STOCKHOLM SCHOOL OF ECONOMICS
Department of Economics
659 Degree project in economics
Spring 2020

Intangible Inputs and Swedish Research Productivity

Hannes Månsson Höglund (24079) and Wilmer Niva Printz (24230)

Abstract: In this paper, we set out to explore research productivity in Sweden and the role of intangible inputs in production. We examine the primitives of an already existing model using financial account data from Swedish firms to evaluate its applicability in the context of Swedish research productivity. This includes measuring the strengths of the relationships between fixed-cost share and the use of intangibles, markups, R&D, and sales growth, respectively. On the basis of this, Swedish firm concentration and research productivity is measured to analyze whether the predicted outcomes of the model can be observed, given the development of the use of intangibles. Our findings include a significant fit of the model, which, given an observed relative increase in investments in intangibles, would predict increasing concentration and decreasing research productivity. While we observe no significant increase in concentration, we find a decreasing research productivity in line with our expectations.

Keywords: Research Productivity, Intangible Inputs, Swedish Paradox, Concentration

JEL: O31, O34

Supervisor:	Federica Romei
Date submitted:	May 14, 2020
Date examined:	May 25, 2020
Discussants:	Sigrid Holmgren and Johan Callermo
Examiner:	Johanna Wallenius

Table of Contents

1. Introduction

An ever-increasing stock of ideas is in economic theory widely considered to be the foundation for sustained economic growth. For example, the innovation of the Spinning Jenny accelerated production during the industrial revolution, the emergence of efficient institutions has through history improved the performance of economies around the world, and new agricultural techniques increased the amount of food that could be produced given the amount of land and workers during the agricultural revolution. Without a constant flow of new ideas like these, economic growth will eventually come to a halt.

The rate at which new ideas are developed has been closely studied in economic literature, and recent concerns have been raised that more and more research efforts are needed to sustain a constant growth of ideas, raising questions like “Are ideas getting harder to find?” (Bloom et al., 2020). In this paper, we will explore one possible explanation for why it might be increasingly difficult to create idea growth, namely an increasing use of intangible inputs in production. De Ridder (2019) explores this explanation, building a model where he introduces a group of high-intangible firms that eventually grow big and concentrate the market, which in turn leads to decreasing research productivity in the long run. While De Ridder tests his model on French and U.S. data, we aim to try his explanation on Swedish research productivity.

In Section 2 we will first provide a short review on economic theories on long-run economic growth and the role of ideas. This will be followed up with a background on the growth of ideas and the use of intangibles in Sweden, and a motivation for why Sweden is of interest with regard to this topic. In Section 3 we will review already existing literature on the topic, with the article from De Ridder (2019) serving as a starting point. On the basis of the background and the literature review we will more specifically define our research focus in Section 4. Section 5 will go through the research method, Section 6 will introduce the data used to perform the research, and Section 7 will present the results. A discussion of the results will be held in Section 8, before reaching a final conclusion.

2. Background

2.1. Ideas and Long-Run Growth

At the heart of most long-run growth models is a function explaining how output is produced within the economy. Often, the Cobb-Douglas production function is used where labor (L) and capital (K) are inputs with a constant return to scale together (output doubles if both capital and labor double) and the stock of ideas (A) measures how efficiently labor and capital are combined to create output (Y). Including the stock of ideas in the equation enables the model to predict two countries with the same level of capital and labor to still have different GDP, due to productivity differences.

$$Y = A \cdot K^{\alpha} \cdot L^{1-\alpha}$$

Included in the stock of ideas goes, among other things, the educational level of the labor force, the efficiency of institutions, and technology, and it is often referred to as TFP (total factor productivity). Most famously, assuming TFP and the labor force to be constant, the Solow growth model predicts that economic growth will eventually stop at a steady state due to the diminishing returns to capital, the level of which depending on the level of TFP, the amount of labor, the investment rate, and the depreciation rate. Since sustained economic growth does not occur endogenously within the model, Solow's growth model and related models are characterized as exogenous growth models.

Economists like Paul Romer later developed endogenous growth models where sustained growth can be explained within the model through an idea production function. The labor force that was previously assumed to fully engage in producing output in the output production function is now divided into one fraction of the labor force producing output, and the other fraction producing ideas (increasing TFP) through research and innovation. The fundamental characteristic of ideas is that they are non-rivalrous, meaning that the use of an idea by one person does not hinder another person to use the same idea simultaneously. On the contrary, objects (such as labor and capital) are rivalrous and cannot be used at multiple places at once. The idea production function states that the stock of ideas (TFP) will grow at a growth rate (g_A) that is dependent on the number of researchers (L_A) and research productivity (z).

$$g_A = z \cdot L_A$$

Due to the ever-growing stock of ideas, the economy will not reach a steady state where growth will come to a halt, but instead experience sustained economic growth. Hence, understanding these terms is fundamental to understanding how long-run growth will proceed. Research productivity is often assumed to be constant, resulting in a constant growth rate of ideas as long as the number of researchers stays constant, in turn leading to a constant growth rate of GDP. However, the assumption of constant research productivity has been questioned in economic literature, which will be further explored in the literature review.

2.2. TFP Growth in Sweden

Sweden is a country of interest with regard to the topic of research productivity as it constitutes a highly developed country that has experienced relatively slow TFP growth since the 1970s, despite extensive research efforts, suggesting a low research productivity. In fact, Sweden has among the highest ratios of R&D spending to GDP in the world, according to gross investment data from the UNESCO Institute for Statistics database. This phenomenon is common among developed countries but has been particularly discussed with regard to Sweden under the name the "Swedish paradox". Ejermo and Kander (2006) attempt to summarize existing literature on the topic, while also searching for the explanation of the weak returns on research and development efforts. The main reasons found are that Sweden has experienced problems in its entrepreneurial climate, where R&D has been concentrated among a few multinational firms, and that high- and medium-tech innovation is at a low level relative to R&D expenditure.

Figure 1 illustrates the effective number of researchers and the growth rate of TFP in Sweden from 1995 and 1985 respectively. By utilizing the theoretical idea production function, this represents L_A and g_A. As seen in the figure, the yearly growth rate of TFP has been relatively stable at low rates close to zero, while the research efforts have been increasing.

Hence, by the logic of the idea production function, this suggest research productivity (z) has been decreasing in recent years.

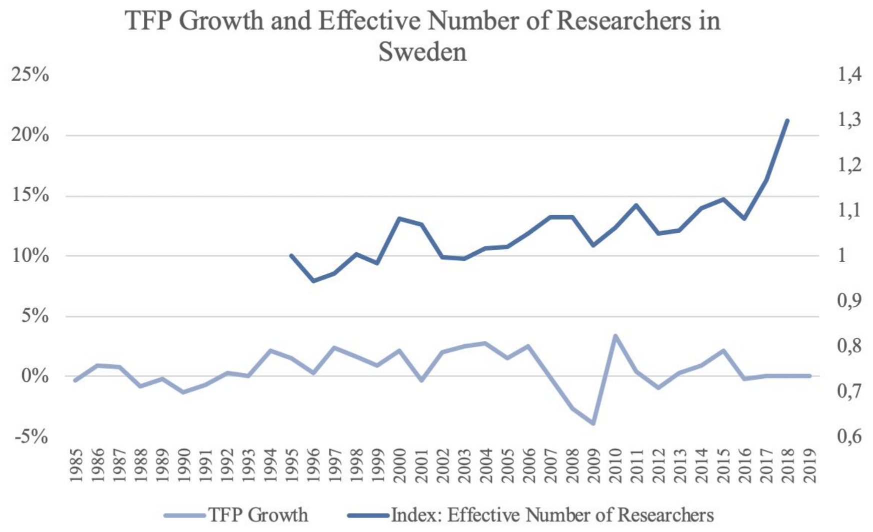

Figure 1: Total factor productivity (TFP) growth (left axis) is retrieved from the OECD Productivity Database, based on National Accounts data. Effective number of researchers (right axis) follow Bloom et al. (2020) and is measured as gross investments in intellectual property products from the National Accounts (Statistics Sweden), deflated by a measure of average wage for persons with two years or more of tertiary education (Statistics Sweden, Appendix A). Index 1995 = 1.

Since both the effective number of researchers and the growth of TFP where relatively stable from around 1995 to around 2010, this speaks for a relatively stable research productivity during this period. After 2010 however, TFP growth has been at slightly lower levels, while the effective number of researchers has been increasing more rapidly, suggesting a decrease in research productivity especially during this later period.

2.3. The Use of Intangibles in Sweden

Building on the concept of non-rivalrousness, De Ridder (2019) and other researchers have explored the hypothetical relationship between an increasing use of intangible inputs in production and a decreasing research productivity. In Sweden, investments in R&D and software and databases together make up the majority of investments in intangibles. Figure 2 shows how investments in these have developed compared to investments in tangible fixed assets in Sweden since 1981. A clear relative increase of investments in intangibles is observed, indicating that this might serve as part of the explanation for the decrease in research productivity observed in the previous section. Exploring the indicated relative increase of investments in intangibles as an explanation for the development of Swedish research productivity will thus be the focus of this paper.

