TO WHAT EXTENT IS INNOVATION DETERMINED BY MARKET STRUCTURE? EXAMINE HOW INDUSTRY CONCENTRATION AFFECTS THE RATE OF TECHNOLOGICAL CHANGE.

Department of Economics, University of Essex

Michael Capleton

Registration Number: 1202670

Supervisor: Dr Katharine Rockett

Abstract

This paper examines the effect of market structure on innovation by empirically examining the relationship between industry concentration and research and development expenditure. Controlling for industry differences, this paper tests for the inverted-U relationship proposed by Aghion et al (2005). Evidence of this relationship is found, and although robust against heteroskedasticity and endogeneity, it may be sensitive to a change away from pooled OLS methods.

Contents

1. Introduction

A large incentive for firms when innovating in the manufacturing process is to reduce their marginal cost, which can lead to much higher price-cost margins and increased profitability. A further benefit for the firms who innovate is to increase market power within their industry, by producing either at a lower marginal cost – and subsequently offering perhaps a lower price – or producing new products. The latter can provide the firm with increased demand as their product is of a better quality than that of their rivals, or tap into new demand in a new market. Improved products without raising prices are a core objective of regulators, and this can be achieved through innovation.[1] Regulatory bodies such as the Competition and Markets Authority (CMA) in the UK and the European Commission in the EU may therefore be interested in which levels of industry concentration are most conducive to innovation.

This paper tests to see whether the level of concentration is a factor in how much firms spend on 'innovating' and if so, to what extent. Section 2 will provide a review of the existing literature in the area of innovation and market structure, discussing thematically the various hypotheses and proposed effects. Section 3 will then describe the dataset to be used in this paper and its strengths and weaknesses. Each variable will be introduced, explained, and critiqued in terms of how well the variable can explain what it proxies. The data does carry limitations, and so these will be discussed also. Section 4 outlines the methods being proposed and tested in the paper. Tests are carried out on a linear relationship, followed by an extension to non-linearity and finally robustness checks. Section 5 outlines and critically discusses the results and outcomes

[1] The European Commision's 'Competition Policy' states that it seeks to create competition that "creates incentives for companies to innovate"

from the empirical testing, before the paper concludes in section 6. Extra tables and graphs referred to in the text are available in the appendix at the end of the paper.

For decades following the work of Joseph Schumpeter in the middle of the twentieth century, a positive relationship between concentration and innovation was debated amongst economists and econometricians, who tested Schumpeter's theory that monopoly power provides the best basis for innovators and that it is monopoly power that is the key incentive for innovators.

For these reasons, a new theory was born in 2005 through the work of Phillippe Aghion, Nick Bloom, Richard Blundell, Rachel Griffith and Peter Howitt, who discussed the possibility of an inverted-U shape relationship between competition and innovation – previously considered by Scherer (1965). Since their paper, this inverted-U hypothesis has been the subject of much testing and debate, with fairly strong evidence in favour of it. This paper tests specifically for the same shape, following a similar approach to Polder & Veldhuizen (2012). Using pooled OLS estimation, models are tested that control for industry differences in employment, wage levels and market locality through an imports measure.

This paper is unique for two main reasons, besides of course the unique model established in the paper. Firstly, none of the literature reviewed in this paper tests in the two countries used here – Germany and Italy. These two EU countries may give us different results than those found previously in the US or UK, partly perhaps due to the differing laws, institutions and regulatory bodies that comprise each country's innovation and entrepreneurial framework. Secondly, where Aghion et al (2005) tests for a relationship between patents and the Lerner index, this paper examines a potential relationship between R&D expenditure and the Herfindahl-Hirschman index (HHI).

These differences in proxies provide a possible robustness check to the results found in Aghion et al (2005).

Using pooled OLS, evidence of the inverted-U relationship is found for both countries. The turning point of this inverted-U is then mathematically estimated using simple calculus to provide an idea of the level of industry concentration that can best produce innovations – a particular point of interest for regulators. Furthermore, these results are tested for two main robustness checks – heteroskedasticity and endogeneity. As heteroskedasticity is suggested in the econometrics, robust regressions are used and the results remain consistent. There is a possible endogeneity concern in terms of our dependent and main independent variables. Successful innovations may lead to an increase in market power – leaving a potential two-way causality problem, highlighted and controlled for in other literature. This paper follows the control technique of Polder & Veldhuizen (2012), introducing lag terms on the concentration variables – the results again remain robust.

These results are strong, however, only with the use of pooled OLS. When a fixed effects model is proposed and tested using the same controls, the results differ rather dramatically, suggesting the results may be sensitive to the model used.

2. Literature Review

This literature review will attempt to give a brief summary of the existing theoretical and empirical research on the possible relationship between market structure and innovation.

2.1. Schumpeterian hypotheses

Schumpeter's work broadly consisted of two views, described as the "two Schumpeters" by Acs & Audretsch (1988) and commonly referred to as Mark I and Mark II. In the former, originally described as "Competitive Capitalism" (Schumpeter, 1934), the new firm was the innovator. The entrepreneurial flair of an individual was the main factor behind an innovation. Schumpeter expected existing firms, often part of large oligopoly industries, to continue with current products and processes; innovation was carried out by entrepreneur, setting up a new firm. Schumpeter Mark II, or "Trustified Capitalism" (Schumpeter, 1947), states that most innovation is undertaken by larger firms with some degree of monopoly power. At this point, Schumpeter argued that simple inventions had already been made; it was now mainly costly, technical innovation that was possible.

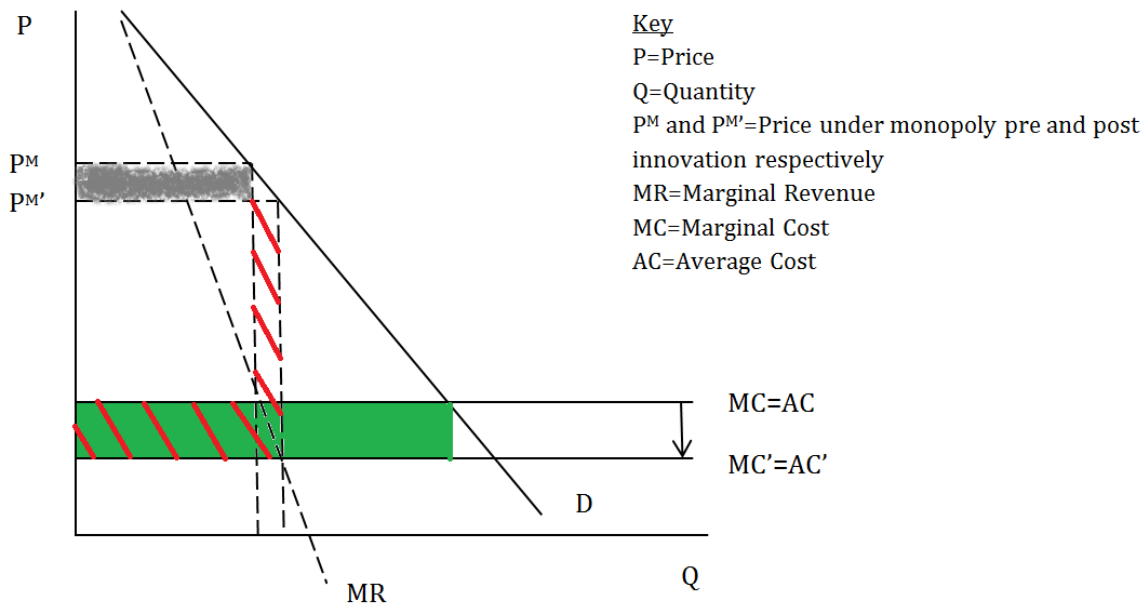
Galbraith (1952) fleshes out and supports Schumpeter's point by arguing that the greater financial demands of these innovations favour the large firms making supernormal profits as they can more easily afford costly research and development. At the time of Galbraith's work, there was very little empirical evidence for or against this notion.

Kamien & Schwartz (1982) review the Schumpeterian hypotheses. Their work distinguishes between two contributors to the link between innovation and monopoly power, which are characterised as anticipation of power and possession of power. For the first point, it is suggested that some monopoly power is necessary to realize supernormal profits; the end goal of innovation. It is then put forward that possession of monopoly power is conducive to innovative activity (page 27, 1982).

Kamien & Schwartz (1982) suggest innovation must be financed internally. This stems from the moral hazard problem of asymmetric information, where a financial institution does not know the feasibility of a R&D project. For an institution to loan a large sum of capital to a firm, it needs to believe the firm will be able to repay the investment. Kamien & Schwartz (1982) asserts this as a further advantage for a firm making monopoly profits – an indicator of previous success. Moreover, it is suggested these monopoly profits allow the firm to attract the most innovative people, as the successful firms can afford to pay higher wages. Contrary to this, Shrieves (page 333, 1978) states that more than half of (R&D) is financed by the federal government, so perhaps firms in competitive markets are able to gain finance for such experiments. Empirically, Scott (1984) attempted to test for whether government financed R&D was actually a substitute or a complement to privately-financed R&D.

