(0.009)	(0.008)	(0.008)
size_inv	-0.384***	-0.375***	-0.383***	-0.382***	-0.381***
	(0.124)	(0.124)	(0.124)	(0.124)	(0.124)
RnD	-0.532***	-0.532***	-0.532***	-0.532***	-0.532***
	(0.075)	(0.075)	(0.075)	(0.075)	(0.075)
CapEx	0.184***	0.182***	0.184***	0.185***	0.187***
	(0.057)	(0.057)	(0.057)	(0.057)	(0.057)
log_c4_country	0.020***	0.033*	0.020***		
	(0.007)	(0.017)	(0.007)		
c4_country_sizexconc		-0.002			
		(0.002)			
c4_country_msxconc			0.008		
			(0.030)		
c4_country				0.032**	
				(0.015)	
c4_country_squ					0.008
					(0.005)
Constant	0.059	0.081	0.058	0.002	0.008
	(0.065)	(0.075)	(0.066)	(0.061)	(0.061)
Observations	9,536	9,536	9,536	9,536	9,536
R-squared	0.223	0.224	0.223	0.222	0.221
Number of isin	1,265	1,265	1,265	1,265	1,265

Robust standard errors in parentheses
*** p<0.01, ** p<0.05, * p<0.1

Table 10.1: Regression table of ROA on country-level C4 industry concentration and control variables.

C4 EU	(1) Model 1	(2) Model 2	(3) Model 3	(4) Model 4	(5) Model 5
market_share_c4_eu	0.043***	0.004	0.087***	0.096*	0.004
	(0.007)	(0.011)	(0.026)	(0.051)	(0.019)
log_age	0.012	0.010	0.012	0.015	0.015
	(0.010)	(0.010)	(0.010)	(0.010)	(0.010)
log_assets	0.007	0.043***	0.007	0.007	0.007
	(0.008)	(0.013)	(0.008)	(0.008)	(0.008)
size_inv	-0.375***	-0.373***	-0.375***	-0.384***	-0.384***
	(0.123)	(0.123)	(0.123)	(0.124)	(0.124)
RnD	-0.532***	-0.531***	-0.532***	-0.533***	-0.532***
	(0.075)	(0.075)	(0.075)	(0.075)	(0.075)
CapEx	0.186***	0.182***	0.186***	0.187***	0.188***
	(0.057)	(0.057)	(0.057)	(0.057)	(0.057)
log_c4_eu	-0.027***	-0.073***	-0.027***		
	(0.006)	(0.018)	(0.007)		
c4_eu_sizexconc		0.007***			
		(0.002)			
c4_eu_msxconc			-0.043**		
			(0.019)		
c4_eu				-0.089*	
				(0.052)	
c4_eu_squ					0.001
					(0.006)
Constant	-0.108*	-0.341***	-0.110*	0.015	0.013
	(0.064)	(0.092)	(0.064)	(0.061)	(0.061)
Observations	9,536	9,536	9,536	9,536	9,536
R-squared	0.226	0.230	0.226	0.221	0.220
Number of isin	1,265	1,265	1,265	1,265	1,265

Robust standard errors in parentheses
*** p<0.01, ** p<0.05, * p<0.1

Table 10.2: Regression table of ROA on EU-level C4 industry concentration and control variables.

HHI Country	(1) Model 1	(2) Model 2	(3) Model 3	(4) Model 4	(5) Model 5
market_share_HHI_country	0.069**	0.050	0.137	0.106***	0.112***
	(0.029)	(0.031)	(0.311)	(0.034)	(0.036)
log_age	0.014	0.015	0.014	0.013	0.014
	(0.010)	(0.009)	(0.010)	(0.010)	(0.010)
log_assets	0.004	-0.082*	0.004	0.002	0.002
	(0.009)	(0.043)	(0.009)	(0.009)	(0.009)
size_inv	-0.383***	-0.374***	-0.383***	-0.385***	-0.388***
	(0.124)	(0.123)	(0.124)	(0.124)	(0.125)
RnD	-0.535***	-0.532***	-0.535***	-0.537***	-0.537***
	(0.076)	(0.075)	(0.076)	(0.075)	(0.075)
CapEx	0.188***	0.195***	0.188***	0.189***	0.190***
	(0.057)	(0.057)	(0.057)	(0.056)	(0.057)
log_HHI_country	-0.044***	-0.101***	-0.043***		
	(0.011)	(0.035)	(0.012)		
HHI_country_sizexconc		0.010**			
		(0.005)			
HHI_country_msxconc			-0.007		
			(0.033)		
HHI_country				-0.000***	
				(0.000)	
HHI_country_squ					-0.000***
					(0.000)
Constant	0.382***	0.849***	0.374***	0.087	0.053
	(0.118)	(0.295)	(0.126)	(0.066)	(0.063)
Observations	9,536	9,536	9,536	9,536	9,536
R-squared	0.226	0.228	0.226	0.229	0.229
Number of isin	1,265	1,265	1,265	1,265	1,265

Robust standard errors in parentheses
*** p<0.01, ** p<0.05, * p<0.1

Table 10.3: Regression table of ROA on country-level HHI industry concentration and control variables.

HHI EU	(1) Model 1	(2) Model 2	(3) Model 3	(4) Model 4	(5) Model 5
market_share_HHI_eu	0.046	-0.022	-0.215	0.073	0.036
	(0.063)	(0.064)	(0.525)	(0.073)	(0.074)
log_age	0.014	0.013	0.014	0.014	0.015
	(0.010)	(0.010)	(0.010)	(0.010)	(0.010)
log_assets	0.006	-0.061*	0.007	0.006	0.006
	(0.009)	(0.032)	(0.009)	(0.009)	(0.009)
size_inv	-0.385***	-0.390***	-0.383***	-0.387***	-0.385***
	(0.125)	(0.126)	(0.124)	(0.125)	(0.125)
RnD	-0.532***	-0.526***	-0.532***	-0.532***	-0.532***
	(0.076)	(0.075)	(0.076)	(0.076)	(0.076)
CapEx	0.186***	0.189***	0.186***	0.188***	0.188***
	(0.057)	(0.057)	(0.057)	(0.057)	(0.057)
log_HHI_eu	-0.033***	-0.092***	-0.034***		
	(0.013)	(0.034)	(0.013)		
HHI_eu_sizexconc		0.010**			
		(0.004)			
HHI_eu_msxconc			0.031		
			(0.057)		
HHI_eu				-0.000***	
				(0.000)	
HHI_eu_squ					-0.000*
					(0.000)
Constant	0.248**	0.648**	0.253**	0.042	0.019
	(0.116)	(0.253)	(0.117)	(0.065)	(0.063)
Observations	9,536	9,536	9,536	9,536	9,536
R-squared	0.223	0.224	0.223	0.222	0.221
Number of isin	1,265	1,265	1,265	1,265	1,265

Robust standard errors in parentheses
*** p<0.01, ** p<0.05, * p<0.1

Table 10.4: Regression table of ROA on EU-level HHI industry concentration and control variables.

10.2 Orbis Panel Regression Results

C4 Country	(1) Model 1	(2) Model 2	(3) Model 3	(4) Model 4	(5) Model 5
market_share_country	0.063***	0.058***	0.063***	0.063***	0.064***
	(0.018)	(0.018)	(0.018)	(0.017)	(0.017)
log_age	0.030***	0.024***	0.029***	0.030***	0.029***
	(0.008)	(0.007)	(0.008)	(0.008)	(0.008)
log_assets	0.034	0.042*	0.035	0.034	0.034
	(0.022)	(0.022)	(0.022)	(0.022)	(0.022)
size_inv	-835.578***	-831.214***	-834.977***	-835.559***	-835.526***
	(130.237)	(129.580)	(130.567)	(130.241)	(130.239)
RnD	-0.000**	-0.000**	-0.000**	-0.000**	-0.000**
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
log_c4_country	-0.022	-1.024**	-0.034		
	(0.050)	(0.425)	(0.059)		
c4_country_sizexconc		0.077**			
		(0.031)			
c4_country_msxconc			0.199		
			(0.281)		
c4_country				-0.043	
				(0.068)	
c4_country_squ					-0.033
					(0.043)
Constant	-0.440	-0.519*	-0.443*	-0.397	-0.406
	(0.264)	(0.266)	(0.266)	(0.294)	(0.283)
Observations	19,775	19,775	19,775	19,775	19,775
R-squared	0.260	0.263	0.260	0.260	0.260
Number of isin	1,709	1,709	1,709	1,709	1,709

Robust standard errors in parentheses
*** p<0.01, ** p<0.05, * p<0.1

Table 10.5: Regression table of ROA on country-level C4 and control variables.

C4 EU	(1) Model 1	(2) Model 2	(3) Model 3	(4) Model 4	(5) Model 5
market_share_eu	0.072	0.037	0.104	0.070	0.070
	(0.074)	(0.080)	(0.071)	(0.073)	(0.073)
log_age	0.030***	0.021***	0.029***	0.030***	0.030***
	(0.008)	(0.007)	(0.008)	(0.008)	(0.008)
log_assets	0.036	0.053**	0.037	0.036	0.036
	(0.022)	(0.024)	(0.022)	(0.022)	(0.022)
size_inv	-834.779***	-827.783***	-832.415***	-834.873***	-834.939***
	(130.342)	(131.572)	(130.678)	(130.410)	(130.457)
RnD	-0.000**	-0.000**	-0.000**	-0.000**	-0.000**
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
log_c4_eu	0.016	-0.415***	0.005		
	(0.018)	(0.089)	(0.022)		
c4_eu_sizexconc		0.032***			
		(0.007)			
c4_eu_msxconc			0.558**		
			(0.242)		
c4_eu				0.011	
				(0.042)	
c4_eu_squ					-0.008
					(0.037)
Constant	-0.433	-0.641**	-0.448*	-0.445*	-0.429
	(0.264)	(0.281)	(0.266)	(0.264)	(0.265)
Observations	19,775	19,775	19,775	19,775	19,775
R-squared	0.260	0.263	0.260	0.260	0.260
Number of isin	1,709	1,709	1,709	1,709	1,709

Robust standard errors in parentheses
*** p<0.01, ** p<0.05, * p<0.1

Table 10.6: Regression table of ROA on EU-level C4 and control variables.

C8 Country	(1) Model 1	(2) Model 2	(3) Model 3	(4) Model 4	(5) Model 5
market_share_country	0.062***	0.057***	0.061***	0.062***	0.061***
	(0.017)	(0.018)	(0.017)	(0.017)	(0.017)
log_age	0.030***	0.024***	0.029***	0.030***	0.030***
	(0.008)	(0.007)	(0.008)	(0.008)	(0.008)
log_assets	0.034	0.040*	0.035	0.034	0.034
	(0.022)	(0.022)	(0.022)	(0.022)	(0.022)
size_inv	-835.567***	-831.524***	-834.834***	-835.548***	-835.521***
	(130.250)	(129.814)	(130.477)	(130.260)	(130.270)
RnD	-0.000**	-0.000**	-0.000**	-0.000**	-0.000**
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
log_c8_country	-0.067	-2.690**	-0.106		
	(0.102)	(1.077)	(0.118)		
c8_country_sizexconc		0.196**			
		(0.080)			
c8_country_msxconc			0.961		
			(0.791)		
c8_country				-0.103	
				(0.125)	
c8_country_squ					-0.073
					(0.074)
Constant	-0.439	-0.497*	-0.442	-0.336	-0.365
	(0.265)	(0.265)	(0.266)	(0.318)	(0.293)
Observations	19,775	19,775	19,775	19,775	19,775
R-squared	0.260	0.263	0.260	0.260	0.260
Number of isin	1,709	1,709	1,709	1,709	1,709

Robust standard errors in parentheses
*** p<0.01, ** p<0.05, * p<0.1

Table 10.7: Regression table of ROA on country-level C8 and control variables.

C8 EU	(1) Model 1	(2) Model 2	(3) Model 3	(4) Model 4	(5) Model 5
market_share_eu	0.070	0.027	0.088	0.068	0.067
	(0.074)	(0.083)	(0.072)	(0.074)	(0.074)
log_age	0.030***	0.018**	0.028***	0.029***	0.029***
	(0.008)	(0.007)	(0.008)	(0.008)	(0.008)
log_assets	0.036	0.054**	0.037	0.036	0.035
	(0.022)	(0.024)	(0.022)	(0.022)	(0.022)
size_inv	-834.916***	-822.694***	-832.347***	-834.928***	-834.917***
	(130.417)	(131.271)	(130.706)	(130.464)	(130.503)
RnD	-0.000**	-0.000**	-0.000**	-0.000**	-0.000**
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
log_c8_eu	0.000	-0.870***	-0.015		
	(0.031)	(0.196)	(0.037)		
c8_eu_sizexconc		0.064***			
		(0.014)			
c8_eu_msxconc			1.132**		
			(0.431)		
c8_eu				-0.028	
				(0.053)	
c8_eu_squ					-0.039
					(0.039)
Constant	-0.435	-0.639**	-0.450*	-0.408	-0.398
	(0.264)	(0.282)	(0.265)	(0.249)	(0.254)
Observations	19,775	19,775	19,775	19,775	19,775
R-squared	0.260	0.264	0.260	0.260	0.260
Number of isin	1,709	1,709	1,709	1,709	1,709

Robust standard errors in parentheses
*** p<0.01, ** p<0.05, * p<0.1

Table 10.8: Regression table of ROA on EU-level C8 and control variables.