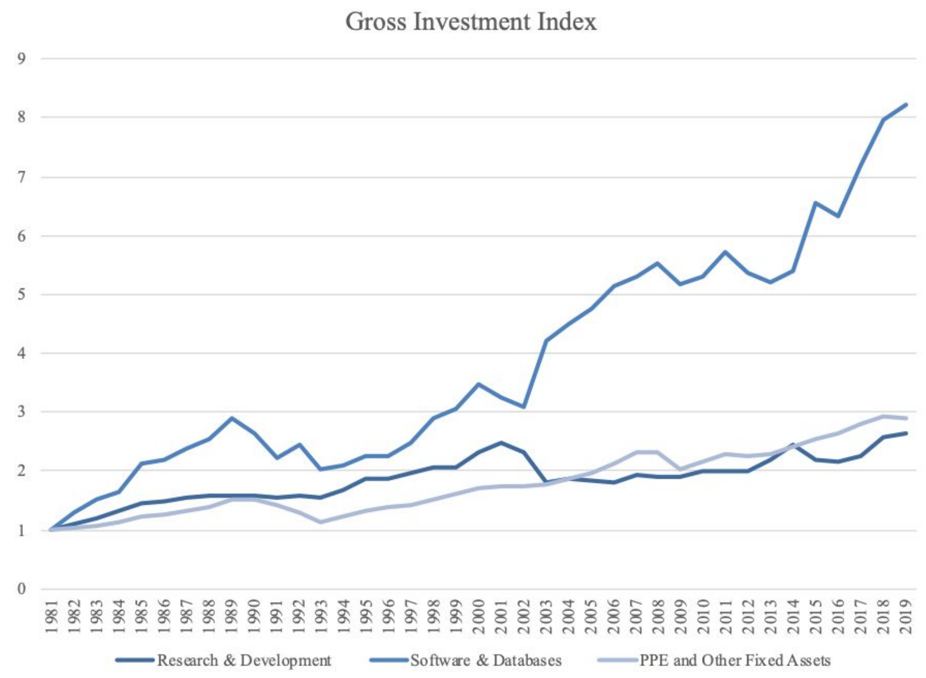

Figure 2: Gross fixed investments from the Swedish National Accounts (Statistics Sweden). Index 1981 = 1.

3. Review of Current Literature

In this section relevant literature on the topic of research productivity will be summarized, with the argument from De Ridder (2019) serving as a starting point both for this literature review and the scope of this paper in general.

3.1 Intangibles and Research Productivity

De Ridder (2019) leads by referring to data showing that TFP growth has in recent times been declining in the US and much of Europe after being above average during the 1990s. The decline has occurred even though research efforts have been increasing, suggesting that research productivity has been declining. This trend is supported by Bloom et al. (2020), who found declining research productivity in the transistor-, agricultural-, and medical sectors in the U.S. These observations are in line with the development in Sweden discussed in the previous section. De Ridder further sets up a model based on four observed positive relationships in firm data between the fraction of total costs that are fixed and the use of intangible inputs, markups, R&D efforts and sales growth. His argument for how an increased use of intangibles as inputs effects research productivity follows a number of steps based on the observed correlations. Firstly, as previously mentioned, he proposes that investments in intangibles in production raise fixed costs and lowers marginal costs due to the feature of non-rivalry in intangibles. Secondly, due to lower marginal costs, high-intangible firms will thus incur higher markups, grow faster, and gain a more dominant market position and invest more in R&D. To exemplify his modeled primitives, he introduces a group of high-intangible firms with lower marginal costs in the model. The model predicts that the introduction of these firms will initially increase research productivity, due to the efficiency of the newly introduced firms. In the long run however, R&D will become more concentrated among these firms, leading to a lower overall R&D effectiveness. The fall in productivity is due to assumed diminishing returns to R&D within firms, supported by Akcigit and Kerr (2018). Reasons for why the efficiency of R&D for growth decreases as it gets concentrated among fewer but larger firms will be further discussed with other literature below. To the point of intangibles, Fernald (2014) analyzes TFP growth during the decades around the recent Great Recession and finds a period of high TFP growth during a decade around the turn of the century, followed by a slowdown. In line with the findings of De Ridder (2019), Fernald found that the temporary boost of high TFP growth was primarily derived from IT-producing or intense-IT-using sectors.

3.2. Intangible Inputs and Concentration

A central prediction of the model of De Ridder (2019) is that the introduction of a group of intangible-intense firms eventually leads to market activity becoming increasingly concentrated among these firms. Central for this prediction is the assumption of heterogeneity among firms in their adoption of intangibles. Korinek and Ng (2019) support this view as they build a model where certain successful firms will create digital innovations for a fixed cost. Because of the non-rival nature of these digital innovations, the production can be scaled up at a lower marginal cost, while the excludability of these innovations (through patenting) leads to these firms becoming more dominant, further concentrating the economic activity. Criscuolo et al. (2019) also highlight the scalability of intangible capital as inputs in production. By analyzing data from nine OECD-countries (including Sweden), they find strong links between intangibles as inputs in production and concentration (where concentration is measured in terms of the fraction of sales that is concentrated among the eight largest groups). Similar links, between the use of IT systems and industry concentration, are found by Bessen (2017). He even provides some support for the causality of this relationship and shows that it is especially the performance of the top firms within each industry that improves the most with increasing use of IT systems, leading to higher concentration.

3.3 Concentration and Research Productivity

Akcigit and Kerr (2018) also build a model incorporating heterogeneity among firms in terms of firm size and investment choices. They divide investments into internal and external, where internal innovations are those that develop or improve already existing product lines within the firm and external innovations are developments of new products. In the article they provide a model showing that as firms grow, they engage relatively more in internal innovation rather than external, and that smaller firms (engaging relatively more in external innovation) are relatively more innovative. Dinopoulos and Syropoulos (2007) poses an argument along the same lines, setting up a model where incumbent firms engage in so called rent-protecting activities to protect their market power. This will lead to more resources being spent on making innovations more difficult to imitate, rather than on the activity of innovation itself, leading to research productivity falling as innovation gets concentrated among a few big firms. Similar findings are found by Torrisi et al. (2016). In their study of survey answers regarding patenting activity, the fraction of patents that were found not to be used by the company increased with firm size (23.5, 23.0 and 43.8 percent respectively among small, medium and large firms). In their paper, patents not used are in turn categorized into “Strategic non-use” referring to patents with a strategic motive of blocking competitors, similar to the rent-protecting activities referred to by Dinopoulos and Syropoulos (2007), and

“Sleeping patents” that are neither used nor have a strategic motive. Together, these studies contribute with a number of reasons for why research productivity decreases as R&D becomes more concentrated among a few big firms.

3.4 Research Productivity in Sweden

In their paper, Ejermo and Kander (2011) study the development of research productivity in Sweden between 1985 and 1998. They define research productivity as granted patents over R&D spending, which is a measure we will use further on in this paper. Similar to the findings of previously mentioned articles, they found research productivity to increase significantly during the period, mainly driven by large increases in research productivity in firms engaged in low- and medium-tech manufacturing, transportation, and chemicals. However, the subsequent slowdown seen in the other articles was not observed here, possibly as a result of the earlier time frame of this study. As mentioned in the background section, Ejermo and Kander (2006) found the generally low Swedish research productivity to partly be due to a strong concentration of R&D to a few big firms, which corresponds well with much of the previously discussed literature.

To sum up, research productivity seems to be dependent to some degree on the use of intangibles, which tends to get concentrated among a few big intangible-intense firms. In turn, R&D becoming more concentrated decreases research productivity due to larger firms engaging relatively more in so called rent-protecting activities (such as strategic patenting) rather than innovation, and in internal innovations rather than external. In Sweden there is a long history of slow TFP growth relative to its high R&D numbers, suggesting a weak research productivity, even if it seems to have increased between 1985 and 1998.

4. Research Focus Specification

As seen in the literature review, there is plenty of research done using micro-data to measure research productivity, most of them contributing to a story of an IT-caused boom in research productivity just before- or around- the turn of the millennium, followed by a seemingly lasting decline in research productivity. The model from De Ridder (2019) suggests the boom and the subsequent decline in research productivity comes from an increased use of intangibles. Since only looking at the period between 1985 and 1998, Ejermo and Kander (2011) might only have managed to capture the boom, and not the eventual subsequent decline. As seen in the background section, there are indications that Swedish research productivity has been declining in recent years, at the same time as investments in intangibles has increased relative to tangible assets, insinuating that De Ridder's predictions might hold true even in Sweden. Hence, using De Ridder's theory to update the story of Swedish research productivity will be the focus of this paper.

We will examine if the relationships that serve as building blocks in the model of De Ridder (2019) hold true using Swedish data. Given a sufficiently significant fit of the relationships, we will then analyze whether the observed relative increase in investments in intangibles has led to the outcomes that the model would predict, namely a higher firm concentration and a decrease in research productivity in the long run. In short: How has research productivity developed in Sweden, and what role does intangibles play?