2.2. The Replacement Effect

Kenneth Arrow's work (1962) appears to oppose Schumpeterian hypothesis. He compares the additional profit to be gained from undergoing some process innovation (that is, reducing marginal and average of production) in perfect competition and monopoly markets. He shows mathematically the profit increase for a monopolist when reducing marginal cost should be less than for a perfectly competitive producer – where we assume that marginal cost is equal to average cost in such markets. This is because the perfectly competitive firm can capture the whole market, given homogeneous goods in the industry, if we assume either perfect intellectual property rights or the possibility of secrecy. A monopolist already earns some (pre-innovation) supernormal profit and just 'replaces' this profit with a small improvement. For this reason a monopolist may have less incentive to innovate and increase its profits, than a perfectly competitive firm who can move to achieve positive profits from an original position of zero profit. Arrow (1962) calls this the 'replacement effect'. This is shown below:



The gain from innovation for the perfectly competitive firm is shown by the fall in MC and AC (filled green). This firm now captures the entire market, charging a marginally lower price than before to attract all consumers, and realising supernormal profits for the first time. The monopolist however only sees a slight improvement in profits through the fall in costs, but already had the entire demand so sees no increase in that respect. This is shown by the red lined area minus the grey shaded area.

2.3. Empirical Research

To summarise from the theoretical review: Schumpeter and Galbraith argued that some monopoly power is necessary for innovative activity, stating a few key reasons as to why conditions in concentrated markets are suitable for technological progress, to summarise the reasons for this from Kamien & Schwartz (1982):

a) Realising extraordinary profits is the incentive for innovation; a period of monopoly power is necessary to achieve these profits.
b) More profit can be realised by a large firm, because of economies of scale. Note that large firms do not necessarily possess market power, but it is usual.
c) Innovation is costly and must often be financed internally – supernormal profits make this affordable. It also allows recruitment of more technical and creative employees.

The exact relationship between market structure and innovation is left open. As a result, this paper now turns to a review of empirical evidence. Throughout empirical research, I found there to be two main measures that have been used for innovation; number of patents (output measure) and level of R&D (input measure). Measures of concentration and competition also vary between papers in this literature review. The data section of this paper goes someway to explaining differences between these proxies, and their strengths and weaknesses.

2.3.1. Testing Schumpeterian Hypotheses

Evidence of a positive linear relationship:

Horowitz (1962) finds that in more highly concentrated industries, firms are more likely to maintain internal research, and less likely to turn to outside organizations to carry out research. He also found from his results that in more highly concentrated markets, research labs are likely to be held by more firms than just the top few. This is an interesting result, as it supposes even the smaller players carry out their own research, in markets dominated by a few large firms. The crucial result regarding the support of Schumpeterian hypothesis though, is based on research expenditure as a proportion of sales. Using his two samples of data, Horowitz (1962) finds that the research expenditure/sales ratio rises linearly with market concentration. Horowitz does however mention that it is not clear whether the spending on R&D (perhaps a new laboratory) has come as result of monopoly profits, or whether market power is created by spending on innovation. This may be a two-way causality problem.

Mansfield (1963) finds more mixed results on the link between concentration and innovation. Taking data from three industries (coal, petroleum and steel) in the early twentieth century, Mansfield tests for whether the largest firms in each industry carry out the majority of innovations. His results varied across industries. Evidence supporting Schumpeterian hypothesis was found in petroleum and coal markets, with the largest four firms responsible for a more than proportionate share of innovations. However, in the steel industry, the result was not consistent.

In a paper by Corsino et al (2008), evidence was found supporting the theory of constant returns to firm size on innovative output. This, alongside Mansfield's findings, brings the correlation between firm size and innovation into question. However, Mansfield's paper concentrates more specifically on the relationship between innovation and monopoly power, not firm size.

Similarly, the work of Shrieves (1978) also finds the relationship to differ across industries. Before testing, Shrieves argues that oligopoly firms possess a double incentive to innovate. For product innovations, the improved product will increase the individual firm's demand curve. Process innovations will lead to a reduction in price less than the reduction in marginal cost as a result of innovative activity – causing a profit gain. Using a sample of 411 firms and R&D intensity as measure of innovation, Shrieves (1978) conducts a regression with a dummy variable of durable goods. The motivation for the inclusion of a dummy variable is to test for differences between industries, which we have seen as present in Mansfield (1963). Overall and without the dummy variable, concentration had a positive and significant regression coefficient on a firm level. This provided his conclusion that there will be more innovations in concentrated industries (page 338, 1978). However, when the model was re-estimated on industry-level the results changed. Accounting for these differences, the results were similar to Mansfield (1963).

Evidence of a negative linear relationship:

Evidence against Schumpeterian hypothesis can be seen in work by Acs & Audretsch (1988). Using innovative output per employee as the dependent variable, the concentration coefficient found is significantly negative. This suggests a negative correlation between market concentration and innovation in reverse to what Joseph Schumpeter and John Kenneth Galbraith proposed. The regression results found provide evidence supporting the notion that competitive markets are a more effective engine for innovation than are highly concentrated industries, offering support to Arrow (1962). They conclude strongly:
"Monopoly power deters innovation" (Page 137, Acs & Audretsch, 1988).
Possible explanations of this result could be the theory of diseconomies of scale. Perhaps in a large firm, it could be argued that employees are not monitored so strictly, because an innovation may not make much difference. Whereas in a small firm an innovation may make a larger difference, and management can more easily monitor staff to try and produce as many innovations as possible. The contrasting results with

Horowitz (1962) may be caused by the change of dependent variable from research expenditure to innovation per employee.

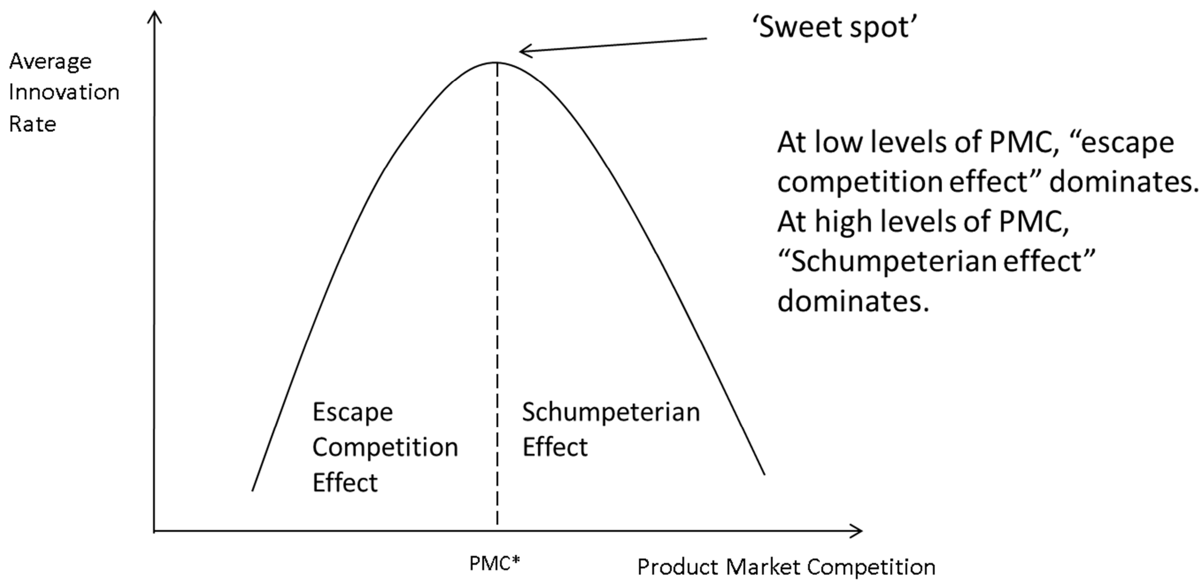
Further evidence of a negative relationship between industry concentration and innovation was found by Blundell et al (1999), who using a panel of 340 British manufacturing firms; found that it was the more competitive industries that produced more innovations. However, evidence supporting the Schumpeterian view that firms with market power innovate more was also found, as within these industries, it was the firms with high market share that were responsible for the majority of innovations.

First suggestion of a non-linear relationship:

A reputable name in Industrial Economics, Scherer (1965), tested for a correlation between number of patents and four-firm concentration ratios. He found a very small, positive link between market concentration and number of patents, pointing out that number of patents may not be an effective enough dependent variable, given the differences across industries. He concludes from his results that if there is a relationship between concentrated industries and technological progress, it is minimal. It appears he is unconvinced that this is strong evidence, expecting there to be no real link between the two. This is confirmed by a concluding statement:
"Inventive output does not appear to be systematically related to variations in market power" (page 1121, Scherer 1965).

2.3.2. The Inverted-U Relationship

Mentioned and hinted to in Scherer (1965) is the possible idea of an inverted U shape relationship between market concentration and innovation. This theory was developed significantly further and justified through two contrasting effects in Aghion et al (2005). The theory from that paper is shown graphically below (not from source).



The reasoning behind the inverted-U theory is proposed in Aghion et al (2005) as two distinct effects. Firstly, at low levels of product market competition (PMC), the escape competition effect dominates, as the motive is strong – similar to the replacement effect in Arrow (1962). Followers also have a low incentive to innovate in such industries because it depends on the amount of "neck and neck" firms – how many are escaping competition (trying to catch up). For these reasons, industries will move into a position with leaders and followers, where the followers are stuck, and this leads to a situation with low innovation. For high levels of PMC, the Schumpeterian effect dominates, with firms led by the incentive to make supernormal profits, and again become stuck in a state of low innovation. Only in a situation where PMC is at a medium level do firms not become stuck in a low innovation state. Aghion et al (2005) found strong evidence of this inverted-U relationship using a pooled OLS log-linear regression model. They use the term 'neck-and-neck' throughout the paper to describe firms at similar technological levels to one another. These are the innovative firms according to Aghion et al. The variable for innovation was number of patents, using 311 firms in the UK between 1973 and 1994. The measure of product market competition here is the average Lerner Indexes across firms in each industry.