HHI Country	(1) Model 1	(2) Model 2	(3) Model 3	(4) Model 4	(5) Model 5
market_share_country	0.070***	0.052**	0.070***	0.080***	0.086***
	(0.018)	(0.020)	(0.017)	(0.019)	(0.020)
log_age	0.029***	0.024***	0.029***	0.029***	0.029***
	(0.008)	(0.007)	(0.008)	(0.008)	(0.008)
log_assets	0.034	0.049**	0.034	0.034	0.034
	(0.022)	(0.023)	(0.022)	(0.022)	(0.022)
size_inv	-835.924***	-834.435***	-835.960***	-835.846***	-835.649***
	(130.403)	(131.361)	(130.501)	(130.358)	(130.289)
RnD	-0.000**	-0.000**	-0.000**	-0.000**	-0.000**
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
log_HHI_country	-0.015	-0.228**	-0.014		
	(0.011)	(0.096)	(0.014)		
HHI_country_sizexconc		0.017**			
		(0.007)			
HHI_country_msxconc			-0.002		
			(0.038)		
HHI_country				-0.051**	
				(0.023)	
HHI_country_squ					-0.049**
					(0.020)
Constant	-0.444*	-0.615**	-0.444*	-0.402	-0.413
	(0.264)	(0.276)	(0.263)	(0.268)	(0.266)
Observations	19,775	19,775	19,775	19,775	19,775
R-squared	0.260	0.263	0.260	0.261	0.261
Number of isin	1,709	1,709	1,709	1,709	1,709

Robust standard errors in parentheses
*** p<0.01, ** p<0.05, * p<0.1

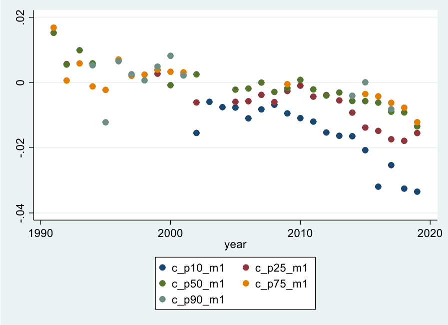
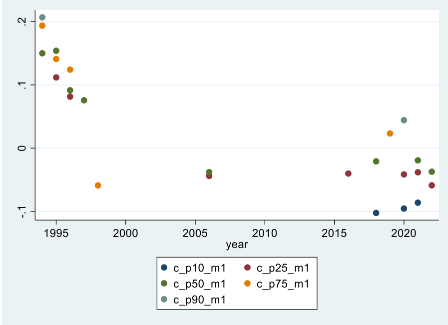
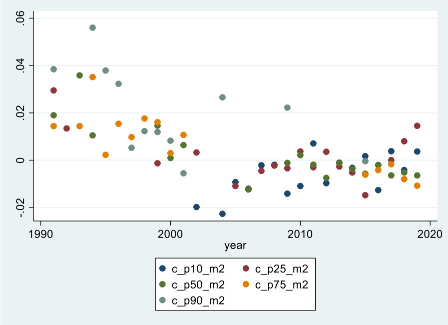
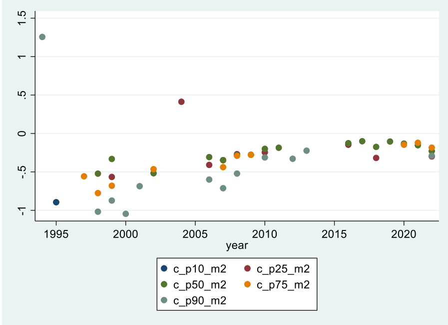
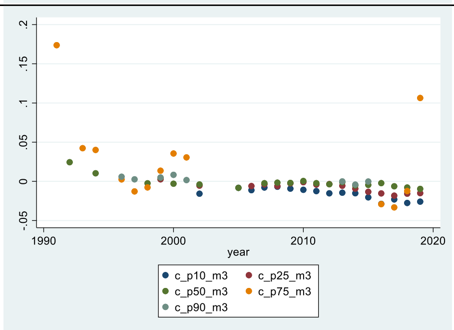
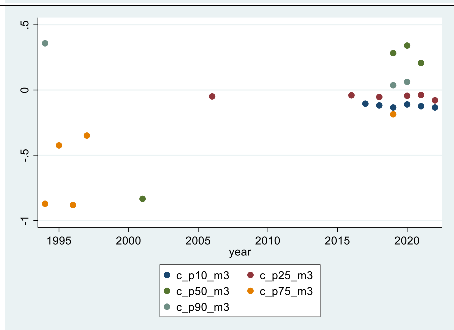
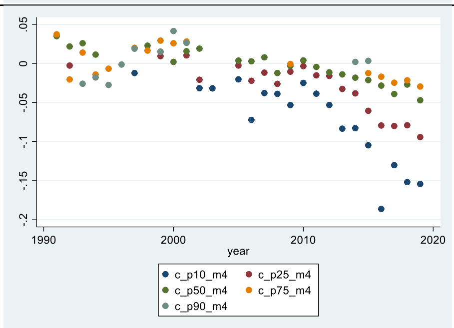
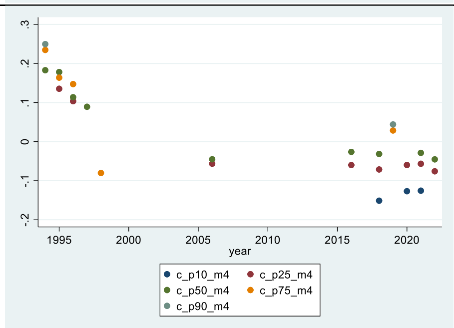
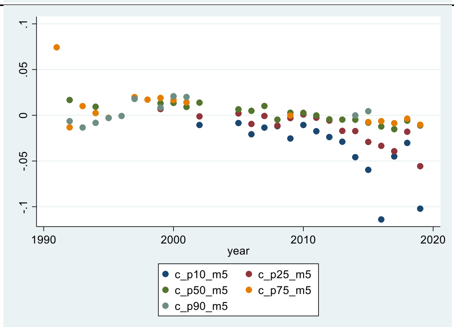
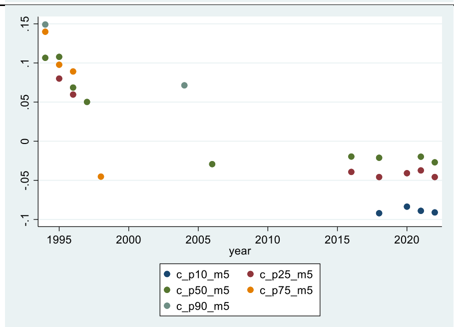
Table 10.9: Regression table of ROA on country-level HHI and control variables.

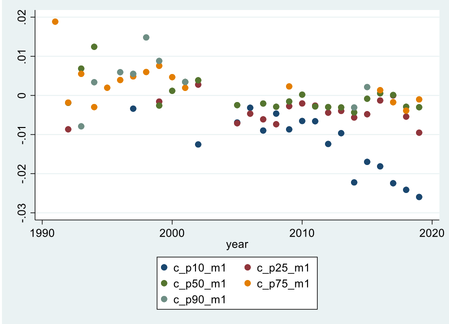
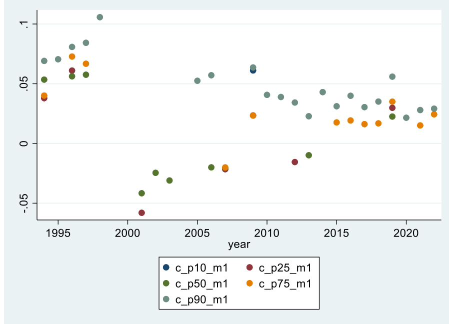
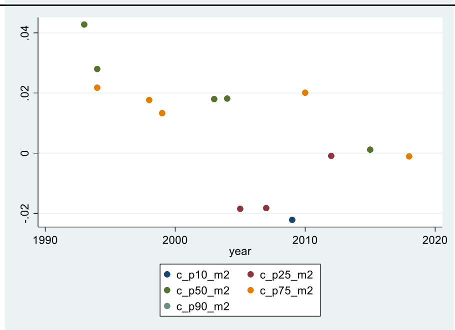
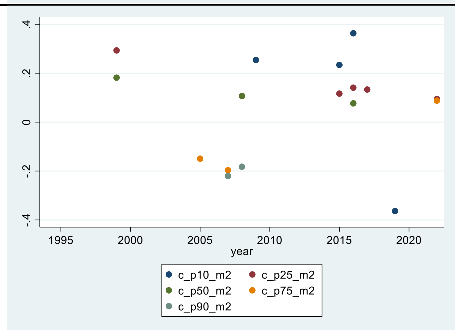
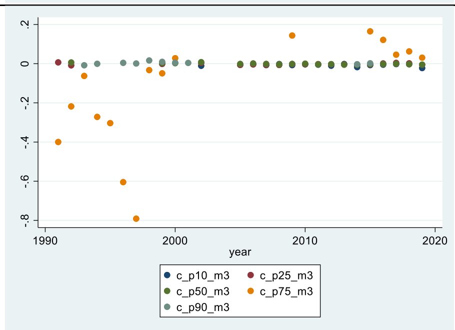
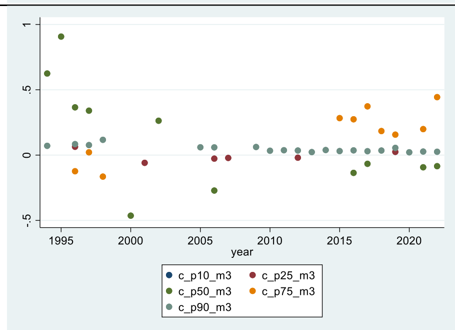
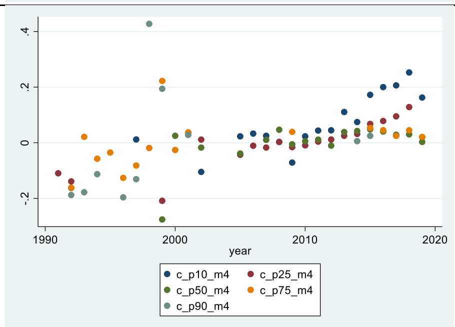
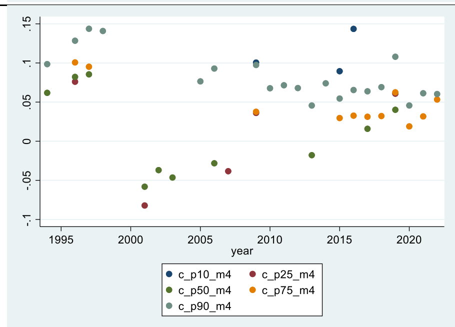
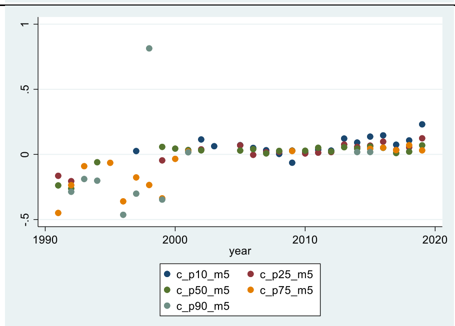
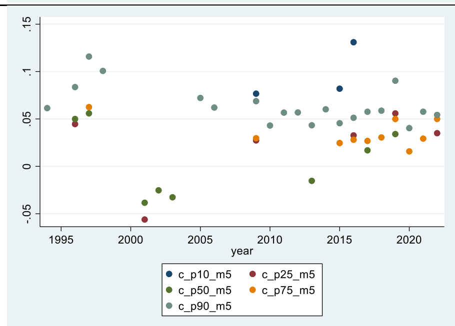
HHI EU	(1) Model 1	(2) Model 2	(3) Model 3	(4) Model 4	(5) Model 5
market_share_eu	0.069	0.046	0.117	0.072	0.081
	(0.073)	(0.079)	(0.073)	(0.073)	(0.075)
log_age	0.030***	0.024***	0.030***	0.030***	0.030***
	(0.008)	(0.007)	(0.008)	(0.008)	(0.008)
log_assets	0.036	0.060***	0.037	0.036	0.035
	(0.022)	(0.022)	(0.022)	(0.022)	(0.022)
size_inv	-834.958***	-827.907***	-833.486***	-834.841***	-834.713***
	(130.304)	(130.792)	(130.580)	(130.381)	(130.371)
RnD	-0.000**	-0.000**	-0.000**	-0.000**	-0.000**
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
log_HHI_eu	0.011	-0.136**	0.008		
	(0.009)	(0.052)	(0.011)		
HHI_eu_sizexconc		0.011***			
		(0.004)			
HHI_eu_msxconc			0.098		
			(0.071)		
HHI_eu				-0.012	
				(0.051)	
HHI_eu_squ					-0.037
					(0.052)
Constant	-0.422	-0.712***	-0.434	-0.431	-0.429
	(0.262)	(0.264)	(0.265)	(0.270)	(0.267)
Observations	19,775	19,775	19,775	19,775	19,775
R-squared	0.260	0.262	0.260	0.260	0.260
Number of isin	1,709	1,709	1,709	1,709	1,709