5. Research Method

5.1 Testing the Fit of the Model

As previously mentioned, the first step of the analysis will be to examine the fit of the relationships that serves as building blocks in De Ridder´s (2019) model, using data on a sample of Swedish firms which will be introduced in the data section. To do this, the fixed-cost share measure has to be calculated since it cannot be directly observed in the data.

5.1.1 Calculating Fixed-Cost Share

Following the equations used by De Ridder (2019), we define fixed cost f to sales revenue $p \cdot y$ as the difference between marginal cost markup and the profit rate, which is operating profit π divided by revenue. Next is the accounting definition of the profit rate for firm i at time t.

$$\frac{\pi_{it}}{p_{it} \cdot y_{it}} = \frac{(p_{it} - mc_{it}) \cdot y_{it}}{p_{it} \cdot y_{it}} - \frac{\tilde{f}_{it}}{p_{it} \cdot y_{it}}$$

Fixed costs to revenues are isolated and markup μ is defined as the ratio of prices to marginal costs. This leaves a definition of fixed costs to revenue.

$$\frac{\tilde{f}_{it}}{p_{it} \cdot y_{it}} = \left(1 - \frac{1}{\mu_{it}}\right) - \frac{\pi_{it}}{p_{it} \cdot y_{it}}$$

Multiplying by revenue and dividing by total operating costs tc results in an equation that defines fixed costs as a share of total operating costs.

$$\frac{\tilde{f}_{it}}{tc_{it}} = \frac{p_{it} \cdot y_{it}\left(1 - \frac{1}{\mu_{it}}\right) - \pi_{it}}{tc_{it}}$$

Sales revenue, operating profits, and total operating costs are observed directly in the sample. We let cost of goods sold estimate the marginal cost times the number of products sold, and sales revenue estimate the price times the number of products sold. Hence, markup is computed as sales over cost of goods sold for firms with income statement classified by function. For firms with income statement classified by nature with no reporting of cost of goods sold, the wage bill and costs for materials and other intermediate inputs in production are used instead.

5.1.2 Testing the Relationships

When the fixed-cost share has been calculated, we will start out by testing the relationship between the use of intangible inputs in production and the fixed-cost share of firms, in which the fixed-cost share is the dependent variable. Financial accounts lack explicit data on expenditures incurred from the use of intangibles. However, the stock of intangible assets maintained by each firm can be obtained from the balance sheet. In contrast to De Ridder´s (2019) measures of software expenses and technology adoption, based on detailed survey data obtained for part of his sample, we will measure the use of intangibles in terms of intangible asset intensity, defined as intangible assets in relation to sales for our complete sample. The background for this measure if further elaborated on in Appendix B.

If the relationship between the fixed-cost share and our measure of intangible asset intensity is found to be significant we can conclude that intangible-intensive firms also tend to have a higher fixed-cost shares. The intuition is that intangible assets incur a certain fixed cost to develop or purchase, while they lead to reductions in marginal costs due to the scalability of intangible inputs in production, as discussed in the literature review.

For the rest of the relationships, the fixed-cost share is the explanatory variable. Firstly, we will examine if firms with higher fixed-cost share have higher markups (due to the lower marginal cost of using intangible inputs). Secondly, we will examine if these firms also engage in more R&D and lastly if they experience higher sales growth. If all relationships hold true, the implications of De Ridder´s (2019) model seem applicable to the context of Swedish research productivity, and we can start comparing the predicted outcomes of the model with reality.

5.2 Concentration

Given what has been observed in terms of the fit of the model and the observed relative increase in investments in intangibles, we will analyze if Sweden seems to have experienced the expected outcomes, firstly in terms of concentration. De Ridder (2019) measures concentration by looking at the distribution of firms according to how many products they produce. He compares the balanced growth paths before and after the introduction of the high-intangible firms and finds a decrease in the share of firms that produce one or two products, while the share of firms that produce three or more products has gone up. Hence, at the new balanced growth path, the economy consists of fewer small firms and activity is more concentrated around the larger firms.

Since we will analyze the predicted outcomes of the model in a reality where changes in intangibles are gradual, as opposed to in a model where high-intangible firms are introduced as a shock, the Herfindahl-Hirschman Index (HHI) will be used to provide us with a measure of concentration on an annual basis ranging from 0.0 to 1.0, instead of the before-and-after comparison made by De Ridder (2019). HHI is calculated by summarizing the squares of all market shares, leading to 1.0 illustrating a market where one firm has 100 percent market share and 0.0 illustrating a market of infinitely many firms exhibiting infinitely small market shares. HHI will be calculated on our sample as a whole and on a number of industries within the sample that will be defined in the data section, and it serves as our primary measure for concentration.

To complement the HHI measure, another standard measure of concentration will be used, namely calculating the concentration ratio. Computing the concentration ratio consists of computing the combined market share of the top firms, most commonly the largest four or eight. This will be calculated on the sample as a whole. Due to the between-industry differences in number of firms in the sample, we will also measure concentration on industry level by computing the combined market share of the five percent largest firms in the sample in each industry, as well as computing the combined market share of the one, two, and five percent largest firms in the sample as a whole. The results from the complementary measures of concentration will be found in Appendix D.

5.3 Research Productivity

Finally, research productivity will be measured as granted patents relative to R&D expenditure deflated by an index of high-skilled labor wages in the spirit of Ejermo and Kander (2011). What R&D expenditure at what time that corresponds to a certain patent filing is normally not observable from outside of a firm or other entity engaged in research. When testing several models to account for the research process, Wang and Hagedoorn (2014) find a one-year lag between R&D and related patent applications to be consistently significant. Further analysis by Dang and Motohashi (2015) confirm that a lag of one year is more significant than either same-period relationships or using longer lag periods. They also estimate the time from application to grant to be 3.87 years on average. Similarly, Bloom et al. (2020) allow for five years from R&D to granted patents. Consequently, we will base our measure of research productivity on R&D expenditure incurred the current year, and on granted patent applications that were filed during the subsequent year.

The choice of which deflator to use for R&D differs in previous research. Naturally, R&D expenditure includes spending on both capital and labor. Hence, to this extent, the use of a

GDP deflator by Hall et al. (2007) can be motivated. More commonly however, R&D is treated as mostly driven by expenditures on wages to the labor involved. Ejermo and Kander (2011) use a wage index for graduate engineers as a deflator, and Bloom et al. (2020) use mean personal income for males with a bachelor's degree or more. We will follow the wage approach and deflate R&D expenditure by an index of mean wages of men and women with two years of tertiary education or more, constructed by data from Statistics Sweden's database Lönestrukturstatistik. As an alternative we will consider a similar wage index based on even higher-skilled labor wages, using doctoral educations. Although the wage levels differ, the indexed series closely mirror each other and thus provide little explanatory difference. The deflator indices are compared in Appendix A.

6. Data Material

6.1 Financial Account Data

To carry out our research, we have retrieved complete and standardized financial statement data on Swedish firms. This data has been collected from Bureau van Dijk's database Amadeus which has available records of public and private European firms from 2010 until 2019. Due to comparatively low data availability for 2019, we have excluded this year as to not give it an unfair lower weight in the analysis. We have excluded data from firms with limited financial data and firms that have missing data for five or more years due to entry, exit or limited data availability. As we set out to estimate research productivity using research and development efforts, we restrict the sample to firms that have reported R&D expenses at least one year during the period. We furthermore drop firm accounts where accounting data has been misrecorded in the data set by excluding years that have negative sales, costs or assets. This leaves a sample of 770 firms, each with balance sheet and income statement data with detailed revenues and costs. The data set also contains additional information on industry classification codes according to NACE Rev. 2 (statistical classification of economic activities in the European community) that enables industry aggregation, and a firm identifier (national trade register number) that allows us to follow firms over time and to match the accounting data with firm-level data from other sources.

Using the NACE Rev. 2 classifications, we group the firms into sectors to make computing of industry concentration and analysis of research productivity on the sector level possible, as well as for making comparisons to other research. With NACE section in parentheses, we find that 670 of the sample firms classify in Manufacturing (C), Wholesale and Retail Trade (G), Information and Communication (J), or Professional, Scientific and Technical Activities and Other Service Activities (M/N/S). The remaining firms are scattered across sectors that do not make for natural aggregation, among them Electricity, Gas and Water (D/E), Construction (F), Real Estate Activities (L) and Financial and Insurance Activities (K).

6.2 Patent Data

Quantifying our research productivity measure requires data on patents in conjunction with resources used in research and development. One possible method is to base the productivity measure on patent applications. We think this is questionable when it comes to measuring the generation of ideas, as there are little criteria to be met in order to file an application. Thus, a patent filing in itself cannot accurately predict whether the patent application is actually granted, withdrawn or denied, and therefore constitutes a highly uncertain measure of actual technological progress. A granted patent constitutes a recognition that a certain technological advance has been made, as a patent can only be granted if there is no “prior art”, that is, no evidence that anyone, anywhere, at any time has described or used a similar technology. Application document data belonging to granted patents is retrieved from the European Patent Office’s PATSTAT database. This database includes legal statuses and documentation on over 100 million patent records from 90 patent issuing authorities. The chosen years correspond to our financial accounts data, starting in 2010 and ending in the last data available, which is halfway through 2019.