There is empirical evidence supporting such a result in the works of Levin et al (1985), who, similarly to Shrieves (1978), factors out for differences between industries using dummy variables. When these factors are considered, his results are consistent with the inverted U relationship theorised above. Levin et al test for a relationship between four-firm concentration ratios (C4) R&D/Sales and innovative output (patents) using both OLS and 2SLS. They include a squared term on the concentration variable and find a negative coefficient on this variable, evidence supporting the inverted-U proposed in Aghion et al (2005). Mathematically, Levin et al find this turning point at levels of C4 between 50 and 60.

A later paper by Haruyama (2006) proposes three further reasons for the inverted-U relationship, building on Aghion et al (2005). The first is that innovation in one industry may hinder or help progress in another, depending on whether the markets are substitutes or complements. Haruyama then argues, using steam engines as an example that innovations are built upon previous progress. His final reason attempts to address the problem of measuring innovation, by accounting for both R&D expenditure and rent protection (patents). In a model looking at both of these factors, an inverted-U theory is proposed.

Scott (1984) finds less conclusive evidence for the inverted U theory. Using privately financed R&D as the dependent variable, it is found that his original result similar to that of Levin et al (1985) disappears when he accounts for other factors affecting the level of R&D. These factors include technological opportunity, condition of entry, and ability to collude – which may cause a sharing of research information. Interestingly, a test is also carried out to find whether government financed R&D is a substitute for privately financed R&D. From Kamien & Schwartz's (1982) review of Schumpeterian hypothesis, it is suggested that firms must finance R&D internally, but others (Shrieves, 1978) suggest that much R&D is financed by the federal government. Scott (1984) concludes that there seems some complementary effect between government-financed R&D and privately-financed R&D. He rejects the hypothesis that the two are substitutable.

Revisiting Aghion et al (2005), Hashmi (2011) tests for the inverted-U relationship between competition and innovation using US data from 1976 to 2001, rather than the UK data in Aghion et al (2005), whilst closely following the empirical approach. Predicting to find similar results, Hashmi actually finds evidence of a positive relationship between competition and innovation. However, he does find that at the more 'neck-and-neck' industries, the positive relationship is steeper – offering some support of Aghion et al (2005).

Polder & Veldhuizen (2012) estimate the non-linear relationship using data from Netherlands between 1997 and 2006. They again follow Aghion et al (2005), by introducing a squared term on the level of competition in their industry-level analysis. They also test for the relationship at firm level in order to test the exact mechanism suggested in Aghion et al (2005). In their industry-level analysis, their results are consistent with the inverted-U shape relationship. They also test for robustness by adding a one-year lag on the competition term – profit elasticity, and their results remain the same.

My paper tests for a relationship between industry concentration and innovation, using HHI and R&D expenditure as the respective proxies. I first estimate a linear relationship to test Schumpeterian hypotheses, controlling for employment and wage level differences between industries. Following Levin et al (1985), Aghion et al (2005) and Polder & Veldhuizen (2012), I then use a squared term on the concentration variable to investigate the inverted-U relationship.

3. Data

Data is formed of three OECD datasets. Firstly, data was obtained on the dependent variable (R&D) from OECD's Business Enterprise Research and Development database on structural analysis of R&D expenditures in industry (SIC Rev.4). This was available for all OECD countries on 100 industries over a period from 2000 to fairly recent, dependent on country. For the explanatory variables, data was available from a similar source on the same 100 industries plus 21 more. The database used here was OECD's structural analysis database. HHI has been used as the measure for industry concentration, and this proved more difficult to find. I have used an estimate for HHI as supplied in the OECD book 'Structural and Demographic Business Statistics 2006' using the size distribution of firms[2]. Finding data on industries that perfectly matched up with the SIC codes from the first two datasets was a challenge, but we are left with two strongly balanced panel datasets for the two separate countries (Germany and Italy) spanning four years (2000-2003) over fifteen industries.

[2] HHI remains fairly stable over time, however there are occasional changes when new firms enter, or existing firms exit a market.

3.1. Summary Statistics

Tables S1 and S2 in the appendix provide the summary statistics for both Germany and Italy on the variables used in the two models. The most notable difference between the countries is the values for *AVEHHI*, our main explanatory variable and measure for concentration. The minimum value for Italy is 1, which represents a perfectly competitive industry. In reality this is very unlikely, but the values are only estimates. It may not be that a value of 1 represents a perfectly competitive market, but instead one that is particularly local. The maximum value is 2998 which represents quite a highly concentrated industry so there is good variation in the industries for Italy, with a standard deviation of 384.99 to support that point. The mean value is 88.17, which is far higher than for Germany – 53.63. It seems that the industries in Germany for which data was possible to obtain were all fairly competitive, as the maximum value for *AVEHHI* is 254; the minimum value is again 1. Similar differences are present with the two measures of R&D, with much higher mean values for both measures in the Italian data than in the German data.

3.2. Variables

This paper aims to assess how concentration can affect the level of innovation in an industry, so we need a variable to represent innovation, which I have chosen as R&D expenditure. This is the dependent variable in all models including in this paper. Some papers use a ratio of R&D expenditure/Sales, but the data was not available for this, and could be an extra step to take this testing further in the future. The dataset contains this in two measures – 'national currency-constant prices' and '2005 dollars-constant prices and PPPs'. The measure I use throughout the paper is the latter, which is from here on labelled *RDConstantPPP.* Purchasing Power Parity (PPP) takes into account the difference in purchasing power between different countries using different exchange rates. This unit of measurement was chosen because it allows comparisons to be drawn if the model were to be extended to the US or UK data by myself or someone else. The data for this variable contains a number of estimates, which may be a problem in drawing comparisons between Germany and Italy – more detail to follow in section 3.3. There also exist alternative measures of innovation, such as output measures (patents) and other input measures (R&D employment). The strengths and weaknesses of the various measures of innovation are set out in section 3.3.

This paper aims to assess how industry concentration affects innovation, so it is necessary to use a measure of concentration as an independent variable. The measure used in this paper is an OECD annual estimation of the Herfindahl-Hirschman Index (HHI). HHI is measured as the sum of the square of the market share of each firm competing in a market. I have used the OECD estimate of HHI, labelled AVE_HHI in the 'Structural and Demographic Business Statistics 2006'. There are alternative measures of concentration used in other research, such as concentration ratios; these are briefly evaluated in section 3.3.

The models outlined in section 4 all contain controls for two variables – wages and employment, and model (3) contains an extra control for imports.

The variable *WAGE* represents the wages and salaries in each of the 15 industries and both countries at current prices, in the national currency (Euro). The level of wages in an industry may well have an effect on the R&D expenditure in an industry, through the following mechanisms: Firstly, an industry with a high wage rate for its manufacturing or innovative staff may also be an industry that is fairly profitable, and so can afford to pay its workers higher wages – or attract the most talented innovators, as argued in

Kamien & Schwartz (1982). Secondly and conversely, a firm that spends heavily on the wages of its employees may have less available to spend on R&D expenditure. These two counteracting effects make it unclear of the sign expected for the variable *WAGE* in the regressions.

EMPN represents the number of persons engaged in the industry, or total employment. This is measured purely in number of people, and so takes integer values. I would expect the coefficient on *EMPN* to be negative, displaying a negative correlation between total employment and R&D expenditure. The justification for this is that I would expect industries that employ more people to be fairly labour intensive (perhaps construction), whereas industries that employ fewer people may need to spend more on R&D as they are more capital intensive and a lot of their manufacturing is carried out by machinery.

A further control is brought into the later model for imports. The variable here *IMPO* is the value of imported goods into the country per industry each year. This is again given in current prices of the national currency (Euro). The reason imports are included is that various industries may have a stronger import component than others. For example, the German market for 'wearing apparel' may be made up of mainly imported goods, and so the manufacturing industry in German 'wearing apparel' may seem fairly concentrated, but in reality there exist many competitors from abroad. There may only be a few firms that manufacture clothes in Germany, but their market is still highly diluted as German consumers purchase from international manufacturers too. The coefficient on *IMPO* should be negative for low levels of HHI and positive for high levels of HHI, meaning the overall coefficient sign is unclear.

3.3. Data limitations

Referring back to the summary statistics, we had vastly differing levels of *AVEHHI* and both measures of R&D between the two countries. One possible explanation for this comes from the documentation page for the OECD databases used in this paper. The Business Enterprise Research and Development (BERD) datasets have guidelines for countries to follow in collation of data, but these are not always followed accurately and consistently. For some countries in the database, estimations may have been made, and for some countries guidelines may have been followed more strictly. Another factor could be the differences in the way countries treat the R&D undertaken by firms. An example given in the documentation is when a firm operating in both the manufacturing and the service sector spend considerably on R&D; this may be attributed to the service sector even though in some cases a considerable proportion of this expenditure is manufacturing R&D.

There is scope for debate as to whether R&D expenditure is the best measure of innovation, or whether an output measure of innovation is more appropriate – as used in other papers including Aghion et al (2005). A possible output measure of innovation is patents, or patents weighted by citations. A potential drawback in using patents as the measure of innovation in industry-level analysis is that patent systems vary between industries. Also, with a high degree of monopoly power, it may not be necessary to protect an innovation, due to the lack of rivals capable of imitating it.