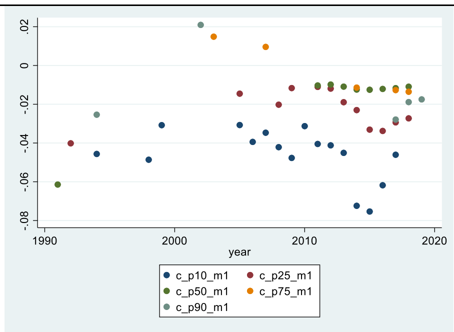
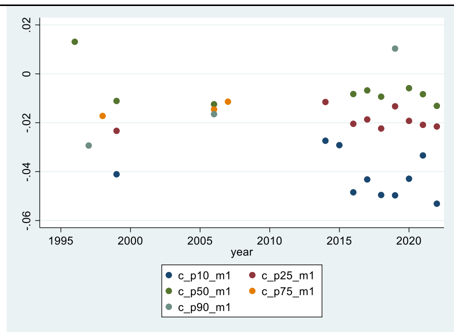
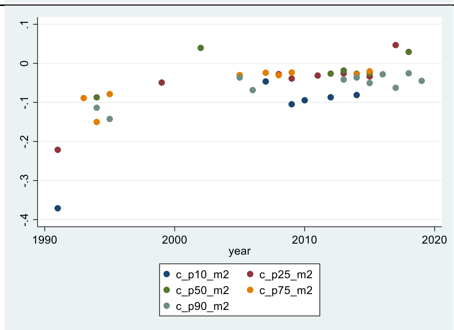
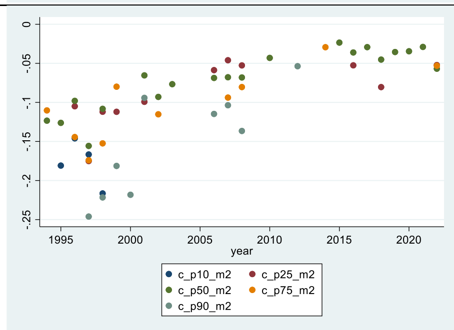
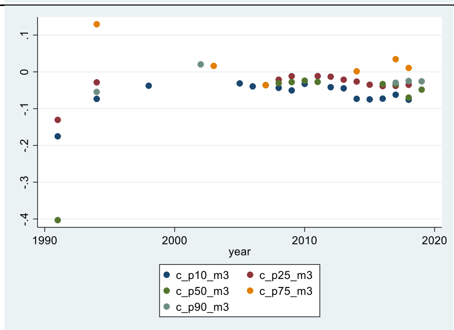
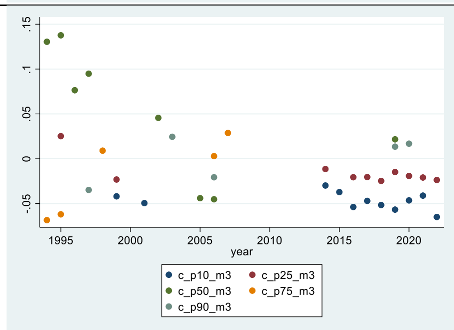
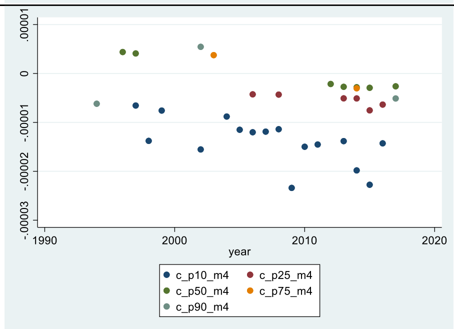
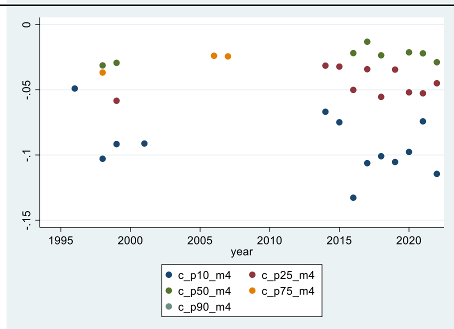
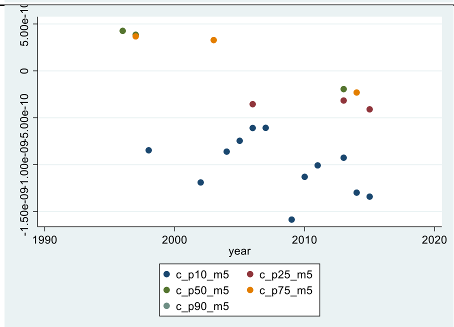
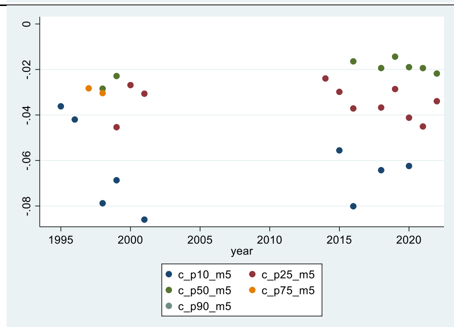
Robust standard errors in parentheses
*** p<0.01, ** p<0.05, * p<0.1

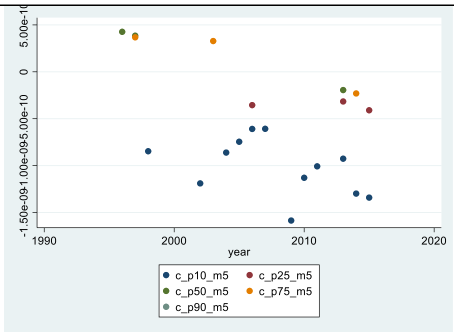
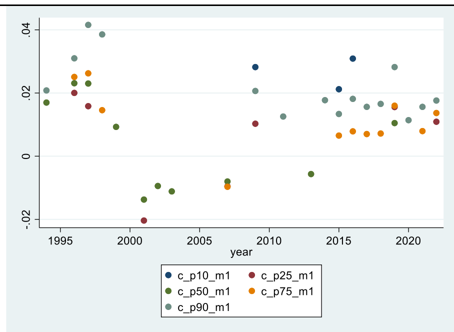
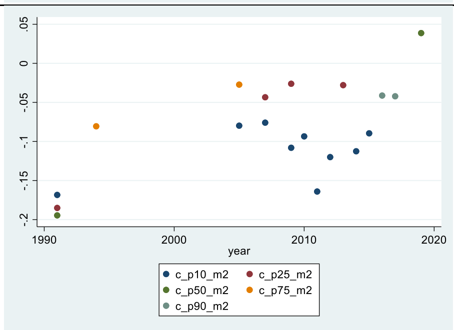
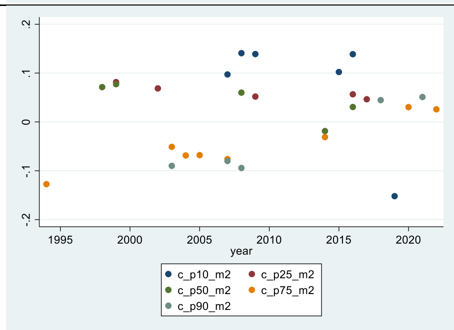
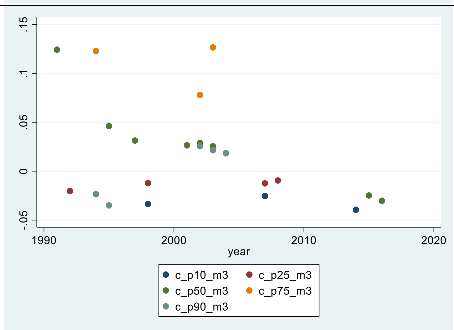
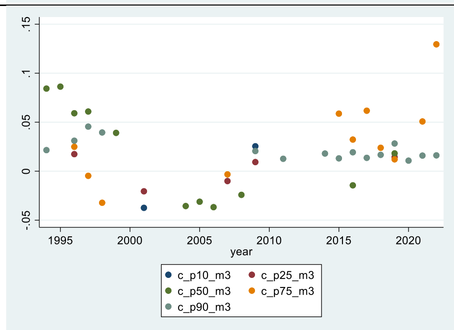
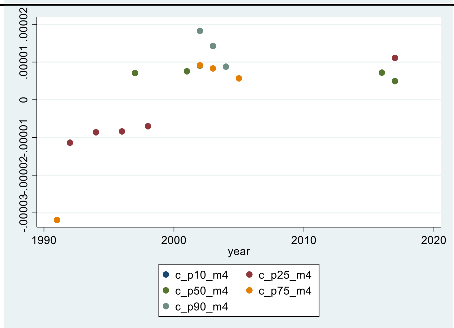
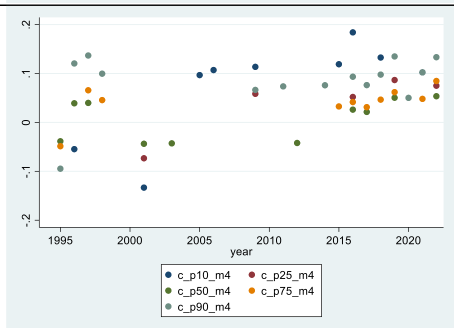
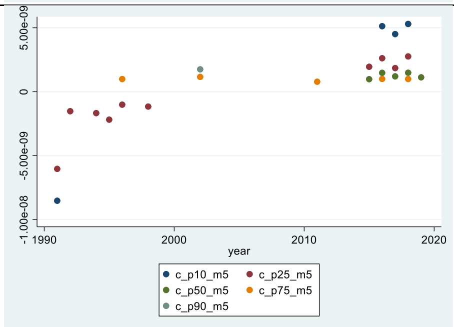
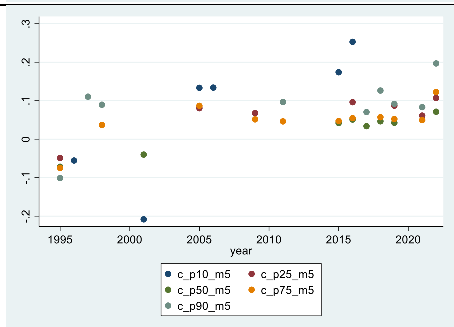
Table 10.10: Regression table of ROA on country-level HHI and control variables.

10.3 Quantile Regressions: Significant Industry Concentration Coefficients

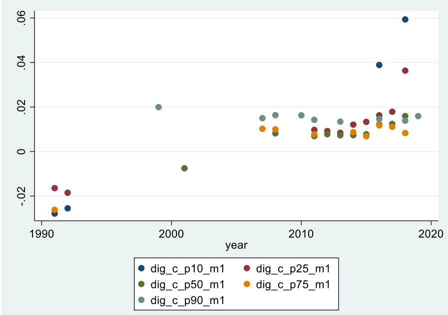
C4 Country	Compustat	Orbis
M1		
M2		
M3		
M4		
M5		

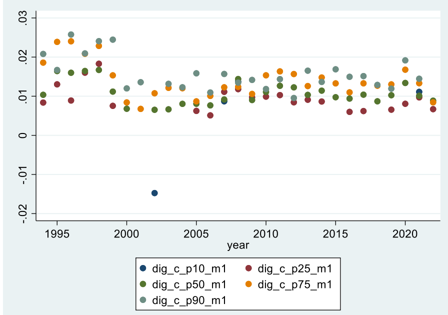
<table>
<tr><td>C4 EU</td><td>Compustat</td><td>Orbis</td></tr>
<tr><td>M1</td><td></td><td></td></tr>
<tr><td>M2</td><td></td><td></td></tr>
<tr><td>M3</td><td></td><td></td></tr>
<tr><td>M4</td><td></td><td></td></tr>
<tr><td>M5</td><td></td><td></td></tr>
</table>

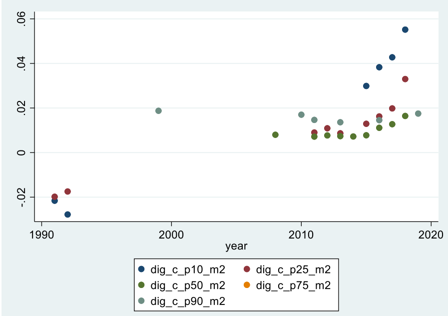
<table>
<tr><th>HHI Country</th><th>Compustat</th><th>Orbis</th></tr>
<tr><td>M1</td><td>
</td><td>
</td></tr>
<tr><td>M2</td><td>
</td><td>
</td></tr>
<tr><td>M3</td><td>
</td><td>
</td></tr>
<tr><td>M4</td><td>
</td><td>
</td></tr>
<tr><td>M5</td><td>
</td><td>
</td></tr>
</table>