In analyzing Swedish research productivity, the origin of patents must be established. But whether to consider the applicant or the inventors in this regard is not an uncomplicated question. Ejermo and Kander (2011) base their analysis on Swedish inventions defined as patents with at least one inventor listed with a Swedish address. In order to compile their data set, computerized matching between inventors' names and addresses and their respective firms was provided by Statistics Sweden. However, referring to previous work by Ejermo and Gabrielsson (2007), they note that no divergence was found when comparing the development of patents with non-Swedish inventors and a Swedish applicant, and patents with one or more Swedish inventors but a non-Swedish applicant. Thus, we do not expect data to render widely different conclusions if origin is established based on the applicant or inventors. EPO assert in their PATSTAT database documentation that the country code for inventors is missing for about 50 percent of the applications. Hence, we find it more reliable to consider the origin of the applicant in classifying patents as Swedish, following also the standard procedure used in patent data reporting from the World Intellectual Property Organization (WIPO), a UN specialized agency. Hence, when retrieving data, we use the country stated in the correspondence address belonging to the applicant as provided in the documentation for the patent application. In the case of firms active in more than one country with patents filed from the headquarters, some overestimation of the patent activity in the country of the headquarters could be present using the chosen approach. Then again, not all inventors that list an address list their personal residential address in the application filings but instead use the company address, which could underestimate the innovation taking place in the country instead if using this approach. We have not engaged in defining the most accurate measure in this paper.

In their nature, patents are connected to technology. The data records kept by EPO and other patent organizations are thus primarily structured by a logic that enables mapping of technological relationships and trends. This is made possible as each filed patent is assigned one or more International Patent Classification (IPC) codes. IPC is an internationally recognized system administered by WIPO that provides a hierarchical structure needed when searching for prior art. To avoid double-counting when assigned multiple IPC codes, patents are individually divided into fractions based on their technological relevance in order to sum to the actual number of patents when aggregating data across technologies. IPC makes analysis of innovation between and within technological fields possible, but it says little about firms or industries, as one feature of technology can be utilized throughout many different industries simultaneously. Hence, a reclassification from the standard is needed to combine the patent data with firm- and industry accounting measures.

For this reason, we obtain comprehensive matching files from the PatLink project, coordinated by Swedish House of Finance (SHoF). The files connect individual patent application identifiers with national trade register numbers of corresponding Swedish firm applicants. The data set covers Swedish firms from 1990 to 2018, thus including the whole sample period of our financial account data. Through the epodoc application number that is uniquely assigned each patent application, we match the patents in the PatLink data set with the patent documentation data retrieved from PATSTAT. This leaves a data set with documentation on granted patents, each with the trade register number of the applicant.

In patent data, one application with several cited inventors holds equally many data records. By excluding non-distinct epodoc identifiers we avoid the risk of counting one invention multiple times. We exclude utility models and design patents from the data by accepting patent of invention as the only type of intellectual property right applied for in the patent documentation. Utility models and design patents are forms of applications that are inherently different from standard patents as they require less stringent conditions to be granted and offer less protection when granted. Utility models protect what is referred to as minor inventions, where the required inventive step from previous technology is lower or even absent, and the novelty requirement may be examined on a local level only in some countries. Design patents cover ornamental designs for objects with practical utility, but the novelty and non-obviousness requirements are not based on the utility itself. Hence, we regard patent of invention to be the only relevant measure.

We assign a reference year to each granted patent that is the year its initial application was filed. If a patent is published and or granted at a later date, the application date provides the best estimate for when the innovation or invention was actually created, as it is the closest in time to the preceding research and development effort of which the patented technology is a product. From the remaining selection of patent data, we count the number of patents yearly for each applicant firm.

6.3 Combined Data Set

Using the trade register number as a link, we assign the yearly number of granted patents to each firm in the financial account data set constructed from the Amadeus database. This leaves us with a combined panel data set containing all the data required to perform our research. Descriptive statistics are provided in Table 1.

Descriptive Statistics

	Mean	Std. Dev.	Median	10th Pct.	90th Pct.
Sales	3,984,651	19,002,124	269,433	29,882	4,942,801
Cost of goods sold	3,179,527	14,801,412	214,818	21,484	4,210,673
Wage bill	139,103	1,360,135	15,952	1,287	87,143
Materials and intermediate inputs	79,189	120,760	34,316	1,196	217,902
Other operating expenses	889,594	4,342,694	66,734	8,752	905,816
Research and development	328,627	2,184,841	13,989	1,206	207,287
Intangible assets	891,445	5,578,258	920	0	443,946
Number of employees	1,291	6,768	90	9	1,150
Granted patents	4	65	0	0	2

Nominal amounts in thousands of SEK, deflated with the GDP deflator. R&D is deflated with the high-skilled labor deflator. Wage bill and material is reported for firms not reporting cost of goods sold.

Table 1: Descriptives of the panel data set following 770 Swedish firms 2010-2018, compiled with data from Amadeus and PATSTAT. The GDP deflator is based on World Bank National Accounts Data and OECD National Accounts data files. The R&D deflator is based on Statistics Sweden's Lönestrukturstatistik, see Appendix A. Variables are further described in Appendix C.

In preparation for the empirical analysis, we compute the fixed-cost share using the definition in Section 5. Figure 3 illustrates the fraction of estimated fixed costs to total operating costs over time.

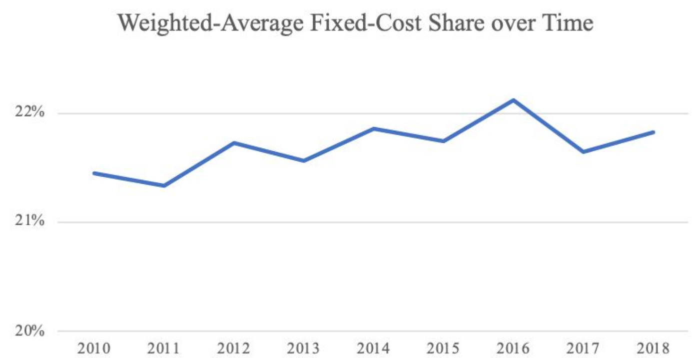

Figure 3: Fixed costs as a percentage of total costs, yearly sales-weighted average across all firms in the sample. Data from Amadeus.

The weighted-average fixed-cost ratio deviates very little from 2010 to 2018. It is consistent with the level among U.S. firms and higher than the ratio for French firms during the same period, as estimated by De Ridder (2019). In Figure 4, the fixed-cost share across sectors is depicted.

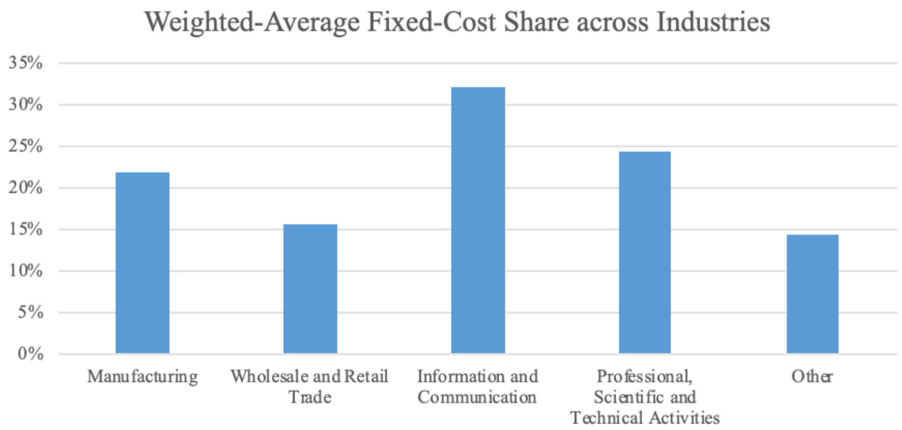

Figure 4: Fixed cost as a percentage of total costs, sales weighted average within each industry 2010-2018. Data from Amadeus.

The Wholesale and Retail Trade sector where variable costs are high exhibit a comparatively low fixed-cost share, while fixed costs in the Information and Communication sector are more important. In comparison to De Ridder's (2019) estimates, Manufacturing and Wholesale and Retail Trade are in line with the US data, while Information and Communication more closely resembles data on French firms.

7. Empirical Analysis

7.1 The Fit of the Model

In this section we apply the regression equations used by De Ridder (2019), either directly or with some variation, on our panel data set. The theoretical framework tested builds on several primitives, the first one being that high-intangible firms have higher fixed-cost shares, due to the scalability feature of non-rival intangible inputs in production changing the cost structure. We therefore set up the fixed-cost share f/tc, as measured in Section 5.1 for firm i in year t as the dependent variable with intangible assets to sales $I/(p \cdot y)$ as the explanatory variable.