Problems also exist with the key explanatory variable in this paper – HHI. Estimation problems aside, HHI may not be a truly accurate measure of concentration and alternatives do exist. As this paper attempts to test the results present in Aghion et al (2005), their measure is a natural point of comparison. Rather than concentration, Aghion et al refer to a relationship between 'product market competition' and innovation. This is measured by the Lerner index, which is given by the price-cost margin over the price. This is a measure of market power, which is also used in the replacement effect analysis in Arrow (1962). The analysis in Aghion et al (2005) is on a firm-level which suits the use of the Lerner Index, however I feel that with the industry-level analysis conducted in this paper, HHI or concentration ratios are a more suitable measure, because they account for the concentration of the industry not the firm individually and industries is the panel in this dataset, not firms. HHI accounts specifically for whether there exists a firm with a large amount of market power in the industry, and so is different to the Lerner Index, which looks at the price-cost margin of each firm in the industry to assess its market power. Concentration ratios are another alternative measure, which attempt to explain the market power of the largest few firms in an industry. I would have liked to test again using concentration ratios as a robustness check, following the same approach. Unfortunately, data on concentration ratios was not available.

HHI is not a good measure of competitiveness if the industry follows Bertrand competition; in this instance Lerner Index is better. However, as the Cournot model is generally fairly well accepted as a model in a lot of manufacturing, due to the existence of capacity constraints, HHI can be considered a reasonable proxy.

Specifically to my data, HHI levels differed between Germany and Italy. This may be partly due to the differing start up conditions for new firms between the countries. Alternatively, as HHI here is a national measure, some of the industries in the dataset may be more local than others. Referring back to the summary statistics, the minimum value of 1 may not truly be estimating a perfectly competitive market, but one that is particularly local. For this particular reason, model (3) in the methodology includes an imports component to account for international firms competing in the domestic market. This will only partly help the issue of HHI as a measure of concentration, and it remains not entirely accurate. As far as I am aware though, all existing measures of concentration do have their flaws.

The endogeneity of concentration is a major concern, as previously highlighted in both Aghion et al (2005) and Hashmi (2011). Successful innovations should increase the market share of the firm through improved products or production methods and hence increase the concentration in the industry. If this is the case, OLS estimates will be biased. Aghion et al used policy variables to control for endogeneity of competition in their model, whilst Hashmi used tariff rates and freight rates on imports. Hashmi finds remarkably different results in his US data when controlling for the endogeneity of competition compared with when he does not – the relationship changes from negative to positive. Interestingly, he concludes that the same problem does not seem to occur in his UK data. In this paper, lags are used to partially address this problem, as in Polder & Veldhuizen (2012).

4. Methodology

Upon examining the literature, this paper aims to test the relationship between industry concentration and innovation. An estimate for the Herfindal-Hirschman Index (HHI) is used as a measure for industry concentration industry, while research and development expenditure (R&D) is used as the measure for innovation, thus the dependent variable in all models.

I examine the possibility of a positive linear relationship, which will provide evidence supporting Schumpeterian hypotheses, which most empirical work before mine have not discovered. Failure to find such a positive linear relationship will further encourage questioning of Schumpeter Mark II, in favour of alternative hypotheses. If a negative linear relationship is found from my model, this will support the work of Acs & Audretsch (1988).

A non-linear model will then be introduced, testing for a possible inverted-U relationship as suggested in Aghion et al (2005). Results will be of interest due to the difference in measures used between the papers. Where Aghion et al use patents as their measure of innovation, I use R&D, and where the Lerner index is used as the main explanatory variable for concentration, the models discussed in this paper use the Herfindahl Index.

4.1. Model Framework

I introduce two separate models in this section, both of which are estimated for Germany and Italy in turn. In sub-section 4.2.1, a linear model is introduced, attempting to estimate a linear relationship between R&D expenditure and industry concentration, while controlling for the level of wages in the industry and the number of persons engaged – total employment. The second model attempts to test for the inverted-U relationship as proposed in Aghion et al (2005) by introducing a squared term on AVEHHI, similar to the approach in Polder & Veldhuizen (2012).

4.1.1. Linear Model

To test for a linear relationship between R&D expenditure and industry concentration, I firstly estimate the following equation using pooled OLS:

(1) $RDConstantPPP = \beta_0 + \beta_1 \text{ AVEHHI} + \beta_2 \text{ WAGE} + \beta_3 \text{ EMPN} + u$

If $\beta_1 < 0$ or the coefficient is statistically significant, this will provide evidence supporting Schumpeterian hypotheses that monopoly power drives innovation. If $\beta_1 > 0$, or if β_1 is statistically insignificant, my data will support most of the previous empirical work as seen in my literature review, with further evidence against the traditional view of a positive relationship between concentration and innovation.

4.1.2. <u>Inverted-U Relationship: Pooled OLS</u>

To test for a non-linear relationship, and the possibility of an inverted-U relationship such as that found in Aghion et al (2005), I estimate the following equation using pooled OLS:

$$(2)\ RDConstantPPP = \beta_0 + \beta_1\ AVEHHI + \beta_2\ HHIsquared + \beta_3\ WAGE + \beta_4\ EMPN + u$$

I follow a similar approach to that of Polder & Veldhuizen (2012) by introducing a quadratic term on my measure of concentration, although they measured competition. The addition of the quadratic term allows us to test for the inverted-U relationship. If the relationship exists, we should see $\beta_2 < 0$. The intercept of the inverted-U will be determined by the control variables. As HHI is always positive, we should also see $\beta_1 > 0$, so that our turning point is a maximum point and thus we have an inverted-U shape relationship. This is because a negative coefficient on *AVEHHI* would suggest

The model will then be extended further to try to account for the problems with HHI mentioned at the end of section 3.3. Including a term for imports or exports may help to control for differences between Germany and Italy in export and import penetration. Unfortunately, data is not available on the imports or exports for Germany, so the following control is only used in analysis of the data for Italy. Model (3) represents this extended control and will be estimated using the Italian data – but not the German data.

$$(3)\ RDConstantPPP = \beta_0 + \beta_1\ AVEHHI + \beta_2\ HHIsquared + \beta_3\ WAGE + \beta_4\ EMPN + \beta_5\ IMPO + u$$

4.1.3. Inverted-U Relationship: Fixed Effects

A Hausman test reveals that a fixed effects regression is more suitable than a random effects regression for Germany in regressing *RDConstantPPP* on *AVEHHI, HHIsquared, WAGE* and *EMPN*. This remains true when *IMPO* is added. The two fixed effects models, using the same variables as with pooled OLS, are given below as model (4) and (5) respectively.

$$(4)\ RDConstantPPP_{it} = \beta_1\ AVEHHI_{it} + \beta_2\ HHIsquared_{it} + \beta_3\ WAGE_{it} + \beta_4\ EMPN_{it} + \alpha_i + u_{it}$$

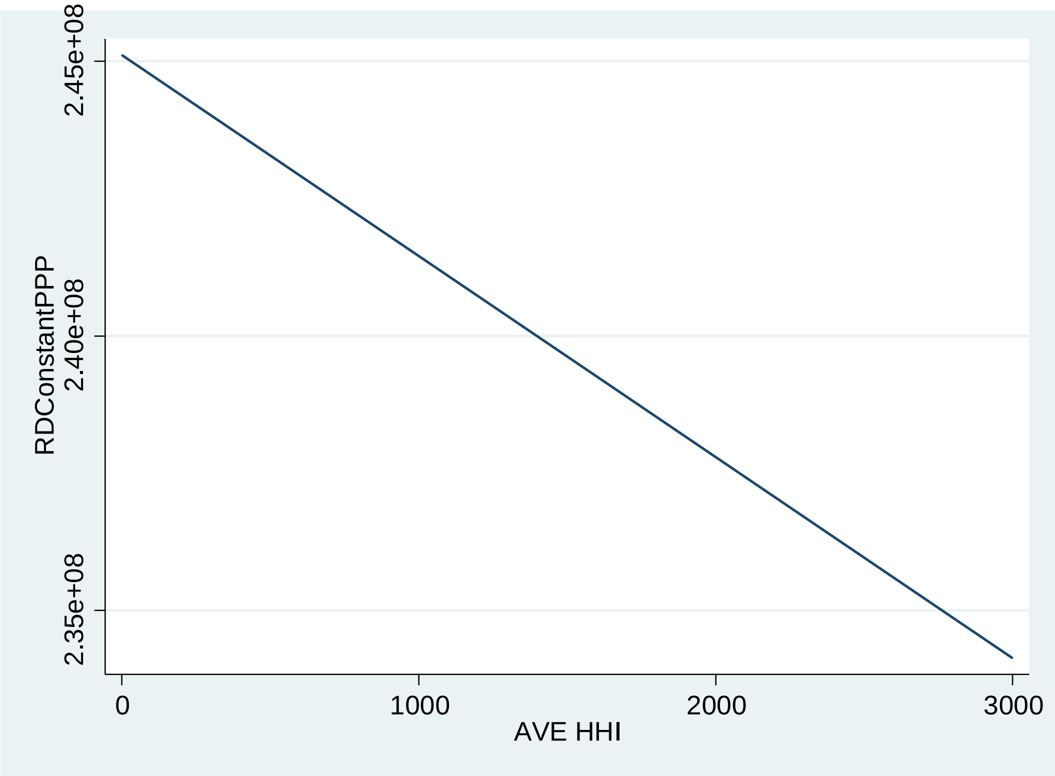
$$(5)\ RDConstantPPP = \beta_1\ AVEHHI + \beta_2\ HHIsquared + \beta_3\ WAGE + \beta_4\ EMPN + \beta_5\ IMPO + \alpha_i + u_{it}$$

The fixed-effects model allows us to study the causes of changes within each industry. By controlling for time-invariant differences between the industries, the coefficients of the fixed-effects model cannot be biased because they omit these characteristics.