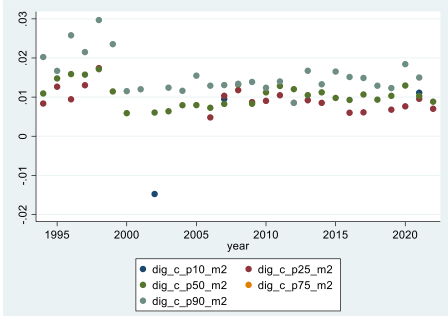
HHI EU	Compustat	Orbis
M1		
M2		
M3		
M4		
M5		

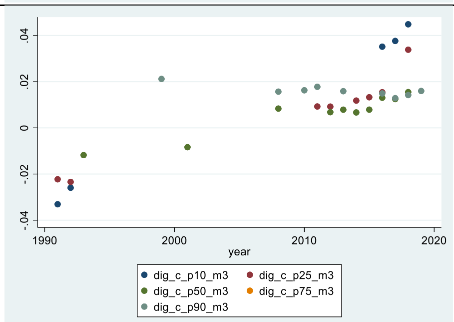
10.4 Quantile Regressions: Significant Digitization Score Coefficients

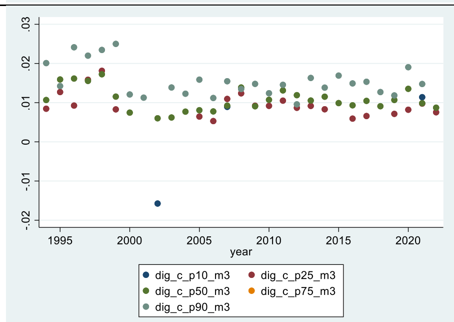
<table>
<tr><td>C4 Country</td><td>Compustat</td><td>Orbis</td></tr>
<tr><td>M1</td><td>


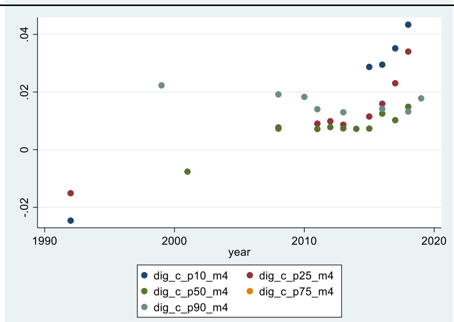
</td><td>


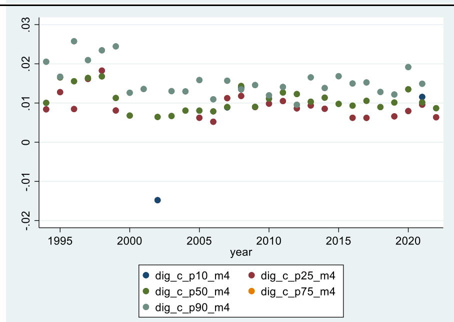
</td></tr>
<tr><td>M2</td><td>


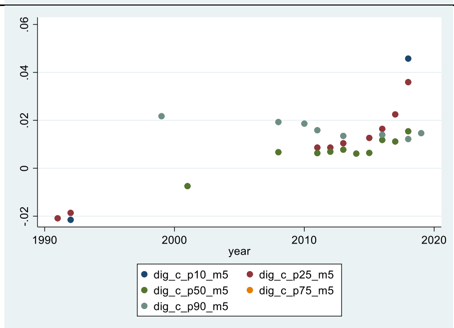
</td><td>


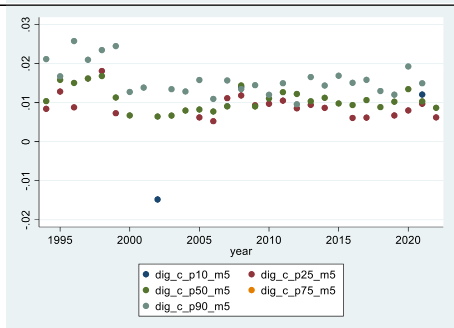
</td></tr>
<tr><td>M3</td><td>


</td><td>


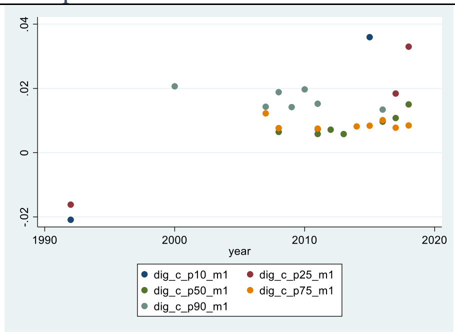
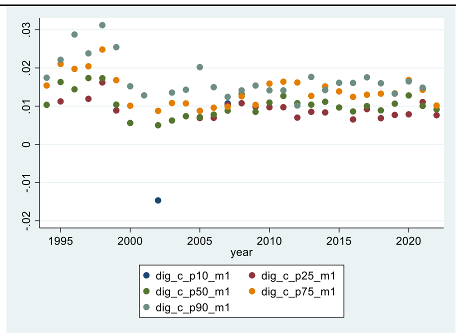
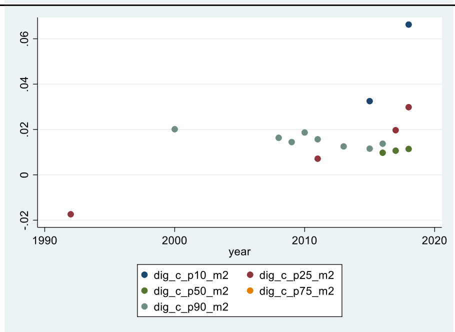
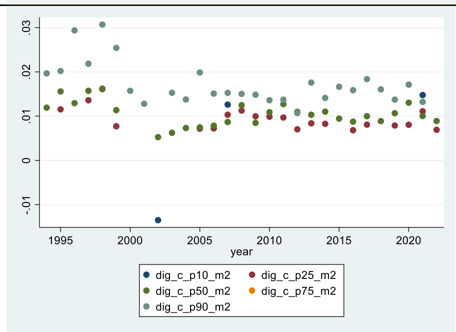
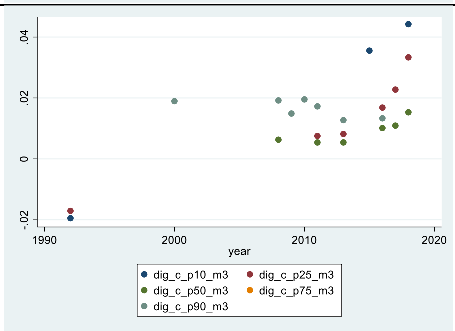
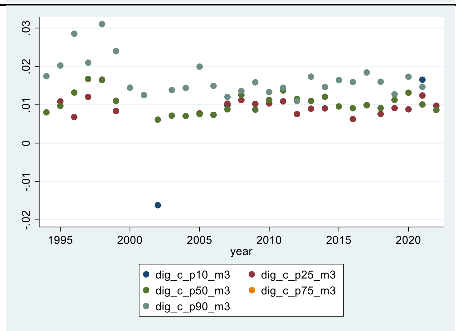
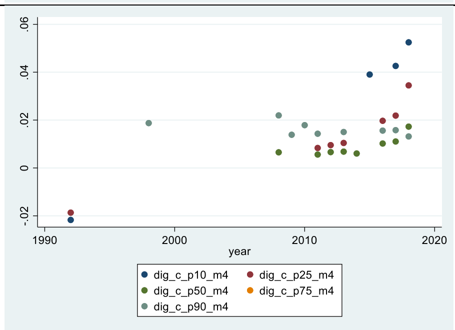
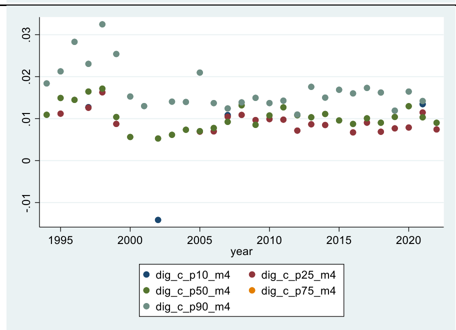
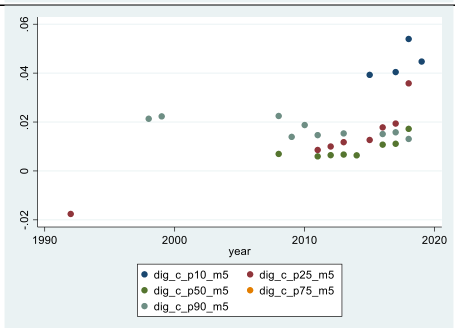
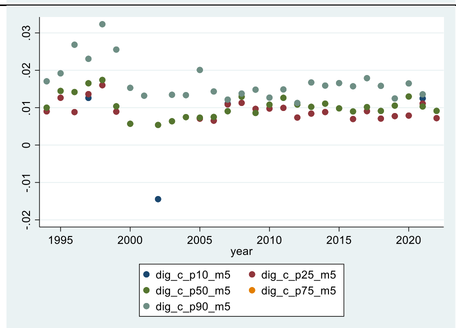
</td></tr>
<tr><td>M4</td><td>


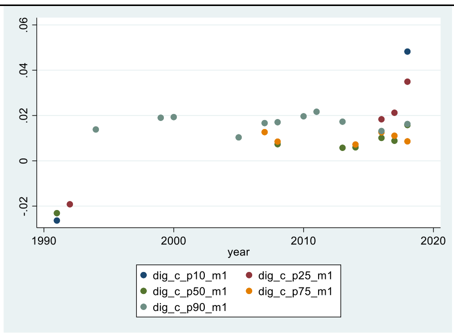
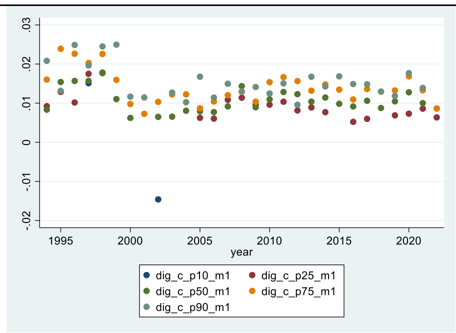
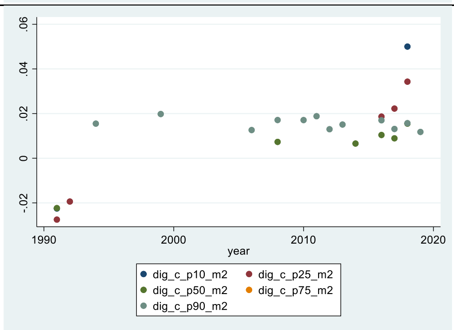
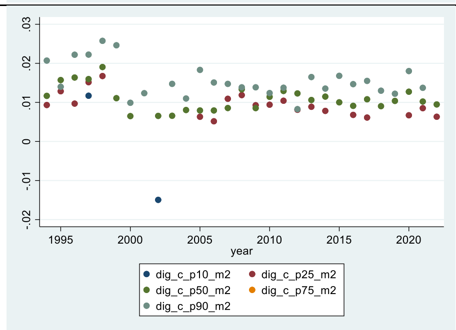
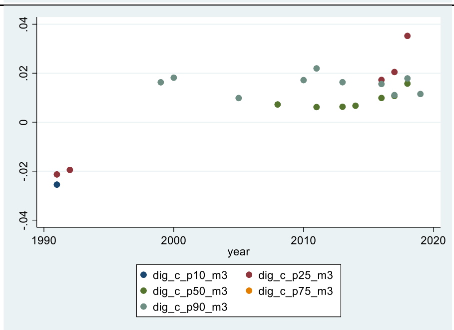
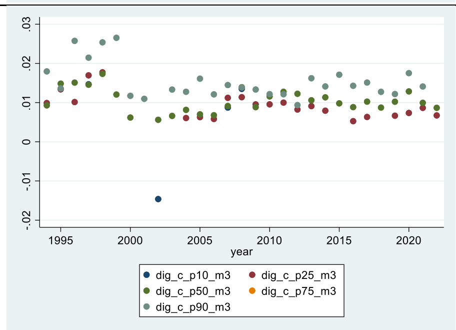
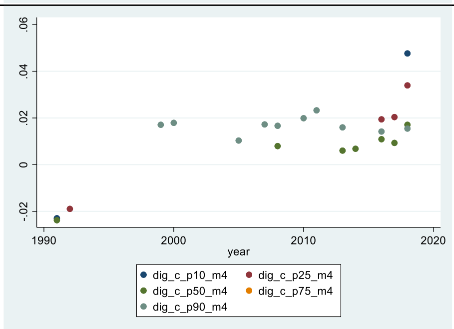
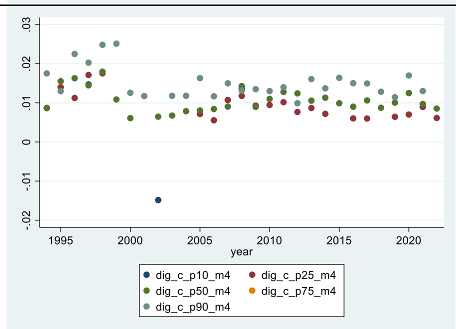
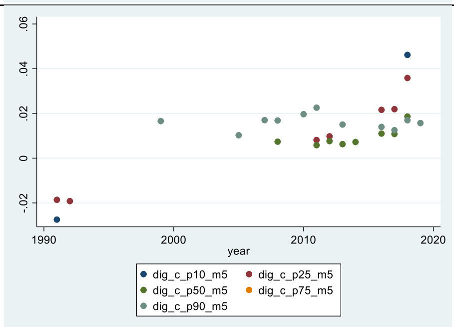
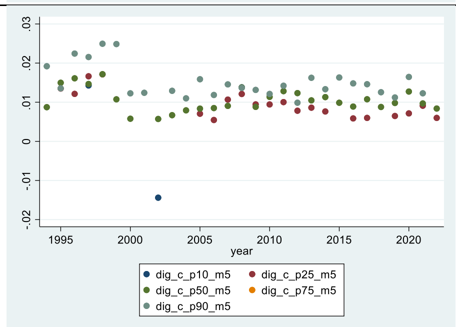
</td><td>