α and ψ are controls for firm- and year-fixed effects, and $g(\cdot)$ is a third-degree polynomial in log real sales controlling for size effect.

$$\frac{\tilde{f}_{it}}{tc_{it}} = \alpha_i + \psi_t + \gamma \cdot \frac{I_{it}}{p_{it} \cdot y_{it}} + \beta' g(p_{it} \cdot y_{it}) + \varepsilon_{it}$$

Table 2 presents the results. We find a significant positive relationship between the intangible asset intensity and fixed-cost share. However, with firm-fixed effects and firm-clustered standard errors, the coefficients are not as robust. Column IV suggest that an increase in intangible assets to sales of 1 percentage point is associated with an 0.016 percentage point increase in the fixed-cost share.

Relationship between Fixed-Cost Share and Intangible Asset Intensity

	I	II	III	IV
Ratio of intangible assets to sales	0.023***	0.023***	0.016***/**	0.016***/**
Standard errors	0.003	0.003	0.003	0.003
Firm-clustered standard errors	0.006	0.006	0.007	0.007
R-squared	0.160	0.160	0.134	0.133
Observations	5,352	5,352	5,352	5,352
Year fixed effects	No	Yes	No	Yes
Firm fixed effects	No	No	Yes	Yes
Size polynomial	Yes	Yes	Yes	Yes

*** and ** represent significance at the 1 and 5% level, respectively. Significance is shown based on regular/clustered standard errors. Variables are winsorized at 1% and 99% tails.

Table 2.

Next, as intangible-intensive firms achieve higher fixed-cost shares, they should also display higher markups due to lower marginal costs. We run the model letting the fixed-cost ratio explain markups, controlling for size through a third-degree polynomial in log real sales $g(\cdot)$, together with firm- and year-fixed effects through α and ψ.

$$\mu_{it} = \alpha_i + \psi_t + \gamma \cdot \frac{\tilde{f}_{it}}{tc_{it}} + \beta' g(p_{it} \cdot y_{it}) + \varepsilon_{it}$$

Presented in Table 3, we find a significant relationship between firm's markup and share of total costs that are fixed. The relationship is consistent also after controlling for fixed effects across time and firms. The result suggests an increase in fixed costs as a fraction of total costs of 1 percentage point is related to an increase in markup of 11.46 percentage points, based on column IV.

Relationship between Markup and Fixed-Cost Share

	I	II	III	IV
Ratio of fixed- to total costs	10.579***	10.540***	11.478***	11.459***
Standard errors	0.283	0.282	0.339	0.339
Firm-clustered standard errors	1.833	1.809	2.394	2.387
R-squared	0.275	0.277	0.203	0.204
Observations	5,544	5,544	5,544	5,544
Year fixed effects	No	Yes	No	Yes
Firm fixed effects	No	No	Yes	Yes
Size polynomial	Yes	Yes	Yes	Yes

*** represent significance at the 1% level. Variables are winsorized at 1% and 99% tails.

Table 3.

Firms with higher fixed- and lower marginal costs are incentivized to use their economy of scale-production by selling more and undercutting their competition on price. De Ridder´s (2019) primitives state that firms take over production from their competition by offering higher-quality products, which are a result of R&D. We assess the expected relationship by regressing R&D intensity, defined as R&D expenditure as a share of same-year sales revenue, on the fixed-cost share. Firm- and year fixed effects are controlled for through α and ψ, and $g(\cdot)$ controls for size through a third-degree polynomial of log real sales.

$$\frac{rd_{it}}{p_{it} \cdot y_{it}} = \alpha_i + \psi_t + \gamma \cdot \frac{\tilde{f}_{it}}{tc_{it}} + \beta' g(p_{it} \cdot y_{it}) + \varepsilon_{it}$$

We find the coefficients to be significant, however less so when accounting for firm-fixed effects with clustered standard errors. The relationship is economically significant, as the sales revenue spent on R&D increases by 0.37 percentage points if fixed costs as a part of total costs increases by 1 percentage point, as stated in column IV.

Relationship between Research and Development Intensity and Fixed-Cost Share

	I	II	III	IV
Ratio of fixed- to total costs	0.539***	0.544***	0.363***/**	0.370***/**
Standard errors	0.047	0.047	0.056	0.056
Firm-clustered standard errors	0.125	0.126	0.155	0.159
R-squared	0.644	0.644	0.641	0.640
Observations	4,330	4,330	4,330	4,330
Year fixed effects	No	Yes	No	Yes
Firm fixed effects	No	No	Yes	Yes
Size polynomial	Yes	Yes	Yes	Yes

*** and ** represent significance at the 1 and 5% level, respectively. Significance is shown based on regular/clustered standard errors. Variables are winsorized at 1% and 99% tails.

Table 4.

As a result of the previous primitive, firms higher in relative fixed costs are prone to exhibit higher sales growth. Hence, the next regression tests the correlation between sales growth and fixed-cost share. Again, we control for firm- and time-fixed effects through α and ψ and for size with $g(\cdot)$.

$$\frac{p_{it} \cdot y_{it}}{p_{it-1} \cdot y_{it-1}} = \alpha_i + \psi_t + \gamma \cdot \frac{\tilde{f}_{it-1}}{tc_{it-1}} + \beta' g(p_{it} \cdot y_{it}) + \varepsilon_{it}$$

As positive sales shocks would automatically have a negative effect on same-period fixed-cost share, we follow De Ridder (2019) and lag the fixed-cost share one period. Similarly, negative sales shocks would be expected to increase same-period fixed-cost share and hence the lagged fixed-cost share and increase next-period sales growth, all else equal. Therefore, two and three period lags are also considered to control for this mechanism and are found to also yield a significant positive relationship between sales growth and fixed-cost share. Given the intuition that higher fixed-cost share leads to sales growth, it is reasonable to make use of a lagging effect in the model as the growth corresponding to a higher fixed-cost share would not be expected to be incurred the same period. This is also noted by De Ridder

(2019). Table 5 illustrates the significant regression results. Column IV tells us that an increase in the fraction of fixed costs in total costs of 1 percentage point is associated with a 1.38 percentage point increase in sales growth the next period.

Relationship between Sales Growth and Fixed-Cost Share

	I	II	III	IV
Lagged ratio of fixed- to total costs	0.609***	0.592***	1.385***	1.383***
Standard errors	0.049	0.048	0.085	0.085
Firm-clustered standard errors	0.096	0.093	0.248	0.247
R-squared	0.019	0.022	0.002	0.002
Observations	4,838	4,838	4,838	4,838
Year fixed effects	No	Yes	No	Yes
Firm fixed effects	No	No	Yes	Yes
Size polynomial	Yes	Yes	Yes	Yes

*** represent significance at the 1% level. Variables are winsorized at 1% and 99% tails.

Table 5.

7.2 Concentration

The theoretical framework resting on the primitives examined in the previous section expects concentration to increase due to the increased relative investments in intangible assets. Figure 5 shows the results from tracking the Herfindahl-Hirschman Index (HHI) across our sample sectors and time. Opposing our expectations, we find no significant change in concentration over the period 2010 until 2018 considering both the aggregated sample of all firms in one pool and separate sectors, with the exception of a slight increase in concentration in Information and Communication since 2015. Further, with an HHI consistently above 0.25, the Information and Communication sector is considered highly concentrated as opposed to the rest of the sectors displaying values corresponding to unconcentrated but not highly competitive industries. Appendix D complements the concentration analysis by considering the CR4 and CR8 (concentration ratio) standard measures, as well as a dynamic floating concentration ratio.

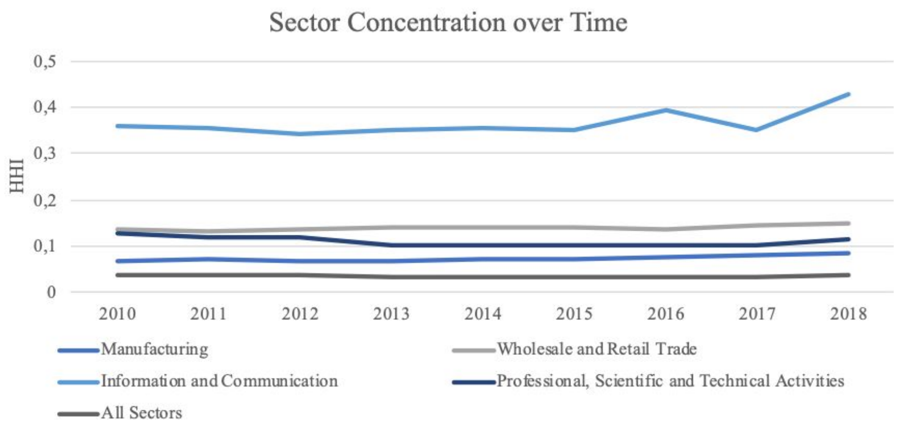

Figure 5: Concentration measured by the sector-specific Herfindahl-Hirschman Index (HHI). Data from Amadeus.

7.3 Research Productivity and Concentration Trends

The acquired data on patents and R&D spending allows us to compute research productivity ratios yearly. The PATSTAT data covers patents granted up until halfway through 2019. By computing research productivity for 2014, we allow for 4.5 to 5.5 years from the R&D activity (dependent on when during the year the R&D was conducted) to the end of the time window for patents to be granted. Based on the discussion on lags in Section 5.3, we consider the research productivity measure for 2015 and forward to be unreliable. Figure 6 displays all firms aggregated and Figure 7 illustrates the productivity development for specific sectors.