5. Results

5.1. Results of the Linear Regression Model

The fitted linear plot below is for the Italian data with *RDConstantPPP* on the y-axis and *AVEHHI* on the x-axis suggests a negative linear relationship between industry concentration and R&D expenditure.



However, the correlation coefficient between the two variables is -0.0041, suggesting the relationship is very weak, if it truly exists. For Germany, the result is similar (Graph 1b, appendix), however, as discussed in section 3.1, the mean of *AVEHHI* is lower in the German data, and this is visible in the x-axis of the two way fitted plot.

This weak or possibly non-existent relationship is supported in the regression results when we estimate model (1) using OLS, shown in Figure 1.

Figure 1: Regression Results for model (1), Italy and Germany, with and without lags

	(1a) RDConstantPPP Italy	(1b) RDConstantPPP Italy	(1c) RDConstantPPP Germany	(1d) RDConstantPPP Germany
AVEHHI	40185.0		6611568.0	
	(102556.7)		(4704286)	
WAGE	0.113***	0.117***	0.467***	0.458***
	(0.0253881)	(0.0292063)	(0.0505151)	(0.0569185)
EMPN	-1092.5***	-1148.7***	-8864.7***	-8752.9***
	(242.1284)	(283.9159)	(988.5878)	(1107.125)
L.AVEHHI [1]		16380.0		6185246.6
		(101462.6)		(5038785)
_cons	59909132.8	53955862.9	-746708302.1	-739571514.5
	(6.51e+07)	(7.46e+07)	(5.14e+08)	(5.71e+08)
N	60	45	56	42
R^2	0.267	0.287	0.623	0.632
Adjusted R^2	0.228	0.235	0.602	0.603

standard errors in parentheses
* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

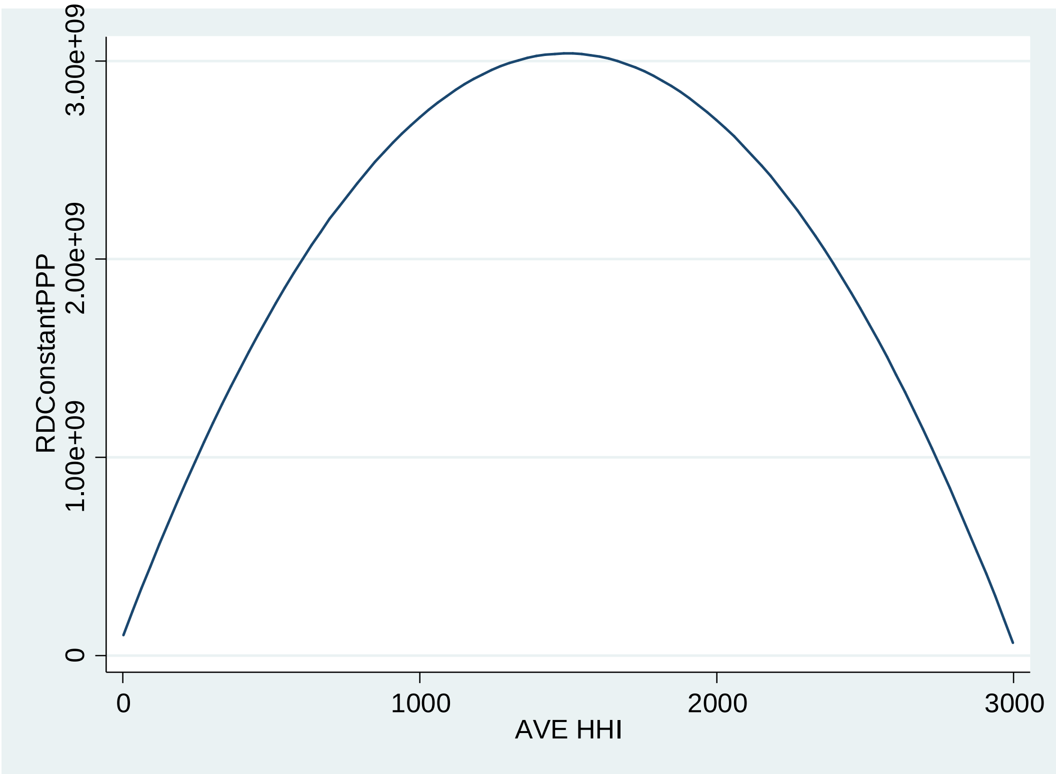
[1] L.AVEHHI = the lag of AVEHHI by one year

Without taking the lags into consideration, *AVEHHI* is statistically insignificant for both countries. This provides evidence towards the theory that the relationship between R&D expenditure and industry concentration cannot be explained linearly. The adjusted R-squared for the Italian results suggest that 22.8% of the variation in *RDConstantPPP* is explained by the model, whereas the German results are very different with an adjusted R-squared of 60.2%.

When a lag term is included on *AVEHHI*, which goes someway to addressing the endogeneity problem, the results remain consistent – making them more robust. Of course HHI does not vary too much over time, only to a small extent, so lags may or may not be the best or most appropriate way to address this concern.

5.2. Evidence of the Inverted-U Relationship

The graph below is a fitted quadratic plot for the Italian data of *RDConstantPPP* and *AVEHHI*. The German equivalent can be found in Graph 2 of the appendix, showing a similar result, but with a higher intercept on the y-axis.



The evidence here strengthens the arguments proposed in Aghion et al (2005) and the difference in choice of measures for innovation and industry concentration (or competition) makes the evidence more robust. Aghion et al graph an inverted-U relationship between product market competition and weighted number of patents, two completely different measures, but predict a similar relationship. The turning point in the graph above (Italy) is seen at around HHI = 1500. According to The U.S. Department of Justice this figure would represent a moderately concentrated marketplace. If we believe the evidence seen above, a suggestion could be that innovation is maximised in 'moderately concentrated' markets. I estimate the maximum point of the inverted-U my model suggests later in the section. However, the graph for Germany (Graph 2b, appendix) estimates the turning point at just above 100, which is a fairly competitive marketplace, although it still follows a similar shape.

5.2.1. Pooled OLS Results

Model (2)

Figure 2 shows the regression results when we estimate model (2) using OLS.

Figure 2: Regression results for model (2), Italy and Germany, with and without lags

	(2a) RDConstantPPP Italy	(2b) RDConstantPPP Italy	(2c) RDConstantPPP Germany	(2d) RDConstantPPP Germany
AVEHHI	4960267.6***		65493077.4***	
	(662251.7)		(1.46e+07)	
HHIsquared	-1633.9***		-250008.3***	
	(218.5891)		(59539.69)	
WAGE	0.133***	0.137***	0.490***	0.481***
	(0.0182368)	(0.0196732)	(0.0442958)	(0.0505257)
EMPN	-1192.3***	-1244.6***	-8925.1***	-8884.3***
	(172.5968)	(189.9841)	(860.6272)	(974.9063)
L.AVEHHI[1]		5159893.3***		58620269.6***
		(716381.3)		(1.57e+07)
L.HHIsquared[2]		-1703.2***		-219148.0**
		(236.1615)		(63052.54)
_cons	-202308633.0***	-210353291.0**	-2.51344e+09***	-2.32341e+09**
	(5.81e+07)	(6.18e+07)	(6.14e+08)	(6.78e+08)
N	60	45	56	42
R^2	0.636	0.690	0.720	0.722
Adjusted R^2	0.610	0.659	0.698	0.692

standard errors in parentheses
* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

[1] L.AVEHHI = the lag of AVEHHI by one year
[2] L.HHIsquared = the lag of HHIsquared by one year

With the addition of the squared term on *AVEHHI*, results look much more promising. *AVEHHI* is now significant at the 99.9% level, so too is the new variable *HHIsquared*. Adjusted R squared has increased from model (1) to model (2) in both countries, most notably for Italy where the value was previously 22.8% (without lags) and is now 61%. This change suggests that the variation in R&D expenditure can be explained better in model (2) where we have a non-linear effect of concentration.

Evidence of the inverted-U relationship can be found upon interpretation of the coefficients on *AVEHHI* and *HHIsquared*. As proposed in section 4.2.2, a positive coefficient on *AVEHHI* and a negative coefficient on *HHIsquared* are required to provide evidence supporting the inverted-U hypothesis. This is exactly what is found in the data. For Italy, the coefficient on *AVEHHI* is 4960267.6 and the coefficient on *HHIsquared* is -1633.9 which makes them economically significant; both are statistically significant at the 99.9% level. For Germany, results remain consistent with the inverted-U. The coefficient on *AVEHHI* is 65493077.4 and the coefficient on *HHIsquared* is -250008.3; both coefficients are again statistically significant at the 99.9% level and the signs of the coefficients are in line with the theory.