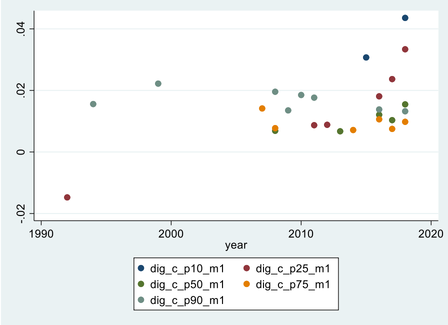
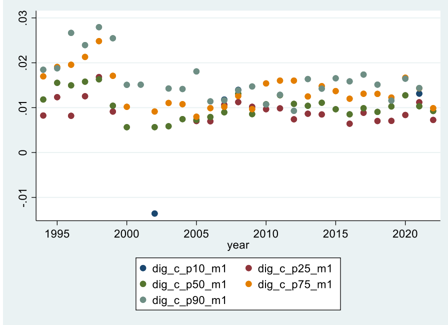
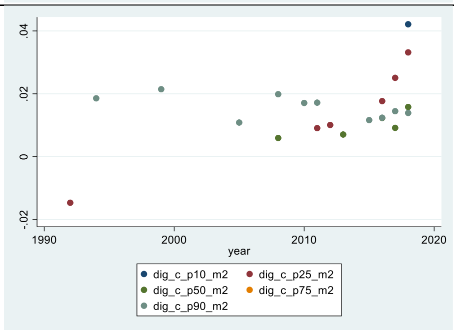
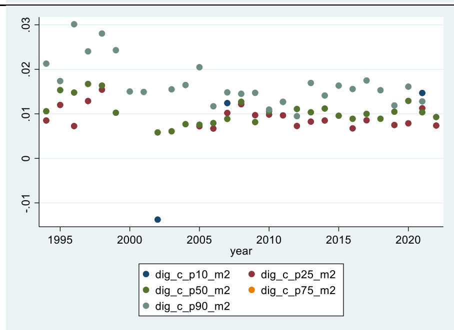
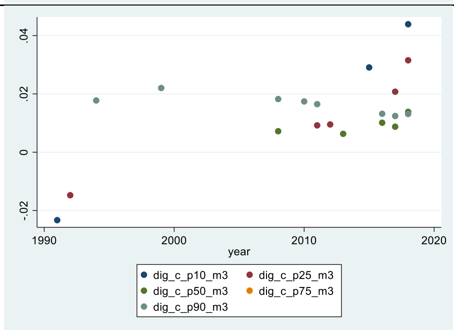
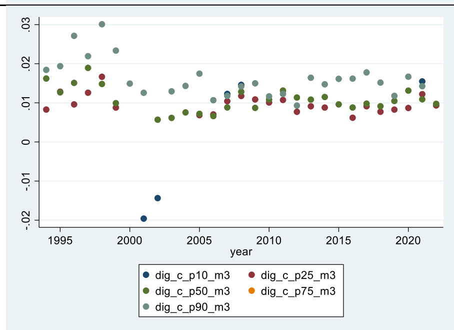
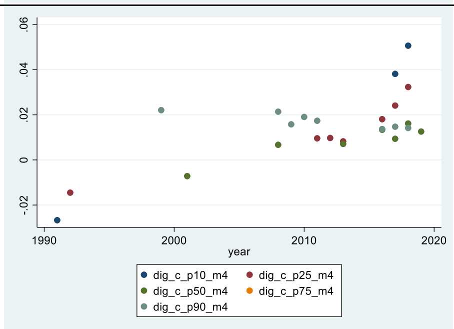
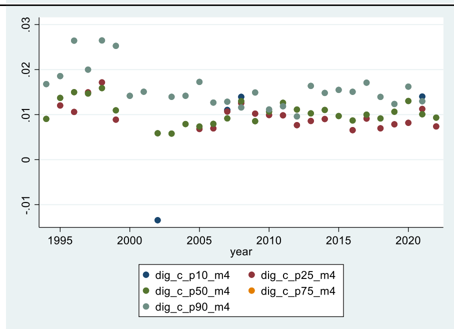
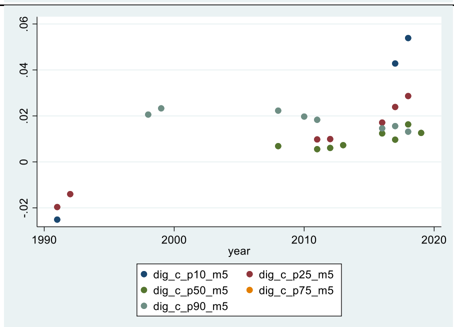
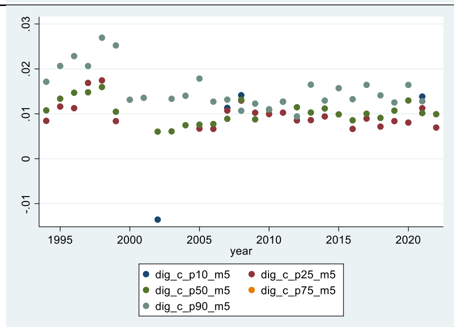
</td></tr>
<tr><td>M5</td><td>


</td><td>


</td></tr>
</table>

<table>
<tr><th>C4 EU</th><th>Compustat</th><th>Orbis</th></tr>
<tr><td>M1</td><td>
</td><td>
</td></tr>
<tr><td>M2</td><td>
</td><td>
</td></tr>
<tr><td>M3</td><td>
</td><td>
</td></tr>
<tr><td>M4</td><td>
</td><td>
</td></tr>
<tr><td>M5</td><td>
</td><td>
</td></tr>
</table>

<table>
<tr><th>HHI
Country</th><th>Compustat</th><th>Orbis</th></tr>
<tr><td>M1</td><td>
</td><td>
</td></tr>
<tr><td>M2</td><td>
</td><td>
</td></tr>
<tr><td>M3</td><td>
</td><td>
</td></tr>
<tr><td>M4</td><td>
</td><td>
</td></tr>
<tr><td>M5</td><td>
</td><td>
</td></tr>
</table>

HHI EU	Compustat	Orbis
M1		
M2		
M3		
M4		
M5		

10.5 Compustat T-Test Results

C4 Country Concentration	Mean p10	Mean p90	p-value: H_A: mean(p10) < mean(p90)
M1	-0.010	-0.001	0.001 ***
M2	-0.006	0.004	0.039 **
M3	-0.008	-0.001	0.001 ***
M4	-0.041	0.001	0.000 ***
M5	-0.014	0.004	0.005 ***

C4 Country Digitization Score	Mean p10	Mean p90	p-value H_A: mean(p10) < mean(p90)
M1	0.005	0.008	0.184
M2	0.006	0.007	0.323
M3	0.005	0.009	0.088 *
M4	0.005	0.008	0.142
M5	0.004	0.008	0.126

C4 EU Concentration	Mean p10	Mean p90	p-value H_A: mean(p10) < mean(p90)
M1	-0.008	0.003	0.000 ***
M2	0.010	-0.003	0.952
M3	-0.005	0.002	0.000 ***
M4	0.018	0.015	0.537
M5	0.003	-0.048	0.869

C4 EU Digitization Score	Mean p10	Mean p90	p-value H_A: mean(p10) < mean(p90)
M1	0.003	0.008	0.055 ***
M2	0.005	0.008	0.180
M3	0.004	0.007	0.197
M4	0.005	0.008	0.140
M5	0.005	0.008	0.222

HHI Country (N=29) Concentration	Mean p10	Mean p90	p-value H_A: mean(p10) < mean(p90)
M1	-0.033	-0.001	0.000 ***
M2	-0.039	-0.035	0.400
M3	-0.044	-0.007	0.000 ***
M4	0.000	0.000	0.000 ***
M5	0.000	0.000	0.000 ***

HHI Country Digitization Score	Mean p10	Mean p90	p-value H_A: mean(p10) < mean(p90)
M1	0.001	0.010	0.001 ***
M2	0.002	0.010	0.004 ***
M3	0.001	0.010	0.001 ***
M4	0.001	0.010	0.001 ***
M5	0.002	0.011	0.003 ***

HHI EU (N=29) Concentration	Mean p10	Mean p90	p-value H_A: mean(p10) < mean(p90)
M1	-0.013	0.001	0.000 ***
M2	-0.046	-0.012	0.012 **
M3	-0.017	0.001	0.000 ***
M4	0.000	0.000	0.083 *
M5	0.000	0.000	0.817

HHI EU Digitization Score	Mean p10	Mean p90	p-value H_A: mean(p10) < mean(p90)
M1	0.003	0.009	0.021 **
M2	0.002	0.010	0.007 ***
M3	0.003	0.009	0.022 **
M4	0.004	0.009	0.053 *
M5	0.005	0.009	0.146

10.6 Orbis T-Test Results

C4 Country Concentration (N = 29)	Mean p10	Mean p90	p-value H_A: mean(p10) < mean(p90)
M1	-0.043	0.016	0.000 ***
M2	-0.259	-0.253	0.479
M3	-0.069	0.020	0.000 ***
M4	-0.060	0.021	0.000 ***
M5	-0.040	0.013	0.000 ***

C4 Country Digitization Score (N = 29)	Mean p10	Mean p90	p-value H_A: mean(p10) < mean(p90)
M1	0.003	0.015	0.000 ***
M2	0.003	0.015	0.000 ***
M3	0.003	0.015	0.000 ***
M4	0.003	0.015	0.000 ***
M5	0.003	0.015	0.000 ***

C4 EU Concentration (N = 29)	Mean p10	Mean p90	p-value H_A: mean(p10) < mean(p90)
M1	0.014	0.039	0.000 ***
M2	0.022	-0.007	0.734
M3	0.000	0.040	0.000 ***
M4	0.029	0.062	0.000 ***
M5	0.027	0.047	0.004 ***

C4 EU Digitization Score (N = 29)	Mean p10	Mean p90	p-value H_A: mean(p10) < mean(p90)
M1	0.004	0.017	0.000 ***
M2	0.004	0.017	0.000 ***
M3	0.005	0.016	0.000 ***
M4	0.004	0.017	0.000 ***
M5	0.004	0.016	0.000 ***

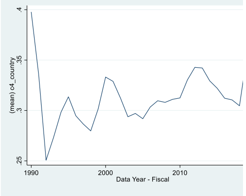
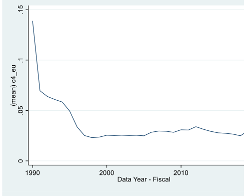
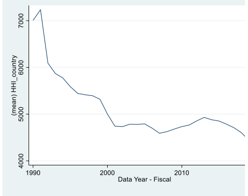
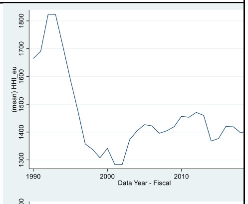
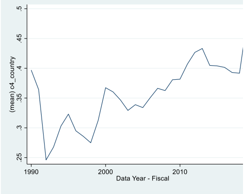
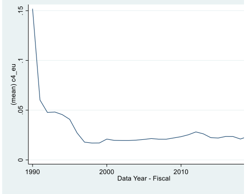
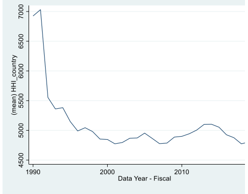
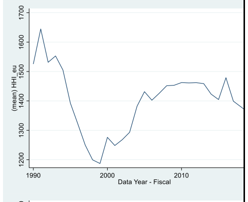
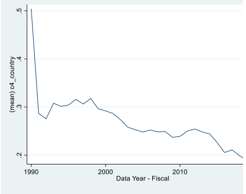
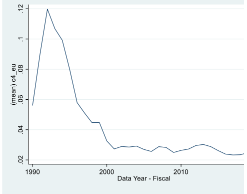
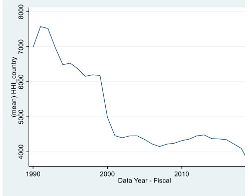
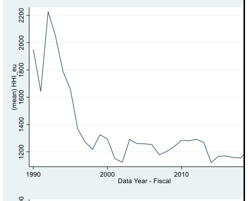
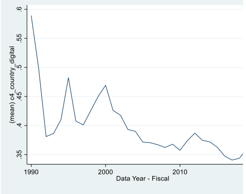
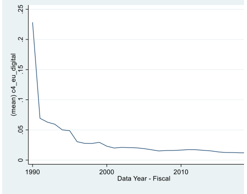
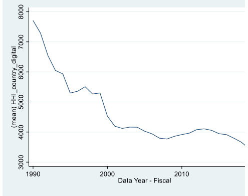
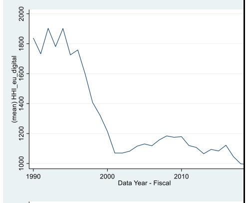
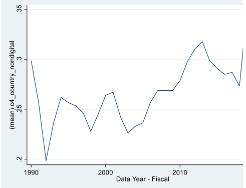
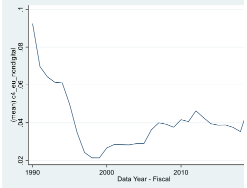
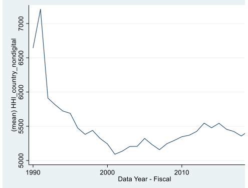
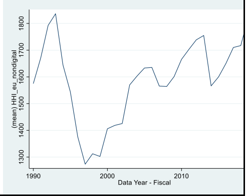
HHI Country (N=29) Concentration	Mean p10	Mean p90	p-value H_A: mean(p10) < mean(p90)
M1	-0.023	-0.002	0.000 ***
M2	-0.048	-0.078	0.993
M3	-0.026	0.000	0.000 ***
M4	-0.055	-0.008	0.000 ***
M5	-0.038	-0.010	0.000 ***