The overall development in Figure 6 reveals a steady increase in R&D spending over the whole period, with slight slowdowns in 2013, 2016 and 2018. After the eight years, deflated R&D expenditure has increased by 20 percent over the whole sample. The rise could have been expected to be higher considering previous observed trends. Ejermo and Kander (2011) recorded an increase just short of 150 percent between 1985 and 1998 among medium- to large Swedish firms. However, most of the growth occurred the last few years after the Swedish crisis in the early 1990s, and during the equivalent number of years 1985 to 1993 the increase amounted to approximately 50 percent. The number of granted patents on the other hand is steady during the first years but is decreasing thereafter, reaching 96, 85 and 61 percent of the 2010 level in 2013, 2014 and 2015 respectively. Although allowing for a total lag from R&D to patent grant of up to five and a half years for R&D conducted in 2014 as previously noted, and one additional year allowance for each year with R&D moving back in time, the number of patents granted is subject to a possible revision upwards. Hence, the drop in research productivity from 2013 to 2014 may be mitigated. Still, Figure 6 reveals a productivity decline in the sample of 17 percent until 2013 and 44 percent until 2014, relative to the 2010 level.

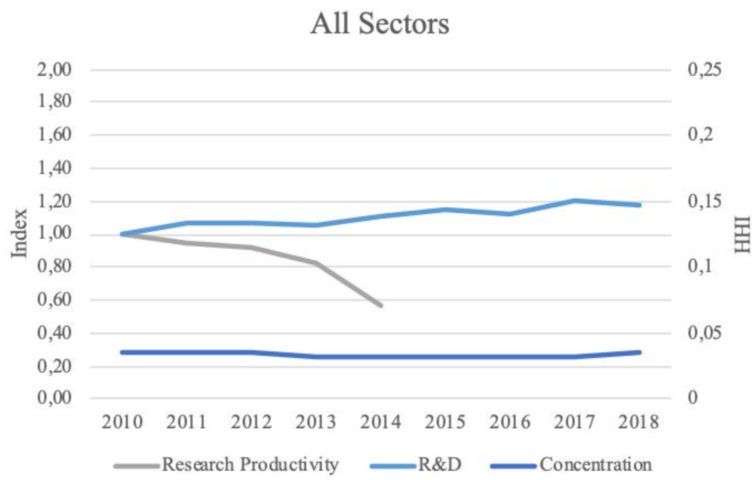

Figure 6: Research productivity (left axis) is granted patents divided by deflated R&D, based on patents with earliest filing year one year after the R&D expenditure. R&D (left axis) is deflated by a high-skilled labor wage index. Left axis index 2010 = 1. Concentration (right axis) is the Herfindahl-Hirschman Index (HHI) for all sectors. The R&D deflator is based on Statistics Sweden's Lönestrukturstatistik, see Appendix A. Patent data from PATSTAT, accounting data from Amadeus.

Figure 7 shows a similar development for the Manufacturing industry as for the total sample, although with a slightly higher R&D growth reaching 42 percent in 2017 and 32 percent in 2018 above the 2010 expenditure. The yearly number of granted patents is relatively stable yet slowly declining, but in combination with increasing R&D resulting in a moderate productivity decline reaching 89 and 44 percent of the 2010 level in 2013 and 2014, respectively.

Out of the sectors examined, Wholesale and Retail Trade exhibit by far the largest decline in research productivity the first year. This is jointly due to a decline in patents and a R&D increase by 13 percent the same year. Productivity continues to decline mainly as a result of fewer patents as R&D is stable. R&D then surges to a spending in 2015 at 42 percent higher than the beginning of the period and before decreasing.

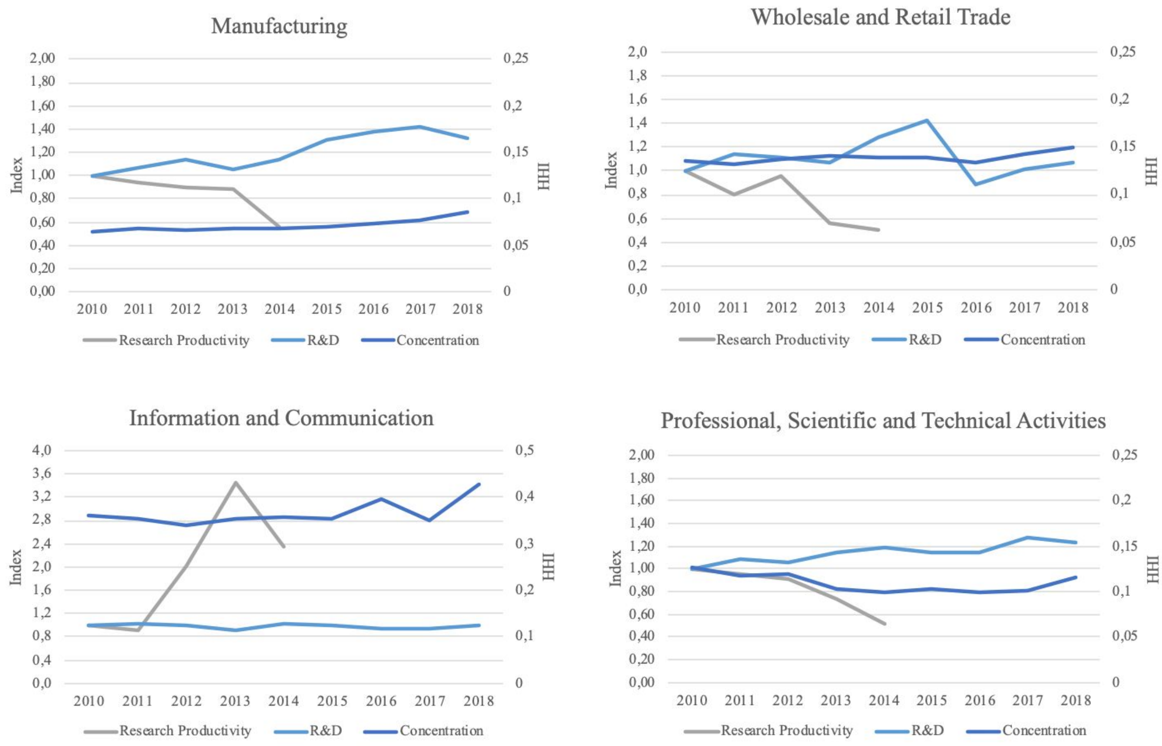

Figure 7: Research productivity (left axis) is granted patents divided by deflated R&D, based on patents with earliest filing year one year after the R&D expenditure. R&D (left axis) is deflated by a high-skilled labor wage index. Left axis index 2010 = 1. Concentration (right axis) is the sector's Herfindahl-Hirschman Index (HHI). The R&D deflator is based on Statistics Sweden's Lönestrukturstatistik, see Appendix A. Patent data from PATSTAT, accounting data from Amadeus.

The Information and Communication sector stands out with a surge in research productivity combined with constant R&D effort over time. The R&D index reaches a high- and low-point of only 1.03 and 0.9, respectively. The productivity incline is explained by three firms with a combined large portion of their sector's patent and R&D activity that all grew their number of granted patents one or more years during the peak period. The effect from this small number of firms is amplified as the particular sector is a comparatively small sample of 43 firms active in R&D. Hence, the effect on the aggregate research productivity measure is insignificant.

As for Manufacturing, R&D is consistently above the initial level for the Professional, scientific and Technical Activities sector with a maximum in 2017 of 128 percent of the expenditure in 2010. Consistent with the full sample as well as the Manufacturing and Wholesale and Retail trade industries, research productivity is declining. The productivity index is down 27 and 48 percent from the 2010 level in 2013 and 2014 respectively.

With the exception for the Information and Communication industry, research productivity is declining in all sectors as well as for all firms combined. This is the case all the while we experience no significant change in concentration.

8. Discussion

Firstly, we can consider the fit of the model to be sufficient based on the data from our Swedish sample since all four relationships were found to be significant. Hence, even in Sweden, high-intangible firms should have a higher fixed-cost share, higher markups, engage more in R&D, and experience a higher sales growth.

Regarding the input in the model, we observed an increase in investments in intangible relative investments in tangibles. There are two major differences between this measure and De Ridder's (2019) introduction of a group of high-intangible firms. Firstly, the relative increase in investments in intangibles in reality has been gradual for a long time, while De Ridder's introduction of a group of high-intangible firms constitutes a shock to the economy. Hence, the predicted outcomes of the model should in reality be expected to be more gradual and not perfectly match the predictions of the model in terms of timing. Secondly, the model includes an inherent firm heterogeneity where the introduced firms are fundamentally different from the already existing firms. Since we do not analyze the extent to which the relative increase in investment in intangibles is heterogeneous across firms, we cannot exclude the possibility that the change is fully homogeneous across firms. However, it seems reasonable to assume that the relative increase in investments in intangible assets is not fully symmetrical across the firms, hence rendering similar results as the model.