When lag terms are introduced on the two concentration variables, results remain strong, thus making them more robust. For Italy the coefficients on *AVEHHI* and *HHIsquared* with lags 5159893.3 and -1703.2 respectively, remaining significant at the 99.9% level. For Germany, the coefficients on *AVEHHI* and *HHIsquared* with lags are 58620269.6 and -219148.0 respectively, and there is a slight fall in significance on the *HHIsquared* coefficient, but it is still significant at the 99% level.

The true length of the lag term on concentration is unclear, and I am unaware of definitive justification for one specific length. The fact this dataset is 4 years long only allows for regressions to be conducted with lags up to 3 years. Furthermore, with a relatively small number of industries available for analysis, extending the lag term further than one year makes results less reliable, as the number of observations becomes very small. Nevertheless, such regressions were carried out and once again we saw some differences between the two countries. For Germany, extending the lag term on *AVEHHI* and *HHIsquared* did not alter results much. The signs of the respective coefficients were still consistent with the inverted-U, and with a two year lag the coefficients on the two variables are still significant at the 99% level. With a three year lag, they are only significant at the 90% level, however as the number of observations is then only 15, these results are less reliable. For Italy, the results are sensitive to an extension of the lag term. The coefficients on *AVEHHI* and *HHIsquared* become statistically insignificant and their signs vary, meaning they are no longer consistent with the inverted-U, so the results may be sensitive to lags for Italy.

Model (3)

With data only available on imports for Italy, but not Germany, model (3) was estimated using both pooled OLS and fixed effects estimation methods. The pooled OLS results are shown below.

	RDConstantPPP Italy
AVEHHI	5375854.7***
	(808720.3)
HHIsquared	-1766.5***
	(262.8294)
WAGE	0.0971**
	(0.0312022)
EMPN	-3.101
	(696.039)
IMPO	0.00696
	(0.0036745)
_cons	-397033850.7***
	(9.86e+07)
N	52
R^2	0.736
Adjusted R^2	0.7075

Standard errors in parentheses * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

First of all, adjusted R squared has increased to 70.75%, suggesting that imports add to the explanatory power of the model. The coefficients on *AVEHHI* and *HHIsquared* remain significant at the 99.9% level, with the added control for imports. They also follow the same economic justification required in the inverted-U theory, with a positive coefficient on *AVEHHI* and a negative coefficient on *HHIsquared*. This additional variable accounts for foreign manufacturers competing in the Italian goods markets. Despite the important coefficients remaining significant, the coefficient on WAGE sees a small decrease, and EMPN has become statistically insignificant. Imports itself is significant at the 90% level, but not at the 95% level, showing it may have some effect on R&D expenditure, therefore suggesting it may be a worthwhile control.

5.2.2. Heteroskedasticity and Robustness

These results provide some evidence supporting the inverted-U theory; however the regression model that produces these results may be subject to heteroskedasticity. For OLS results to be unbiased, multiple linear regressions (MLR) contain assumptions. MLR5 states that homoskedasticity is present, so that the error term u has the same variance given any values of the explanatory variables:

$$Var(u|x_1, x_2, \dots, x_k) = \sigma^2 \text{ , with } k \text{ explanatory variables}$$

The residual versus fitted plots in graph 3 and 4 of the appendix show some initial suggestion of heteroskedasticity. As the plots move toward the top right corner, there exist a few points that are not in line with the general downward sloping pattern – this could suggest heteroskedasticity. Upon performing a Breusch-Pagan test it seemed there was clear heteroskedasticity present in the model for both countries. To account for this problem, the regressions were run once more, but this time I accounted for heteroskedasticity using the robust regression technique; results are seen below in Figure 2.

Figure 2: Robust Regression for model (2), Italy and Germany, with and without lags

	(2e) RDConstantPPP Italy	(2f) RDConstantPPP Italy	(2g) RDConstantPPP Germany	(2h) RDConstantPPP Germany
AVEHHI	4960267.6***		65493077.4***	
	(794492.8)		(1.54e+07)	
HHIsquared	-1633.9***		-250008.3***	
	(263.1512)		(59250.7)	
WAGE	0.133***	0.137***	0.490***	0.481***
	(0.160906)	(0.174032)	(0.823271)	(0.1004287)
EMPN	-1192.3***	-1244.6***	-8925.1***	-8884.3***
	(150.7157)	(162.3676)	(1542.172)	(1861.97)
L.AVEHHI		5159893.3***		58620269.6**
		(617583.3)		(1.69e+07)
L.HHIsquared		-1703.2***		-219148.0**
		(204.4649)		(63061.79)
_cons	-202308633.0***	-210353291.0***	-2.51344e+09***	-2.32341e+09**
	(3.88e+07)	(4.36e+07)	(6.37e+08)	(7.38e+08)
N	60	45	56	42
R^2	0.636	0.690	0.720	0.722

Robust standard errors in parentheses
* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Our results remain consistent after accounting for a minor heteroskedasticity problem. Evidence of an inverted-U relationship is still present, with a positive coefficient on *AVEHHI* and a negative coefficient on *HHIsquared*, whilst all explanatory variables in the model are statistically significant at the 99.9% level without lags. This time the introduction of lags sees a slight fall in significance on both *HHIsquared* and *AVEHHI*, but both are still statistically significant at the 99% level.

5.2.3. Fixed Effects

Table 1 in the appendix shows the results of the regressions on models (4) and (5). Models (4) and (5) are in effect variants of models (2) and (3), using the fixed effects estimation method. These results are far less convincing than the pooled OLS results, and really show no evidence of the inverted-U relationship that was estimated by Aghion et al (2005) and backed up by the pooled OLS results in this paper. Neither model (4) or (5) provides statistically significant coefficients on either of the concentration variables, failing to support the inverted-U theory or really show any relationship at all. A possible reason for this could be that with the existence of heterogeneity, which may be present in this paper, results can vary between estimation methods. Furthermore, the small number of time periods makes it even more likely that there is some heterogeneity between the estimators and the error term. Note that similarly to model (3), model (5) is only estimated using Italian data, due to the unavailability of imports data for Germany. The difference in results between estimation methods is a key concern, and further investigation into why this affects the results so substantially is a top priority moving forward in this work.

5.3. Estimating the Turning Point of the Inverted-U

Let us remind ourselves of model (2):

$$RDConstantPPP = \beta_0 + \beta_1 HHI - \beta_2 HHI^2 + \beta_3 \text{WAGE} - \beta_4 \text{EMPN} + \text{u}$$

Using the regression results from section 5.2.1 for model (2), where the signs of β_1 and β_2 are as predicted by the inverted-U model, it may be possible to predict a level of HHI for which R&D expenditure is maximised.

The estimated equation for Italy is shown below:

$RDConstantPPP$ = -202308633.0 + 4960267.6 HHI - 1633.9HHI^2 + 0.133 WAGE - 1192.3 EMPN + u

A turning point can be found by deriving *RDConstantPPP* with respect to *HHI*. The first order conditions are:

$$\frac{\partial\ RDConstantPPP}{\partial HHI} = \mathbf{4960267.6 - 2(1633.9)}\boldsymbol{HHI} = \mathbf{0}$$

This gives an *HHI** value of 1517.922639.

To ensure this is indeed a maximum point as it should be in an inverted-U shape, we take the second order derivative:

$$\frac{\partial^2\ RDConstantPPP}{\partial HHI^2} = \mathbf{-3267.8 < 0}$$

Thus, the Italian data and the model predict that an *HHI* level of 1517.92 will maximise R&D expenditure. Referring back to Graph 1a, the level of *HHI* predicted as the turning point from the two way quadratic fitted plot was indeed at around 1500, so this is consistent.

Similarly for Germany:

$RDConstantPPP$ = -2.51344e+09 + 65493077.4 HHI - 250008.3 HHI^2 + 0.490 WAGE - 8925.1 EMPN + u

$$\frac{\partial\ RDConstantPPP}{\partial HHI} = \mathbf{65493077.4 - 2(250008.3)}\boldsymbol{HHI} = \mathbf{0}$$

This gives an *HHI** value of 130.9818062.

To ensure this is indeed a maximum point as it should be in an inverted-U shape, we take the second order derivative:

$$\frac{\boldsymbol{\partial}^{\mathbf{2}}\ RDConstantPPP}{\boldsymbol{\partial} HHI^{2}} = \mathbf{-500016.6} < \mathbf{0}$$

Thus, the German data and the model predict that an *HHI* level of 130.98 will maximise R&D expenditure. Graph 2b in the appendix predicts the maximising level of HHI to be just over 100, so our value of 130.98 falls slightly higher than that. The issue with this result is that the German data only contains industries with HHI levels of 254 and lower, so all industries for which we have data from Germany are quite highly competitive. Therefore, the results may not be as reliable for the German data as they are for Italy, where the data contained a wider range of industry concentration. This difference also suggests that the results may be very specific to the country used. As a result, no single policy response can really be recommended in finding the 'sweet spot', as its position varies. However, I do believe that a larger sample size for both countries may bring the results closer together.

6. Conclusion

This paper examines how industry concentration affects the rate of technological change – innovation, using R&D expenditure as the measure of innovation. The inverted-U relationship as proposed in Aghion et al (2005) is tested, looking for a result where neither perfectly competitive markets nor monopolists innovate much, but the moderately concentrated markets are most progressive.