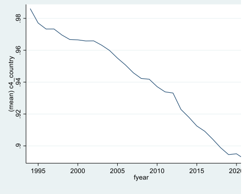
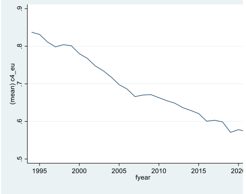
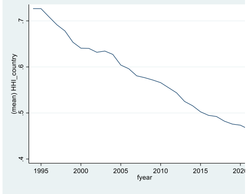
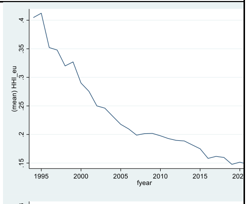
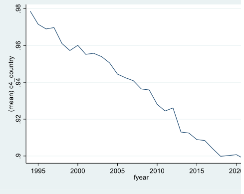
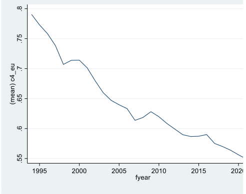
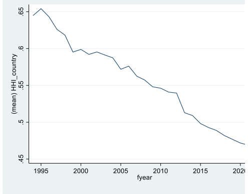
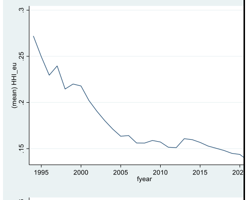
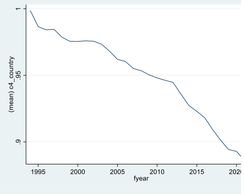
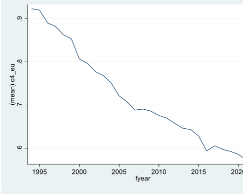
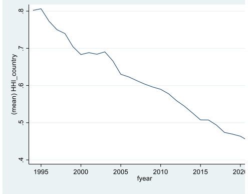
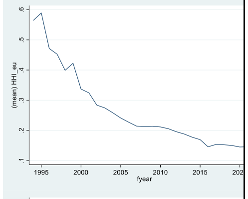
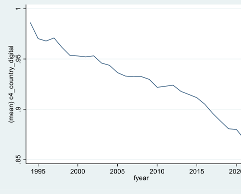
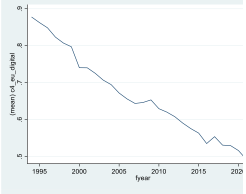
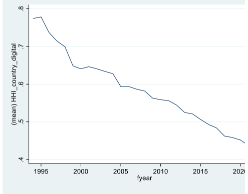
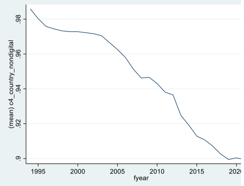
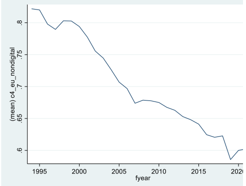
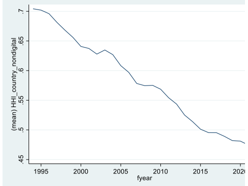
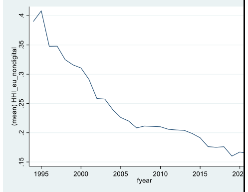
HHI Country Digitization Score	Mean p10	Mean p90	p-value H_A: mean(p10) < mean(p90)
M1	0.003	0.015	0.000 ***
M2	0.003	0.015	0.000 ***
M3	0.003	0.015	0.000 ***
M4	0.003	0.015	0.000 ***
M5	0.003	0.015	0.000 ***

HHI EU (N=29) Concentration	Mean p10	Mean p90	p-value H_A: mean(p10) < mean(p90)
M1	0.009	0.013	0.058 *
M2	0.026	-0.003	0.926
M3	0.006	0.014	0.001 ***
M4	0.055	0.046	0.751
M5	0.068	0.038	0.969

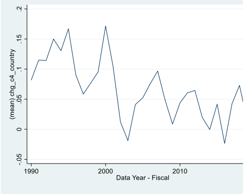
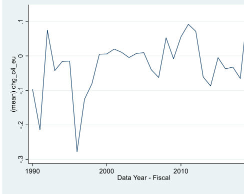
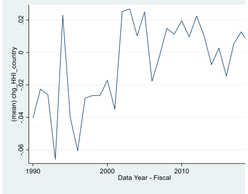
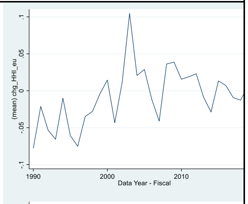
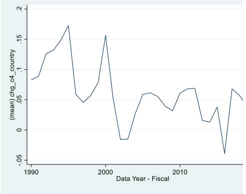
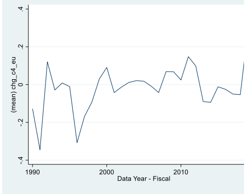
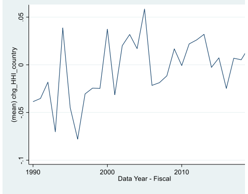
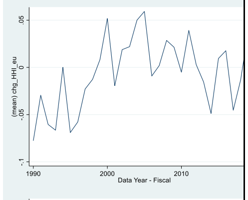
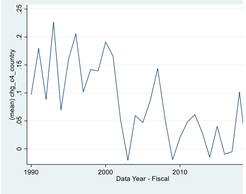
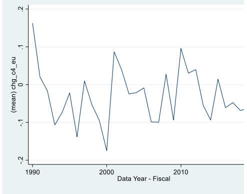
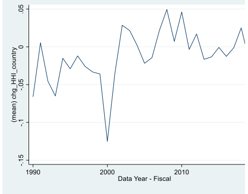
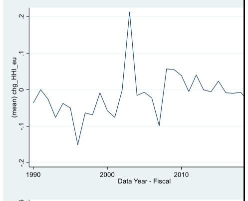
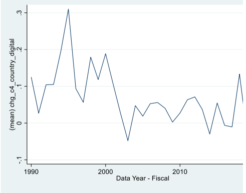
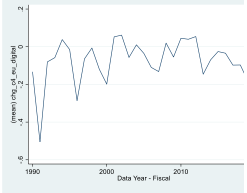
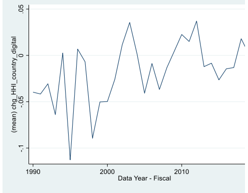
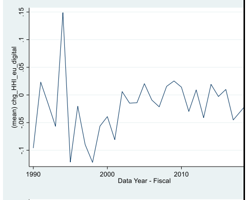
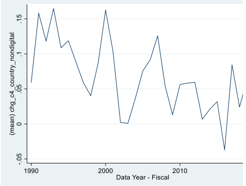
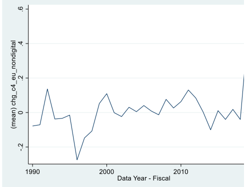
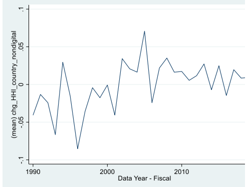
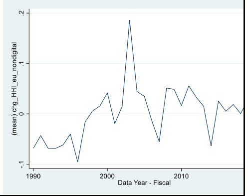
HHI EU Digitization Score	Mean p10	Mean p90	p-value H_A: mean(p10) < mean(p90)
M1	0.005	0.016	0.000 ***
M2	0.004	0.016	0.000 ***
M3	0.005	0.016	0.000 ***
M4	0.005	0.015	0.000 ***
M5	0.005	0.015	0.000 ***

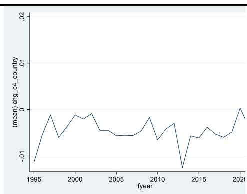
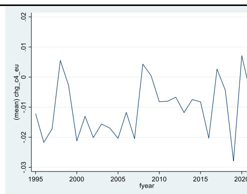
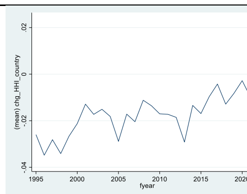
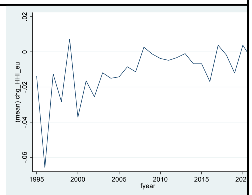
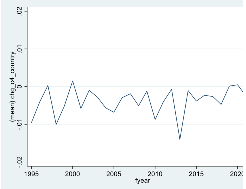
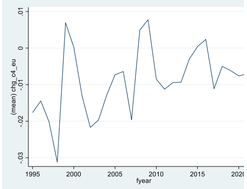
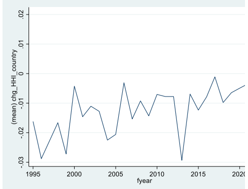
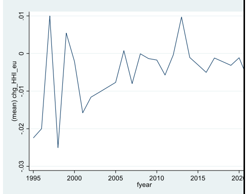
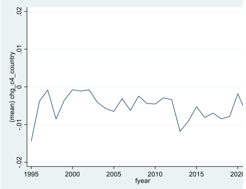
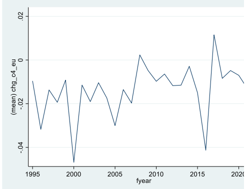
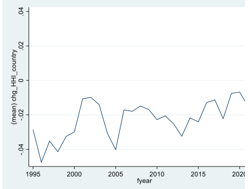
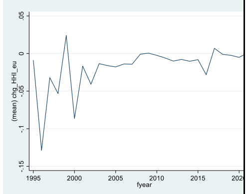
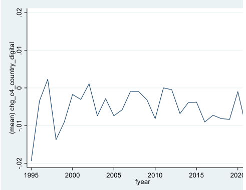
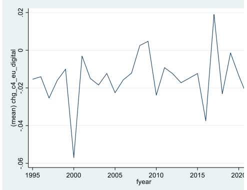
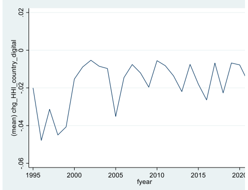
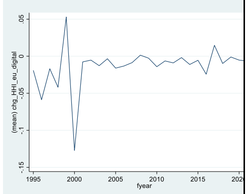
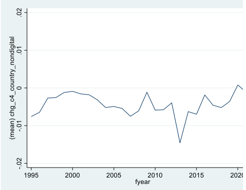
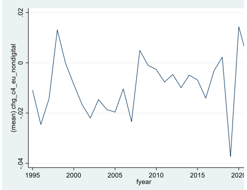
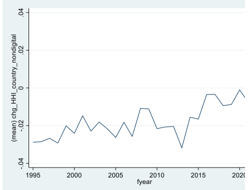
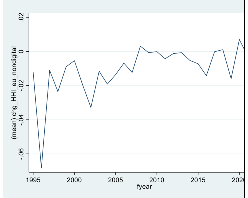
10.7 Absolute Industry Concentration Trends

Compustat Absolute Concentration	C4 Country	C4 EU	HHI Country	HHI EU
All Industries				
Manufacturing				
Services				
Digital				
Non-Digital				