Assuming a similar firm heterogeneity as constructed in the model of De Ridder (2019), we should expect an increased concentration, which has not been observed during the analyzed time period. Hence, an explanation for the seemingly constant firm concentration since 2010 might be that the relative increase in investments in intangibles has been relatively homogenous in recent years. Another explanation for the relatively constant concentration might be that Sweden has already transitioned to the new balanced growth path with high concentration as predicted by De Ridder (2019). As argued by Ejermo and Kander (2006) when both analyzing new data and analyzing findings from previous works, Sweden has for a long time experienced problems with R&D being highly concentrated (opposing the relatively low levels of concentration found in our sample), and it particularly increased during the period 1965-1994. However, we cannot confirm that this increase stems from an increasing use of intangibles. As well, an uncertainty regarding our findings on concentration is that we have only analyzed a sample of Swedish firms, and not the concentration of all Swedish firms.

Further, even if concentration has been observed to be constant within the sample, we have observed a decrease of research productivity in our sample between 2010-2014, using granted patents over R&D spending as a measure. It is difficult to tell if this is part of a long-run trend or if it is a short run fluctuation, but it matches the indications of a decreased research productivity particularly during these years observed in the background section. The decrease can stem from reasons other than an increased concentration, from an increase in concentration that we failed to capture, or from a combination of the two.

Finally, even if the results on research productivity seem to correspond to the indications from Section 2, we must acknowledge the limitations of using patents as a measure for new ideas in the economy. Firstly, there are many factors included in TFP that are not patented, like educational level of the labor force and the efficiency of institutions for example. Thus, technology increase in the form of patents only explain a fraction of TFP growth. Secondly, there are patents that do not contribute to an increased technological level of the economy. As discussed in the literature review, firms can engage in rent-protecting activities such as strategic patenting. The findings of Torrosi et al. (2016) provides an indication of the extent of this problem. In their paper they analyze survey answers regarding the motives behind- and the use of- patents in Europe, Israel, USA, and Japan between 2003 and 2005. With regards to motives, they categorized a patent as a blocking patent if the firm answered above three on a five-point scale on the question if blocking competitors was an important motive behind the innovation. In terms of use, a patent can be used internally in the firm's own products or externally by for example selling or licensing the patent to other firms. If any of the uses was fulfilled, a patent was categorized as used, rather than non-used. On the basis of these categorizations, they form three categories of patents. Firstly, the category "Commercial use" include patents that are used in any way, regardless of if they have a blocking motive or not. Since used patents often also fills a blocking function, even if it is not the main motive, the authors see no reason to divide the used patents into blocking and non-blocking ones. Secondly, "Strategic non-use" refers to non-used patents with a blocking motive, and thirdly, "Sleeping patents" refers to patents that are neither used nor have a blocking motive. For our analysis, only the dichotomy of use vs. non-use is relevant, since it is the use of a patent that contributes to innovation and the development of the ideas in the economy. Whether a non-used patent has strategic motives or not is not relevant for our analysis. Hence, category two are three will further serve as a common category of non-used patents. Further, they present their finding in terms of the percentage of patents that are in the different categories in total, in differences geographical areas, across different technological categories, and across different firm sizes. Overall, only 60.6 percent are found to be used, while the number is 62.0 percent for Europe and Israel. With regards to firm size, they find that the use is significantly lower among large firms, while both the strategic non-use and the sleeping patent are more frequent (76.5, 77.0, 56.2 percent of patents are found to be used among small, medium sized and large firms, respectively). This supports the theories and findings discussed in the literature review regarding how research productivity decreases as innovation gets concentrated among a few big firms. In terms of different technological categories, the use was highest in Consumption and Construction (75.8 percent) and in Process engineering (66.9 percent), and lowest in Chemicals and Pharmaceuticals (52.9 percent). To make our measure of research productivity more accurate we would have had to adjust the patent data using these findings as estimations for how many of the patents that are actually used across different sectors and firm sizes.

9. Conclusion

To conclude, Sweden has long experienced slow TFP growth despite extensive R&D efforts, often referred to as the Swedish Paradox. While previous research has found research productivity to increase in Sweden between 1985-1998, both recent data and theories on the role of intangibles on research productivity suggests it has been, or should have been, declining in recent years. Hence, we have sought to update the story of Swedish research productivity and explored the use of intangibles as an explanation for the development.

Our findings include a significant fit of the relationships serving as building blocks in the model of De Ridder (2019), hence supporting that high-intangible firms have higher fixed cost-share, have higher markups, engage more in R&D, and experience higher sales growth. In line with our expectations, research productivity, measured as granted patents over R&D spending, has been found to decrease within our sample during 2010-2014, even if it does not seem to have been a result of an increased concentration. However, related literature suggests that Sweden has already experienced a long period of increasing concentration, and that this has been a problem for Swedish research productivity for a long time.

References

Akcigit, U. and Kerr, W. R. (2018). Growth Through Heterogeneous Innovations. *Journal of Political Economy, 126(4),* 1374-1443.

Bessen, J. (2017). Information Technology and Industry Concentration. *Boston University School of Law, Law & Economics Paper Series, 17-41.*

Bloom, N., Jones, C. I., Van Reenen, J., and Webb, M. (2020). Are Ideas Getting Harder to Find? *American Economic Review, 110(4),* 1104-1144.

Criscuolo, C., Bajgar, M., and Timmis, J. (2019). Supersize Me: Intangibles and Industry Concentration. *Manuscript in preparation.*

Dang, J. and Motohashi, K. (2015). Patent Statistics: A Good Indicator for Innovation in China? Patent Subsidy Program Impacts on Patent Quality. *China Economic Review, 35(C),* 137-155.

De Ridder, M. (2019). Market Power and Innovation in the Intangible Economy. *Cambridge Working Papers in Economics, 1931.*

Dinopoulos, E. and Syropoulos, C. (2007). Rent Protection as a Barrier to Innovation and Growth. *Economic Theory, 32(2),* 309-332.

Ejermo, O. and Gabrielsson, J. (2007). International Cooperation for Swedish Inventors – An Exploratory Study. *Paper presented at EPIP (European Policy for Intellectual Property).*

Ejermo, O. and Kander, A. (2006). The Swedish Paradox. CIRCLE Electronic Working Paper Series. 2006/01, CIRCLE: Lund, Sweden.

Ejermo, O. and Kander, A. (2011). Swedish Business Research Productivity. *Industrial and Corporate Change, 20(4),* 1081-1118.

Fernald, J. G. (2014). Productivity and Potential Output Before, During, and After the Great Recession. *NBER Macroeconomics Annual 2014, 29,* 1-51.

Korinek, A. and Ng, D. X. (2019). Digitization and the Macro-Economics of Superstars. *Manuscript in preparation.*

Torrisi, S., Gambardella, A., Giuri, P., Harnoff, D., Hoisl, K., and Mariani, M. (2016). Used, Blocking and Sleeping Patents: Empirical Evidence from a Large-Scale Inventor Survey. *Research Policy, 45(7),* 1374-1385.

Wang, N. and Hagedoorn, J. (2014). The Lag Structure of the Relationship between Patenting and Internal R&D Revisited. *Research Policy, 43(8),* 1275-1285.

Appendix A. Wage Index Deflators

Figure A1 provides a visual comparison of the data on mean yearly wages for men and women with two different levels of education, in this form as indexed deflator series. Both education levels are alternative proxies for high-skilled labor. For this figure, the index is normalized in year 2000 for comparison. Monthly average wage data is obtained from the database Lönestrukturstatistik, compiled by Statistics Sweden. Data is available from 1991 and 1995 for the higher and lower level of education considered, respectively. Throughout the paper, the deflator used for R&D is based on the lower of the two education levels.

Figure A1: Inverted wage development. Index 2000 = 1. Data from Statistics Sweden's Lönestrukturstatistik.

Appendix B. Capitalized Costs of Intangibles

The cost of intangibles used by a firm is expensed the corresponding accounting period and included among the costs on the income statement. As intangibles also can be held to create benefits beyond the current period, the acquisition costs can be capitalized to the balance sheet as intangible assets, incurring an expense when amortized over their useful life. Along with R&D, patented technologies and several other asset subgroups, IT systems and software can thus be found on the balance sheets of firms. This is generally stated by the wide-reaching and accepted accounting standards IFRS and GAAP.

When purchasing software and IT-systems with perpetual licenses and finite subscription licenses that stretch further than the accounting period, firms can have the cost of the license, installation and testing capitalized. For externally or internally developed software, conditional on certain accounting criteria, firms can have external development fees, costs to obtain, and other costs derived directly from obtaining the asset capitalized as an intangible asset, when developed or obtained for the purpose of internal use. This regards stand-alone software and systems. That is the case when the hardware used to utilize the software can operate in its own right without the software, such as computers and Enterprise Resource Planning (ERP) systems.

Appendix C. Original Data Variables

Amadeus Financial Account Data

Definitions from the database documentation.

Sales is total sales revenues. The BvD Code is TURN.