Using a panel of 15 industries over 4 years in both Germany and Italy, with HHI as the measure of concentration, regressions were carried out to see if the same result was found in this paper. Using countries untested in the literature I have read, we have a unique dataset. The different measures used for both innovation and concentration compared with Aghion et al (2005) also makes for a unique and interesting setting.

Following the approach outlined in Polder & Veldhuizen (2012), a squared term for concentration was added to a model which is built to test solely for concentration effects, whilst controlling for total employment, wages and imports in the industries.

Evidence of the inverted-U relationship is found using pooled-OLS analysis. The estimated location of the maximum turning point on the inverted-U differed between the two countries, with HHI values of around 130 for Germany and 1518 for Italy. This difference was to be expected after the initial review of the dataset; the two countries had differing summary statistics. Unfortunately, there was a lack of highly concentrated industries in either panel, which biased the estimated turning points towards low values, but should not have affected the resulting shape.

These results were subject to two main forms of robustness checks for the two main issues faced in this area of work – heteroskedasticity and the endogeneity of concentration. There were signs of heteroskedasticity when running a 'het test' and when plotting the residuals versus the fitted. To account for these differences robust regressions were conducted and results remained consistent. The endogeneity concern comes from a two-way causality problem that successful innovations can lead to market power. To address this, Polder & Veldhuizen's technique of adding a lag term on concentration variables was used, and results again remained consistent with the inverted-U. However, the results are sensitive to the pooled OLS estimation method used, as they do not hold under the fixed effects approach tested in this paper. Further investigation is required to find out why this is, and additional work using a fixed effects model would be very interesting.

Further testing is needed before the inverted-U theory can be considered a definite model. At low levels of concentration it does seem there is an increase in innovation as industries become less competitive, perhaps evidence supporting the so-called "Schumpeterian effect" (Aghion et al, 2005). High levels of concentration were not tested so well in this paper, so further research with more representation of the highly concentrated industries is required before firm conclusions can be drawn. Tests would also need to be carried out on other countries. For further robustness and if data allowed, this paper would be improved with a second measure of concentration – concentration ratios, for example – but this was not possible.

7. Bibliography

- **Acs, Zoltan J. and David B. Audretsch.** *"Testing the Schumpeterian Hypothesis".*

  Eastern Economic Journal, Volume XIV, No. 2 (1988)

- **Aghion, Philippe., Nick Bloom, Richard Blundell, Rachel Griffith and Peter Howitt.** *"Competition and Innovation: An Inverted-U Relationship".*

  May Edition, The Quarterly Journal of Economics (2005)

- **Arrow, Kenneth.** *"Economic Welfare and the Allocation of Resources for Invention".*

  'The Rate and Direction of Inventive Activity: Economic and Social Factors', Princeton University Press (1962)

- **Blundell, Richard.**, **Rachel Griffith and John Van Reenen.** *"Market Share, Market Value and Innovation in a Panel of British Manufacturing Firms"*

  Review of Economic Studies, 66, 529-554 (1999)

- **Corsino, Marco., Giuseppe Espa and Rocco Micciolo.** *"R&D, Firm size, and Product Innovation Dynamics"*

  DISA Working Paper, Department of Computer and Management Sciences, University of Trento (2008)

- **Galbraith, John Kenneth.** *"American Capitalism: The Concept of Countervailing Power".*

  Boston: Houghton Mifflin (1952)

- **Haruyama, Tetsugen.** *"An Inverted U Relationship Between Competition and Innovation: A Revisit"*

  Graduate School of Economics, Kobe University (2006)

- **Hashmi, Aamir Rafique.** "*Competition and Innovation: The Inverted-U Relationship Revisited"*

  Department of Economics, National University of Singapore (2011)

- **Kamien, Morton I. and Nancy L. Schwartz.** *"Market Structure and Innovation"*

  Cambridge University Press (1982)

- **Levin, R.C., W.M Cohen, and D.C Mowery.** *"R & D Appropriability, Opportunity, and Market Structure: New Evidence on Some Schumpeterian Hypotheses"*

  The American Economic Review, Vol. 75, No. 2 (1985)

- **Mansfield, Edwin.** *"Size of Firm, Market Structure and Innovation".*

  Page 556-576, Journal of Political Economy, Vol. 71, no.6. (1963)

- **Polder, Michael., and Erik Veldhuizen.** *"Innovation and Competition in the Netherlands: Testing the Inverted U for Industries and Firms"*

  Journal of Industry, Competition and Trade, 12(1), 67-91 (2012)

- **Scherer, Frederic Michael.** *"Firm Size, Market Structure, Opportunity and the Output of Patented Inventions".*

  American Economic Review, 55, pp. 256-266 (1965)

- **Schumpeter, Joseph A.** *"The Theory of Economic Development"*

  Cambridge, Massachusetts: Harvard University Press (1934)

- **Schumpeter, Joseph A.** *"The Creative Response in Economic History"*

  Journal of Economic History 7(2), pp. 149-159 (1947)

- **Scott, J.T.** *"Firm Versus Industry Variability in R&D Intensity"*

  Z. Grilches (ed.), R&D, Patents, and Productivity, University of Chicago Press, Chicago (1984)

- **Shrieves, Robert.** *"Market Structure and Innovation: A New Perspective".*

  Journal of Industrial Economics, 26, pp. 329-347. (1978)

Data from OECD

- STAN database for structural analysis (ISIC Rev.4) – www.stats.oecd.org

- STAN R&D expenditures in industry (ISIC Rev.4) – www.stats.oecd.org

- 'Structural and Demographic Business Statistics 2006' available online at www.sourceoecd.org

8. Appendix

Summary Statistics

Table S1: Summary Statistics for Germany

Variable	Observations	Mean	Standard Deviation	Minimum	Maximum
RDConstantPPP	60	1.75e+09	3.51e+09	6073281	1.44e+10
AVEHHI	56	53.625	64.29635	1	254
HHIsquared	56	6935.89	15104.21	1	64516
WAGE	60	2.04e+10	2.91e+10	5.36e+08	1.12e+11
EMPN	60	842733.3	1504948	27378	5928000
IMPO	0	N/A	N/A	N/A	N/A

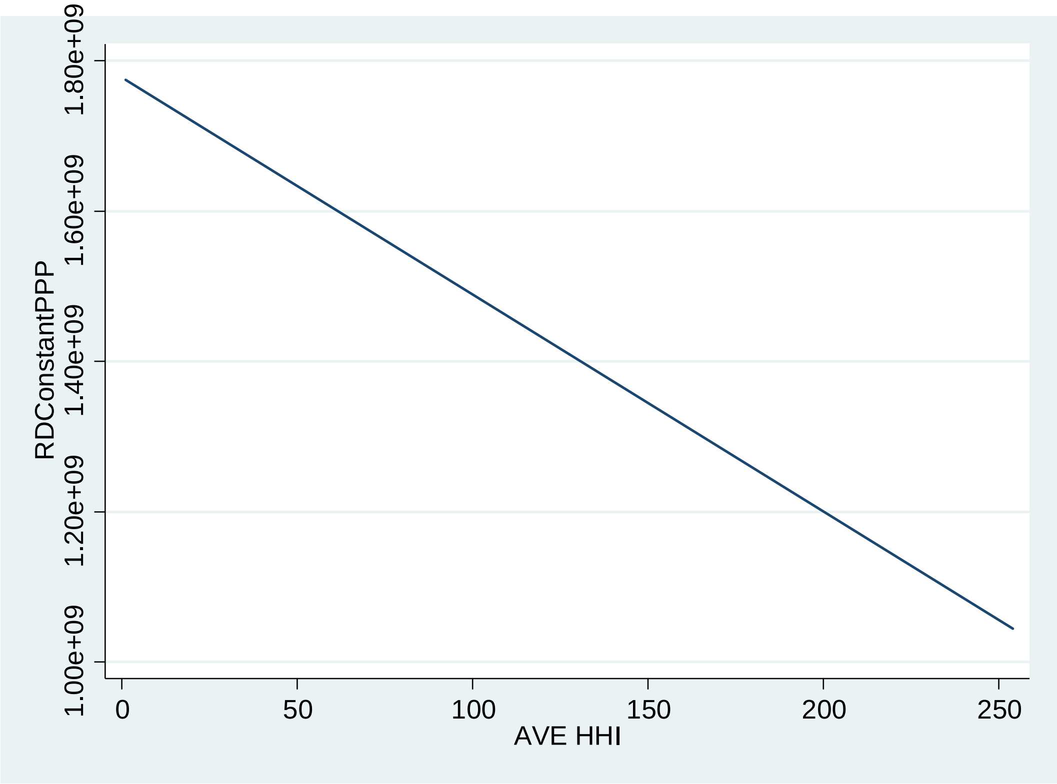
Table S2: Summary Statistics for Italy

Variable	Observations	Mean	Standard Deviation	Minimum	Maximum
RDConstantPPP	60	2.45e+08	3.41e+08	265282	1.08e+09
AVEHHI	60	88.16667	384.9885	1	2998
HHIsquared	60	153519.3	1159884	1	8988004
WAGE	60	6.59e+09	8.53e+09	8.09e+08	3.56e+10
EMPN	60	514906.7	892564.9	36800	3570600
IMPO	52	1.20e+10	9.48e+09	5.58e+07	3.31e+10