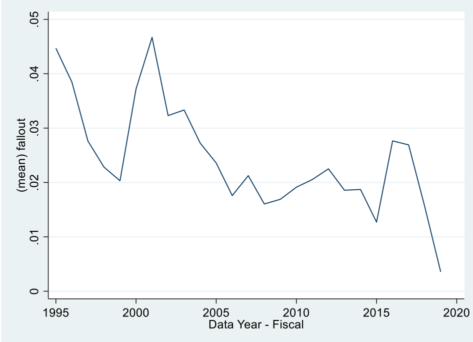
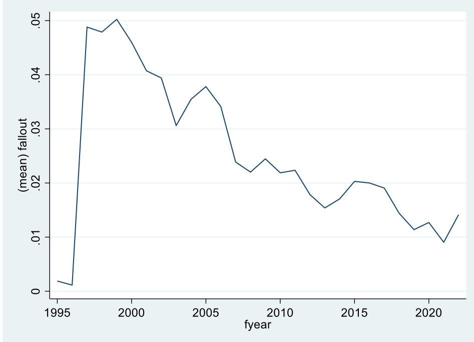
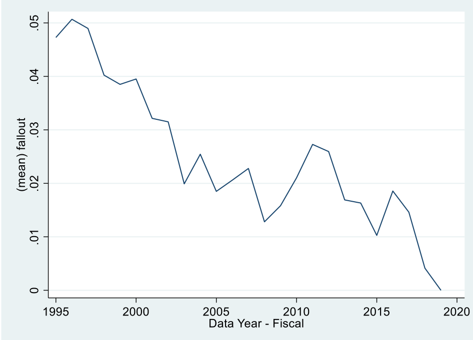
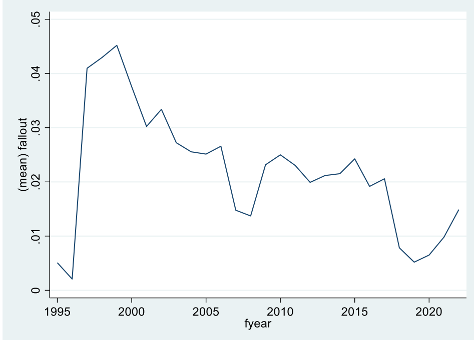
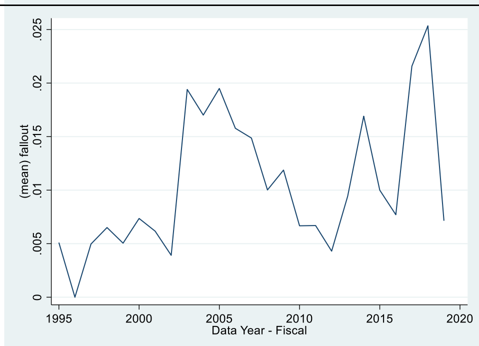
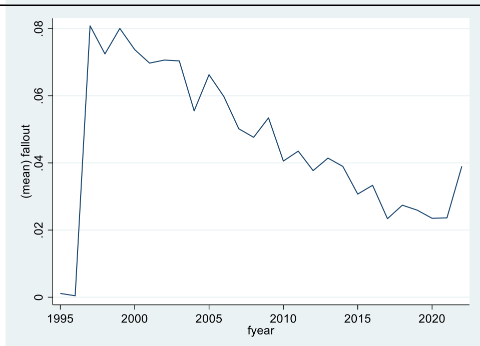
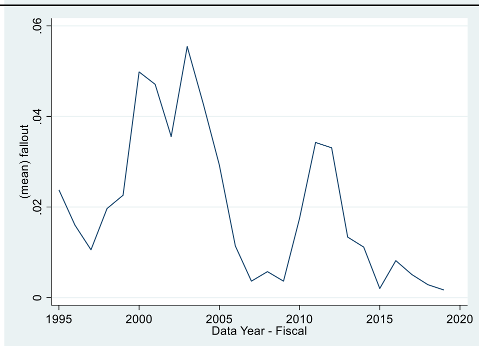
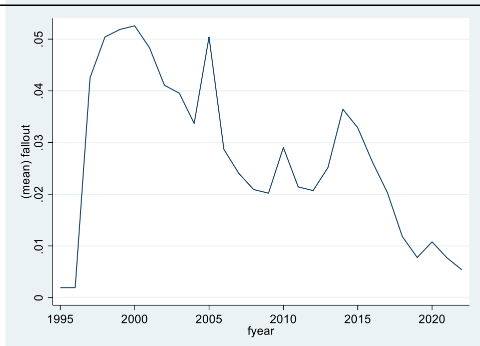
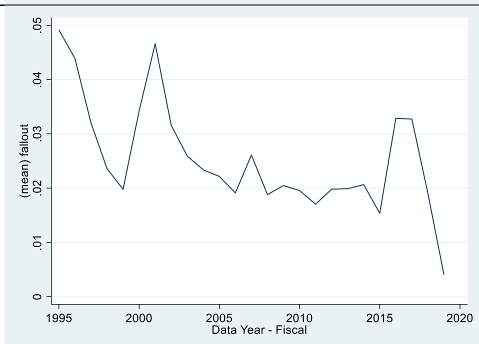
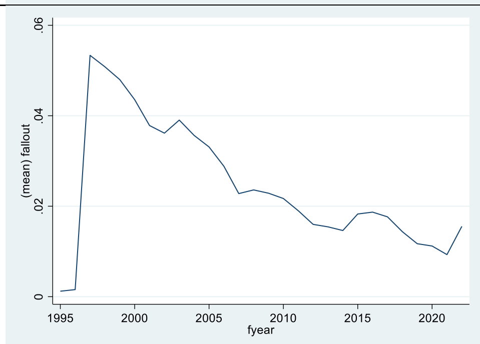
Orbis Absolute Concentration	C4 Country	C4 EU	HHI Country	HHI EU
All Industries				
Manufacturing				
Services				
Digital				
Non-Digital				

10.8 Average Changes in Industry Concentration

Compustat Concentration Changes	C4 Country	C4 EU	HHI Country	HHI EU
All Industries				
Manufacturing				
Services				
Digital				
Non-Digital				

Orbis Concentrati on Changes	C4 Country	C4 EU	HHI Country	HHI EU
All Industries				
Manufacturi ng				
Services				
Digital				
Non-Digital				

10.9 Details Fallout Ratio

<table>
<tr><th>Fallout Ratio</th><th>Compustat</th><th>Orbis</th></tr>
<tr><td>Total</td><td></td><td></td></tr>
<tr><td>Manufacturin
g</td><td></td><td></td></tr>
<tr><td>Services</td><td></td><td></td></tr>
<tr><td>Digital</td><td></td><td></td></tr>
<tr><td>Non-Digital</td><td></td><td></td></tr>
</table>

10.10 Robustness Checks

VARIABLES	(1) Model 1	(2) Model 2	(3) Model 3	(4) Model 4	(5) Model 5
market_share	-0.003	-0.009	0.000	0.010	0.012
	(0.033)	(0.033)	(0.108)	(0.035)	(0.038)
log_age	0.002	0.003	0.002	0.002	0.002
	(0.012)	(0.012)	(0.012)	(0.012)	(0.012)
log_assets	0.021**	0.020**	0.021**	0.021**	0.020**
	(0.009)	(0.009)	(0.009)	(0.009)	(0.009)
size_inv	-0.238*	-0.237*	-0.238*	-0.237*	-0.239*
	(0.142)	(0.143)	(0.142)	(0.143)	(0.143)
RnD	-0.501***	-0.501***	-0.501***	-0.502***	-0.502***
	(0.087)	(0.087)	(0.087)	(0.087)	(0.087)
CapEx	0.163**	0.164**	0.163**	0.163**	0.163**
	(0.064)	(0.064)	(0.065)	(0.064)	(0.064)
log_concentration	-0.043***	-0.058	-0.043***		
	(0.014)	(0.040)	(0.015)		
sizexconc		0.003			
		(0.005)			
msxconc			-0.003		
			(0.065)		
concentration				-0.025***	
				(0.008)	
concentration_squ					-0.004**
					(0.002)
Constant	-0.043	-0.037	-0.043	-0.016	-0.045
	(0.064)	(0.065)	(0.064)	(0.065)	(0.065)
Observations	5,868	5,868	5,868	5,868	5,868
R-squared	0.284	0.284	0.284	0.284	0.283
Number of isin	798	798	798	798	798

Robust standard errors in parentheses
*** p<0.01, ** p<0.05, * p<0.1

Table 10.11: Panel regression results using the post-2000 Compustat sample excluding industries 26 and 58.

VARIABLES	(1) Model 1	(2) Model 2	(3) Model 3	(4) Model 4	(5) Model 5
market_share	-0.107	-0.141	0.195***	-0.117	-0.128
	(0.132)	(0.150)	(0.069)	(0.141)	(0.148)
log_age	-0.015	-0.033	-0.017	-0.015	-0.014
	(0.025)	(0.032)	(0.025)	(0.024)	(0.024)
log_assets	0.192**	0.287*	0.195**	0.192**	0.192**
	(0.096)	(0.147)	(0.098)	(0.096)	(0.097)
size_inv	-98.177	-98.141	-98.113	-98.173	-98.169
	(164.543)	(164.495)	(164.500)	(164.540)	(164.537)
log_concentration	0.048	-1.887	-0.021		
	(0.081)	(1.173)	(0.055)		
sizexconc		0.160			
		(0.101)			
msxconc			1.016*		
			(0.550)		
concentration				0.124	
				(0.162)	
concentration_squ					0.132
					(0.143)
Constant	-2.192**	-3.279**	-2.246**	-2.296**	-2.274**
	(1.035)	(1.608)	(1.062)	(1.159)	(1.123)
Observations	67,696	67,696	67,696	67,696	67,696
R-squared	0.041	0.042	0.041	0.041	0.041
Number of isin	5,451	5,451	5,451	5,451	5,451

Robust standard errors in parentheses
*** $p<0.01$, ** $p<0.05$, * $p<0.1$

Table 10.12: Panel regression results using the post-2000 Orbis sample excluding industries 26 and 58.

VARIABLES	(1) Model 1	(2) Model 2	(3) Model 3	(4) Model 4	(5) Model 5
market_share	-0.015	-0.052	-0.078	0.007	0.021
	(0.036)	(0.040)	(0.129)	(0.040)	(0.046)
log_age	0.008	0.009	0.008	0.008	0.008
	(0.011)	(0.011)	(0.011)	(0.012)	(0.012)
log_assets	0.020**	0.013	0.021**	0.020**	0.019*
	(0.010)	(0.010)	(0.010)	(0.010)	(0.010)
size_inv	-0.425*	-0.415*	-0.423*	-0.423*	-0.426*
	(0.242)	(0.241)	(0.241)	(0.242)	(0.243)
RnD	-0.540***	-0.535***	-0.540***	-0.541***	-0.541***
	(0.092)	(0.091)	(0.092)	(0.091)	(0.092)
CapEx	0.159***	0.171***	0.159***	0.158***	0.158**
	(0.061)	(0.061)	(0.061)	(0.061)	(0.061)
log_concentration	-0.059***	-0.170***	-0.061***		
	(0.016)	(0.053)	(0.017)		
sizexconc		0.019***			
		(0.007)			
msxconc			0.046		
			(0.080)		
concentration				-0.042***	
				(0.012)	
concentration_squ					-0.009***
					(0.003)
Constant	-0.035	0.004	-0.036	0.013	-0.028
	(0.073)	(0.074)	(0.073)	(0.076)	(0.074)
Observations	6,922	6,922	6,922	6,922	6,922
R-squared	0.280	0.284	0.280	0.281	0.280
Number of isin	836	836	836	836	836

Robust standard errors in parentheses
*** p<0.01, ** p<0.05, * p<0.1

Table 10.13: Panel regression results using the post-2000 Compustat sample consisting only of manufacturing industries.

VARIABLES	(1) Model 1	(2) Model 2	(3) Model 3	(4) Model 4	(5) Model 5
market_share	0.040	-0.006	0.202	0.046	0.044
	(0.082)	(0.060)	(0.129)	(0.084)	(0.081)
log_age	0.023	0.012	0.022	0.024	0.024
	(0.016)	(0.011)	(0.016)	(0.017)	(0.017)
log_assets	0.059***	0.127***	0.060***	0.059***	0.059***
	(0.017)	(0.037)	(0.017)	(0.017)	(0.017)
size_inv	-19.817	-18.969	-19.769	-19.779	-19.750
	(33.873)	(33.470)	(33.851)	(33.878)	(33.875)
log_concentration	-0.154	-1.409	-0.189*		
	(0.097)	(0.990)	(0.107)		
sizexconc		0.104			
		(0.074)			
msxconc			0.438**		
			(0.180)		
concentration				-0.276*	
				(0.159)	
concentration_squ					-0.198*
					(0.107)
Constant	-0.878***	-1.659***	-0.900***	-0.627***	-0.718***
	(0.136)	(0.526)	(0.136)	(0.229)	(0.187)
Observations	30,953	30,953	30,953	30,953	30,953
R-squared	0.025	0.032	0.026	0.025	0.025
Number of isin	2,266	2,266	2,266	2,266	2,266

Robust standard errors in parentheses
*** p<0.01, ** p<0.05, * p<0.1

Table 10.14: Panel regression results using the post-2000 Orbis sample consisting only of manufacturing industries.