Cost of goods sold is cost of goods sold, production and services. Costs directly related to the production of the goods sold and depreciation of those costs. The BvD Code is COST

Wage bill is employee costs including pensions. The BvD Code is STAF.

Materials and intermediate inputs is raw costs for materials, consumables and goods for resale. The BvD Code is MATE.

Other operating expenses is all costs not directly related to the production of goods sold such as commercial costs, administrative expenses etc. and depreciation of those costs. The BVD Code is OOPE.

Operating profit is operating P/L [=EBIT]. All operating revenues less all operating expenses. The BvD Code is OPPL.

Research and development is total amount of expenses research and development activities. The BvD Code is RD.

Intangible assets is total intangible fixed assets. The BvD Code is IFAS.

Number of employees is total number of employees included in the company's payroll. The BvD Code is EMPL.

Notes: For year consistency with firms using split financial years, fiscal year ends through March 31 of the subsequent year are standardized as if the fiscal year end is December 31. Financial account data on Swedish firms is provided by UC and harmonized by Bureau van Dijk / Moody's Analytics.

Appendix D. Complementary Measures of Concentration

Together with the Herfindahl-Hirschman Index (HHI) used in the paper, the concentration ratio of the four (CR4) and eight (CR8) largest firms constitute standard the market concentration measures. The concentration ratios consider the combined market share for largest firms in terms of sales. Figure A2 illustrates the concentration ratios for the aggregated sample of firms.

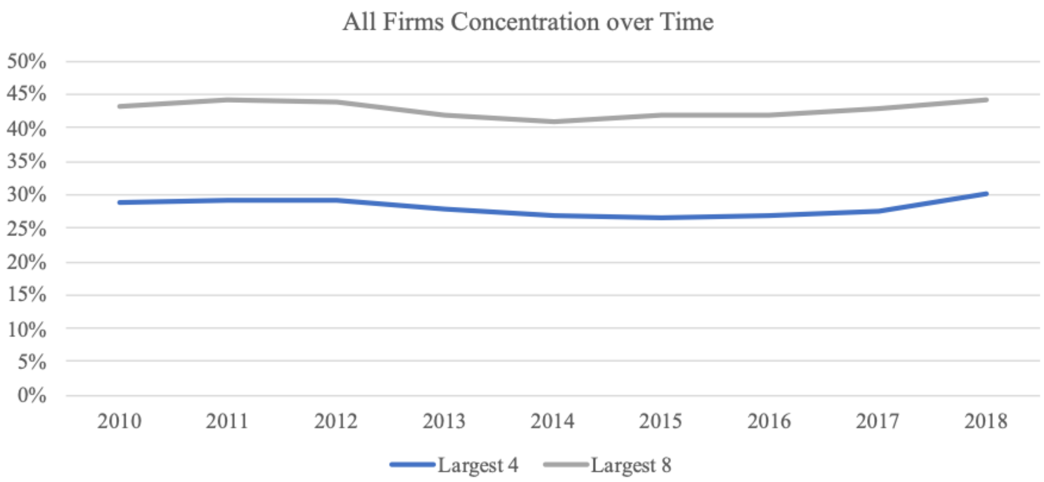

Figure A2: CR4 and CR8 for all sample firms pooled. Data from Amadeus.

Figure A3 shows an alternative concentration ratio with a floating number of firms for the combined market share. The number of firms included is thus based on a chosen percentage and several cut-off points for inclusion in the group are considered.

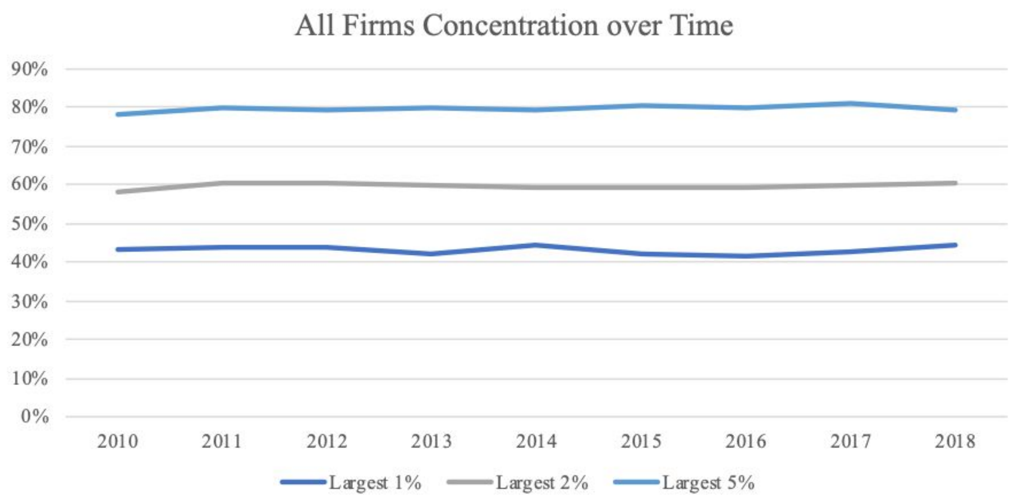

Figure A3: Combined market share in terms of sales for the largest one, two and five percent of firms. Data from Amadeus.

In Figure A4, the same measure based on the largest five percent is computed within each sector. We can conclusively say that the trend in concentration based on market share in terms of sales is indifferent to the measurement techniques applied, as there is no significant change for specific sectors or for the aggregated sample over the period.

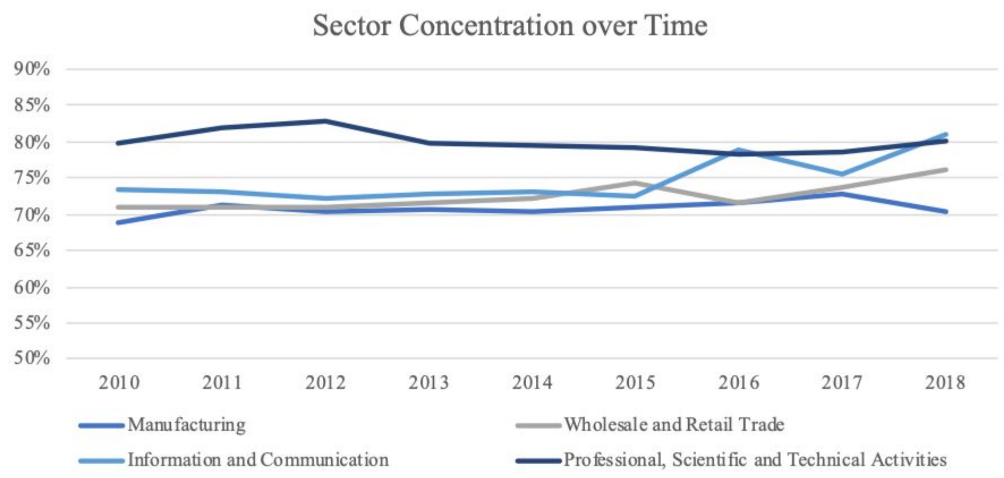

Figure A4: Combined market share in terms of sales for the largest five percent of firms within each industry. Data from Amadeus.

To contrast the measures of sales, Figure A5 consider concentration in R&D. The combined R&D expenditure share of top five percent spenders within each industry. In the Information and Communication sector, one firm alone is disproportionately dominant in our sample with regard to R&D expenditure, hence the high relative concentration.

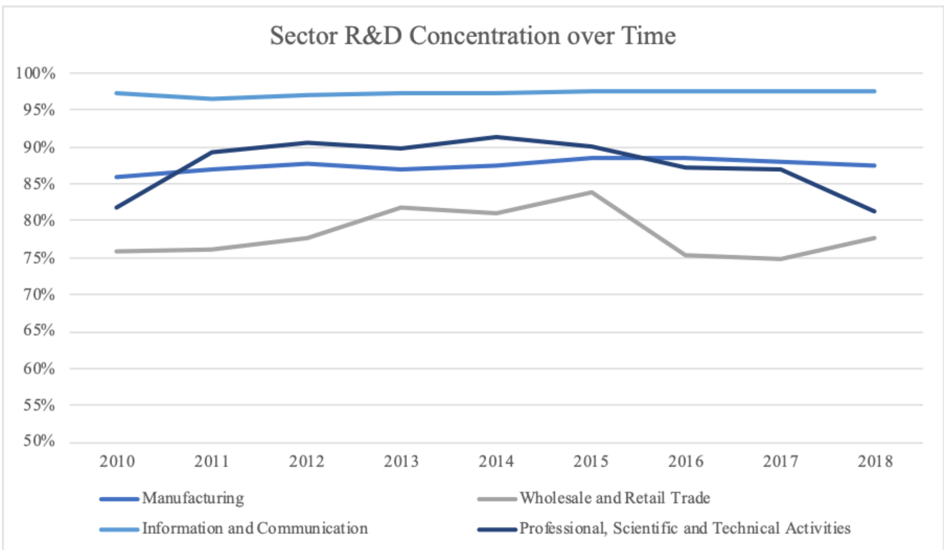

Figure A5: Combined share R&D expenditure for the largest five percent of firms within each industry. Data from Amadeus.