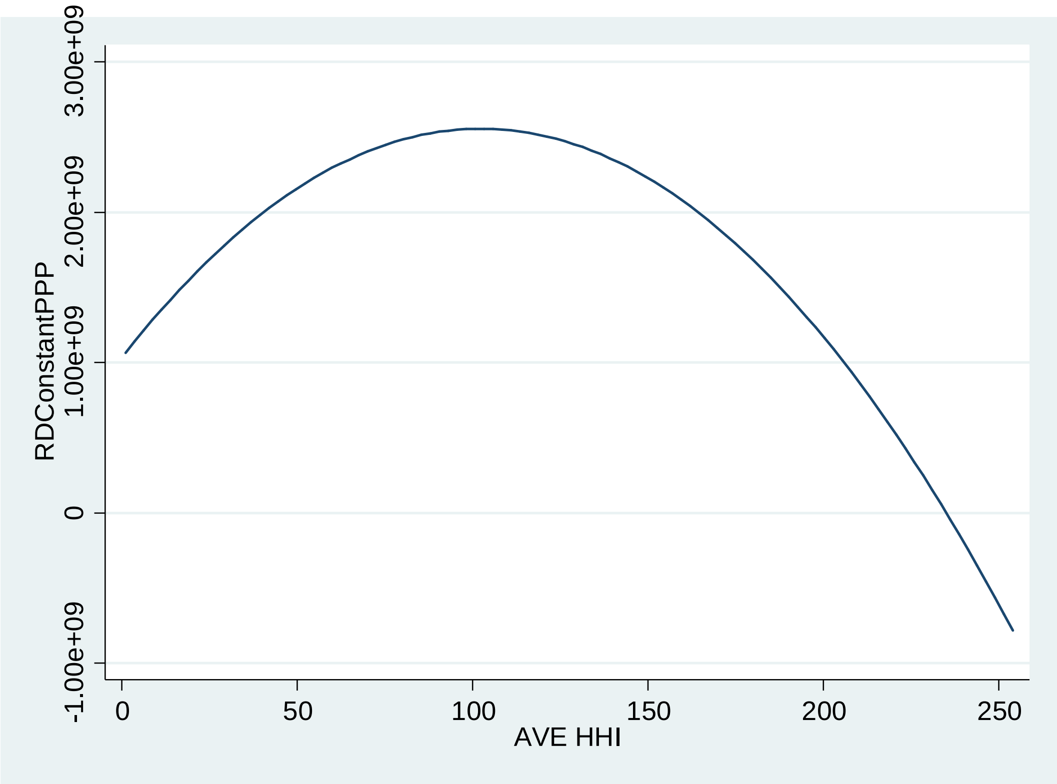
List of Variables

Variable	Description	Measure
RDConstantPPP	R&D expenditure by industry	2005 Constant Prices – US Dollars ($) – PPP
AVEHHI	HHI estimate from OECD	1-10000- Estimation
HHIsquared	The square of AVEHHI	Estimation
WAGE	Annual wages and salaries by industry	Euro (€)
EMPN	Total number of persons engaged in employment by industry	Integers – number of people
IMPO	Annual value of imported goods into the country by industry	Euro (€)

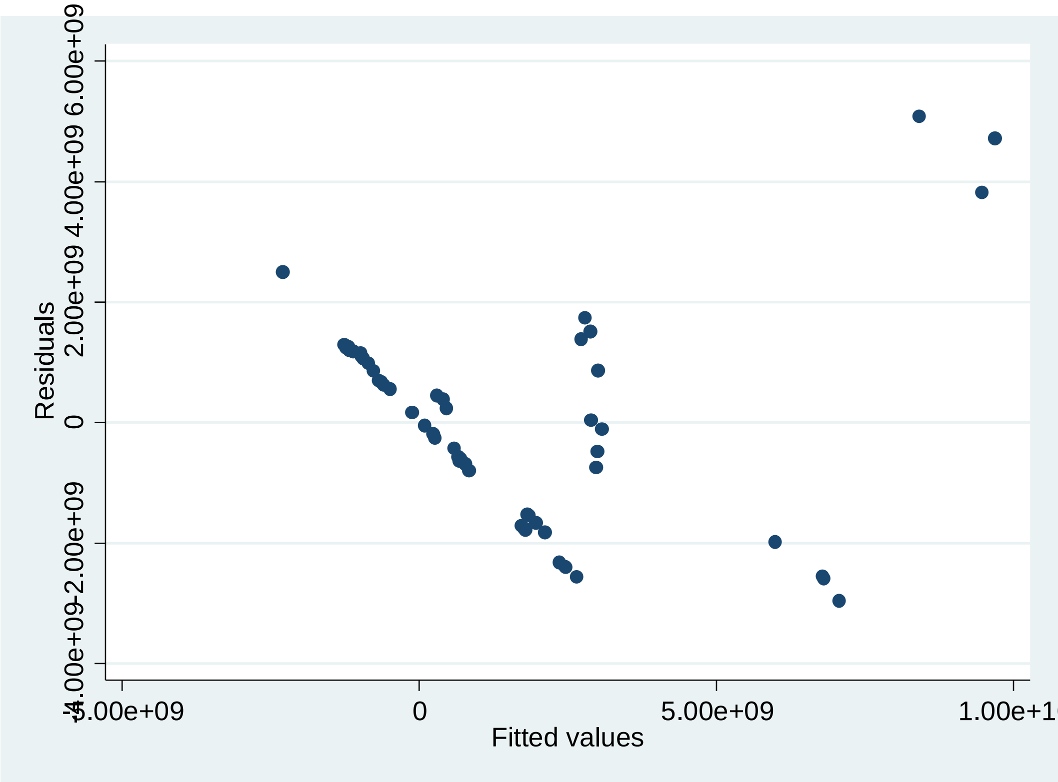
Graph 1b: Two way linear fitted plot of *RDConstantPPP* and *AVEHHI* for Germany



Graph 2b: Two way quadratic fitted plot of *RDConstantPPP* and *AVEHHI* for Italy



Graph 3: Germany: Residuals versus fitted plot



Graph 4: Italy: Residuals verus fitted plot

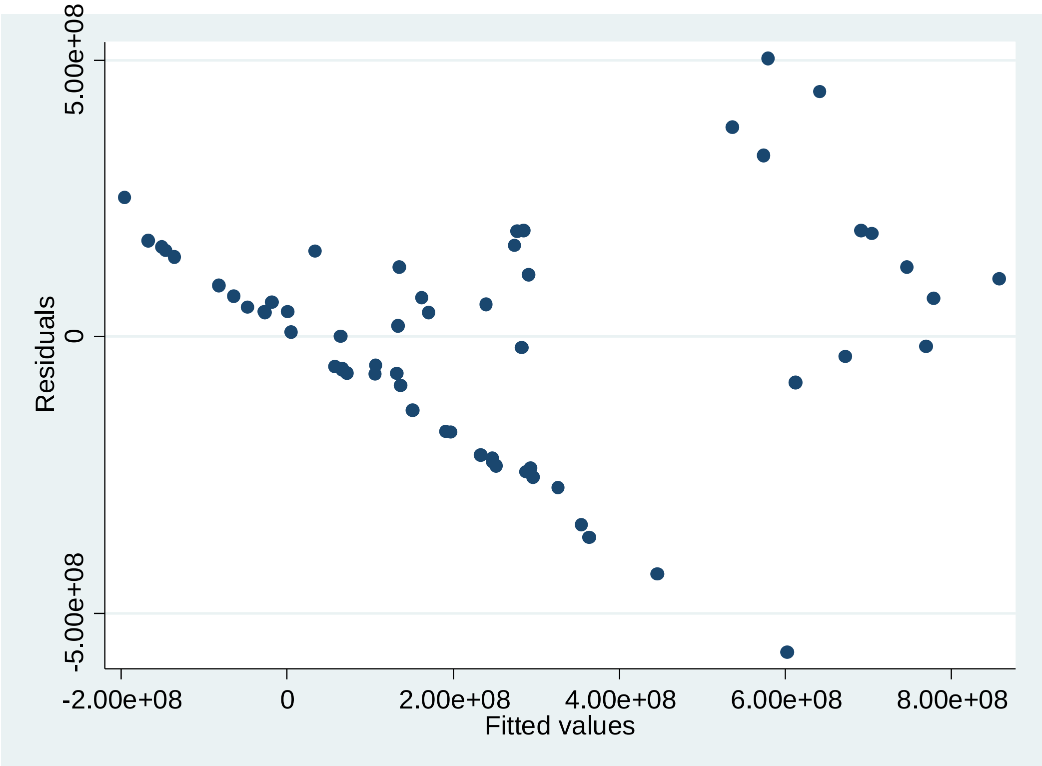


Table 1: Fixed Effects Results – Model (4) and (5)

	(4) RDConstantPPP Italy	(4) RDConstantPPP Germany	(5) RDConstantPPP Italy
AVEHHI	823580.0	-519178.5	112983.9
	(935346.9)	(6184838)	(897863.2)
HHIsquared	-260.5	4005.7	-36.54
	(300.7714)	(24448.82)	(288.1315)
WAGE	0.139***	0.0764	0.162*
	(0.032881)	(0.0571573)	(0.0609657)
EMPN	-2849.8**	-1458.4	2290.8
	(994.4948)	(1244.923)	(3239.307)
_cons	761897597.7**	1.30010e+09*	-618865010.8
	(2.64e+08)	(4.86e+08)	(0.011608)
IMPO			-0.0123
			(0.011608)
N	60	56	52
R^2	0.384	0.048	0.473

standard errors in parentheses * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$