VARIABLES	(1) Model 1	(2) Model 2	(3) Model 3	(4) Model 4	(5) Model 5
market_share	0.535***	0.609***	1.047**	0.610***	0.577***
	(0.181)	(0.179)	(0.434)	(0.195)	(0.194)
log_age	0.024	0.024	0.023	0.025	0.025
	(0.023)	(0.023)	(0.023)	(0.023)	(0.023)
log_assets	-0.032*	-0.035**	-0.037**	-0.033*	-0.032*
	(0.017)	(0.016)	(0.018)	(0.017)	(0.017)
size_inv	-0.397***	-0.410***	-0.412***	-0.402***	-0.405***
	(0.097)	(0.100)	(0.102)	(0.097)	(0.099)
RnD	-0.377***	-0.376***	-0.382***	-0.380***	-0.380***
	(0.092)	(0.092)	(0.090)	(0.092)	(0.091)
CapEx	0.192	0.185	0.198	0.195	0.193
	(0.133)	(0.132)	(0.133)	(0.133)	(0.133)
log_concentration	-0.047	-0.007	-0.034		
	(0.032)	(0.063)	(0.036)		
sizexconc		-0.009			
		(0.010)			
msxconc			-0.413		
			(0.282)		
concentration				-0.041**	
				(0.021)	
concentration_squ					-0.008*
					(0.004)
Constant	0.174	0.190	0.191	0.227*	0.197*
	(0.117)	(0.117)	(0.120)	(0.116)	(0.118)
Observations	2,097	2,097	2,097	2,097	2,097
R-squared	0.142	0.143	0.144	0.142	0.141
Number of isin	389	389	389	389	389

Robust standard errors in parentheses
*** p<0.01, ** p<0.05, * p<0.1

Table 10.15: Panel regression results using the post-2000 Compustat sample consisting only of services industries (45-99).

VARIABLES	(1) Model 1	(2) Model 2	(3) Model 3	(4) Model 4	(5) Model 5
market_share	-0.117	-0.163	0.265*	-0.148	-0.175
	(0.173)	(0.207)	(0.142)	(0.196)	(0.216)
log_age	0.002	-0.023	-0.002	0.003	0.003
	(0.038)	(0.058)	(0.040)	(0.037)	(0.036)
log_assets	0.224	0.331	0.229	0.225	0.225
	(0.148)	(0.235)	(0.151)	(0.149)	(0.149)
size_inv	-92.778	-92.710	-92.679	-92.761	-92.749
	(160.259)	(160.175)	(160.187)	(160.246)	(160.238)
log_concentration	0.237	-1.857	0.165		
	(0.183)	(1.700)	(0.134)		
sizexconc		0.174			
		(0.156)			
msxconc			1.386		
			(0.984)		
concentration				0.508	
				(0.393)	
concentration_squ					0.448
					(0.354)
Constant	-2.447	-3.649	-2.514	-2.885	-2.754
	(1.493)	(2.461)	(1.539)	(1.815)	(1.721)
Observations	35,633	35,633	35,633	35,633	35,633
R-squared	0.046	0.047	0.046	0.046	0.046
Number of isin	3,270	3,270	3,270	3,270	3,270

Robust standard errors in parentheses
*** p<0.01, ** p<0.05, * p<0.1

Table 10.16: Panel regression results using the post-2000 Orbis sample consisting only of services industries (45-99).

VARIABLES	(1) Model 1	(2) Model 2	(3) Model 3	(4) Model 4	(5) Model 5
market_share	0.053	0.001	0.301	0.073	0.078
	(0.057)	(0.052)	(0.220)	(0.065)	(0.067)
log_age	0.024	0.027*	0.024	0.024	0.025*
	(0.015)	(0.015)	(0.015)	(0.015)	(0.015)
log_assets	0.014	0.007	0.012	0.013	0.013
	(0.012)	(0.012)	(0.013)	(0.012)	(0.012)
size_inv	-0.372***	-0.351***	-0.379***	-0.374***	-0.376***
	(0.129)	(0.127)	(0.130)	(0.129)	(0.130)
RnD	-0.457***	-0.450***	-0.459***	-0.458***	-0.458***
	(0.076)	(0.075)	(0.077)	(0.076)	(0.076)
CapEx	0.167*	0.178**	0.163*	0.164*	0.162*
	(0.086)	(0.085)	(0.086)	(0.086)	(0.086)
log_concentration	-0.060***	-0.208***	-0.051**		
	(0.020)	(0.059)	(0.021)		
sizexconc		0.026***			
		(0.008)			
msxconc			-0.201		
			(0.160)		
concentration				-0.044***	
				(0.014)	
concentration_squ					-0.010***
					(0.003)
Constant	-0.047	-0.023	-0.043	0.004	-0.035
	(0.086)	(0.084)	(0.087)	(0.091)	(0.087)
Observations	5,543	5,543	5,543	5,543	5,543
R-squared	0.221	0.228	0.221	0.222	0.221
Number of isin	714	714	714	714	714

Robust standard errors in parentheses
*** p<0.01, ** p<0.05, * p<0.1

Table 10.17: Panel regression results using the post-2000 Compustat sample consisting only of Northern European firms.

VARIABLES	(1) Model 1	(2) Model 2	(3) Model 3	(4) Model 4	(5) Model 5
market_share	0.498***	0.490***	0.214***	0.514***	0.532***
	(0.108)	(0.112)	(0.071)	(0.112)	(0.117)
log_age	-0.053*	-0.055*	-0.051*	-0.054**	-0.055**
	(0.027)	(0.029)	(0.027)	(0.027)	(0.027)
log_assets	-0.171***	-0.159**	-0.175***	-0.172***	-0.172***
	(0.059)	(0.077)	(0.060)	(0.059)	(0.059)
size_inv	-1,559.612***	-1,559.457***	-1,560.007***	-1,559.647***	-1,559.679***
	(222.128)	(222.217)	(222.022)	(222.118)	(222.106)
log_concentration	-0.047	-0.306	0.014		
	(0.062)	(0.759)	(0.064)		
sizexconc		0.021			
		(0.059)			
msxconc			-1.061***		
			(0.286)		
concentration				-0.163	
				(0.133)	
concentration_squ					-0.184*
					(0.111)
Constant	2.352***	2.205**	2.406***	2.480***	2.455***
	(0.708)	(0.925)	(0.720)	(0.741)	(0.731)
Observations	43,663	43,663	43,663	43,663	43,663
R-squared	0.722	0.722	0.722	0.722	0.722
Number of isin	3,270	3,270	3,270	3,270	3,270

Robust standard errors in parentheses
*** p<0.01, ** p<0.05, * p<0.1

Table 10.18: Panel regression results using the post-2000 Orbis sample consisting only of Northern European firms.

VARIABLES	(1) Model 1	(2) Model 2	(3) Model 3	(4) Model 4	(5) Model 5
market_share	0.004	0.004	-0.239**	0.020	0.008
	(0.044)	(0.048)	(0.116)	(0.048)	(0.053)
log_age	0.019	0.019	0.020	0.019	0.019
	(0.015)	(0.015)	(0.015)	(0.015)	(0.015)
log_assets	-0.018	-0.018	-0.015	-0.018*	-0.018*
	(0.011)	(0.012)	(0.011)	(0.011)	(0.011)
size_inv	-1.237***	-1.237***	-1.200***	-1.239***	-1.240***
	(0.396)	(0.401)	(0.382)	(0.389)	(0.386)
RnD	-0.791***	-0.791***	-0.789***	-0.792***	-0.793***
	(0.060)	(0.061)	(0.060)	(0.060)	(0.060)
CapEx	0.190**	0.190**	0.184**	0.195**	0.198**
	(0.077)	(0.076)	(0.076)	(0.077)	(0.077)
log_concentration	-0.047**	-0.047	-0.054***		
	(0.020)	(0.064)	(0.021)		
sizexconc		-0.000			
		(0.010)			
msxconc			0.153***		
			(0.058)		
concentration				-0.025**	
				(0.011)	
concentration_squ					-0.002
					(0.002)
Constant	0.184**	0.184**	0.177**	0.210**	0.177**
	(0.081)	(0.084)	(0.079)	(0.082)	(0.082)
Observations	3,072	3,072	3,072	3,072	3,072
R-squared	0.340	0.340	0.343	0.339	0.337
Number of isin	408	408	408	408	408

Robust standard errors in parentheses
*** p<0.01, ** p<0.05, * p<0.1

Table 10.19: Panel regression results using the post-2000 Compustat sample consisting only of Southern European firms.

VARIABLES	(1) Model 1	(2) Model 2	(3) Model 3	(4) Model 4	(5) Model 5
market_share	0.085***	0.084***	0.114***	0.088***	0.088***
	(0.025)	(0.025)	(0.031)	(0.025)	(0.025)
log_age	0.020**	0.019**	0.020**	0.020**	0.020**
	(0.009)	(0.009)	(0.009)	(0.009)	(0.009)
log_assets	0.011*	0.014*	0.012*	0.011*	0.011*
	(0.006)	(0.008)	(0.006)	(0.006)	(0.006)
size_inv	-85.793	-85.789	-85.717	-85.815	-85.800
	(57.812)	(57.823)	(57.847)	(57.805)	(57.811)
log_concentration	-0.028*	-0.088	-0.035**		
	(0.016)	(0.096)	(0.017)		
sizexconc		0.005			
		(0.007)			
msxconc			0.089		
			(0.067)		
concentration				-0.053*	
				(0.032)	
concentration_squ					-0.039
					(0.027)
Constant	-0.191**	-0.224**	-0.197**	-0.143	-0.161*
	(0.083)	(0.108)	(0.084)	(0.095)	(0.091)
Observations	19,105	19,105	19,105	19,105	19,105
R-squared	0.039	0.039	0.039	0.039	0.039
Number of isin	1,440	1,440	1,440	1,440	1,440

Robust standard errors in parentheses
*** p<0.01, ** p<0.05, * p<0.1

Table 10.10: Panel regression results using the post-2000 Orbis sample consisting only of Southern European firms.

VARIABLES	(1) Model 1	(2) Model 2	(3) Model 3	(4) Model 4	(5) Model 5
market_share	0.419*	0.310	0.977**	0.560**	0.577**
	(0.242)	(0.225)	(0.442)	(0.248)	(0.242)
log_age	-0.018	-0.018	-0.021	-0.021	-0.020
	(0.023)	(0.024)	(0.023)	(0.023)	(0.023)
log_assets	-0.023	-0.025*	-0.030**	-0.023*	-0.021*
	(0.015)	(0.014)	(0.015)	(0.014)	(0.013)
size_inv	-0.316*	-0.325**	-0.328**	-0.317*	-0.318*
	(0.160)	(0.155)	(0.163)	(0.162)	(0.164)
RnD	-0.311	-0.314	-0.295	-0.296	-0.284
	(0.270)	(0.261)	(0.275)	(0.264)	(0.268)
CapEx	0.111	0.134	0.112	0.124	0.125
	(0.097)	(0.099)	(0.099)	(0.096)	(0.096)
log_concentration	-0.062	-0.154	-0.042		
	(0.044)	(0.114)	(0.051)		
sizexconc		0.020			
		(0.018)			
msxconc			-0.431		
			(0.293)		
concentration				-0.054**	
				(0.025)	
concentration_squ					-0.012**
					(0.005)
Constant	0.249**	0.250***	0.264**	0.310***	0.256***
	(0.096)	(0.089)	(0.103)	(0.104)	(0.091)
Observations	506	506	506	506	506
R-squared	0.175	0.182	0.182	0.180	0.180
Number of isin	122	122	122	122	122

Robust standard errors in parentheses
*** p<0.01, ** p<0.05, * p<0.1

Table 10.11: Panel regression results using the post-2000 Compustat sample consisting only of Eastern European firms.

VARIABLES	(1) Model 1	(2) Model 2	(3) Model 3	(4) Model 4	(5) Model 5
market_share	0.346**	0.270**	0.530**	0.352**	0.348**
	(0.159)	(0.120)	(0.265)	(0.159)	(0.153)
log_age	0.036	0.008	0.033	0.036	0.036
	(0.080)	(0.068)	(0.080)	(0.081)	(0.082)
log_assets	0.041	0.150*	0.043	0.041	0.041
	(0.065)	(0.089)	(0.065)	(0.065)	(0.065)
size_inv	35.227***	35.357***	35.248***	35.219***	35.213***
	(7.080)	(7.054)	(7.072)	(7.079)	(7.077)
log_concentration	-0.296	-1.963	-0.338		
	(0.211)	(1.749)	(0.231)		
sizexconc		0.165			
		(0.154)			
msxconc			0.504		
			(0.474)		
concentration				-0.499	
				(0.330)	
concentration_squ					-0.371
					(0.238)
Constant	-0.730*	-1.720*	-0.756*	-0.264	-0.418
	(0.397)	(0.985)	(0.406)	(0.541)	(0.471)
Observations	12,202	12,202	12,202	12,202	12,202
R-squared	0.189	0.193	0.189	0.189	0.189
Number of isin	1,286	1,286	1,286	1,286	1,286

Robust standard errors in parentheses
*** p<0.01, ** p<0.05, * p<0.1

Table 10.12: Panel regression results using the post-2000 Orbis sample consisting only of Eastern European firms.