Paper to be presented at

DRUID15, Rome, June 15-17, 2015

(Coorganized with LUISS)

# Cost spreading in the photovoltaic industry- Testing the Cohen and Klepper model

**Ann-Kathrin Blankenberg**
University of Kassel
Institute of Economics
ak.blankenberg@uni-kassel.de

Abstract

This article provides a new test of the Cohen & Klepper (1996) model of cost spreading which explains the relationship between innovation, firm size and R&D and leads to economies of scale in innovation at the example of the photovoltaic industry in Germany. First, it is analyzed whether the cost spreading mechanism serves as an explanation for size advantages in this industry. This is related to the assumption that the incentives to invest in R&D increase with the ex-ante output. Furthermore I investigate whether firms that plan to grow will have more innovative activities. The analysis is based on an unbalanced panel data set of German photovoltaic solar cell producers covering the time period 2000-2011. The results indicate that cost spreading serves as an explanation for size advantages in this industry and furthermore growth plans lead to higher innovative activities.

Jelcodes:O31,O38

## 1 Introduction

The discussion about the advantages and disadvantages of firm size has a long tradition in academic and political circles. The "Schumpeterian" hypotheses frame the discussion about the relationship between innovation, firm size and research and development (R&D). In the past decades, based on this, different empirical patterns have been observed. Some claim that large firms produce fewer innovations although they spend more on it. Others argue that there must be an advantage of size. The question of whether size is an advantage in competition has not been conclusively answered. Based on this research, Cohen and Klepper (1996a) argue that if the returns of R&D are conditioned by firm size, large firms would have a big advantage over small firms. Based on studies which analyze the relationship between firm size and R&D they summarize the robust findings into four stylized facts and create a model based on the cost spreading argument to explain these findings. Cost spreading means "[…] the larger the firm then the greater the level of output over which it can apply the fruits of its R&D and hence the greater its returns from R&D (Cohen and Klepper, p.926, 1996a)." The model offers the opportunity to explore the relationship between firm size, innovation and R&D and it is possible to show that large firms have an advantage of size which is attributable to cost spreading.

This article provides a new test of the Cohen and Klepper (1996a) model of cost spreading leading to economies of scale in innovation. The contribution of this article is twofold. First, it analyzes whether the cost spreading mechanism serves as an explanation for size advantages in the photovoltaic industry. This is related to the assumption that the incentives to invest in R&D increase with the ex-ante output. But we still lack information about the relationship between growth plans and innovative behavior. Planned data of firms has not been used in any study investigating the cost spreading mechanism to date. But, if cost spreading is a working mechanism in this industry, it might also be observable in the strategic decisions, in the planning data. So, in a second step I will analyze whether firms that plan to grow exhibit more innovative activity. This article is based on an unbalanced panel sample of photovoltaic solar cell producers in Germany, covering the time period 2000-2011 and also includes the planning data of the firms for the time period 2005-2011. The dynamics of the photovoltaic industry in Germany offer a good opportunity to test the cost-spreading model. This industry is a highly innovative industry and the total number of patents experienced a sharp increase following the implementation of demand-inducing policy instruments in recent years. The output of the firms also increased strongly. In such an evolving industry product and process innovations play an important role and there is still great potential for innovative activities. What is more, this industry has been highly supported by public policy. Thus, it should be examined whether the patterns differ from recent observed patterns. The results of this article indicate that the photovoltaic firms in Germany have an advantage of size which is attributable to cost spreading. Furthermore, the examination of the planned data indicates that firms which plan to grow will have more innovations.

The paper is organized as follows. Section 2 starts with a review of the literature. This is followed by a description of the cost spreading model of Cohen and Klepper (1996a) and a short outline of its testing in the literature (section 3). The section concludes with the presentation of the new approach to test the cost spreading argument and an outline of the hypotheses based on the theory. Section 4 describes the data, defines the variables used to analyze the innovation, firm size and growth plans relationship and gives an overview of the econometric models. Section 5 presents the results. The paper ends with a conclusion in section 6.

## 2 Overview about the innovation research

The research about innovation is to a great extent influenced by the writings of Schumpeter in the last century and started in the 1960s (Gilbert 2006). In his early work Schumpeter (1912/1934) assumed that the entrepreneur carries out the new combinations. This perspective is changed in his later work. Influenced by the changes in the economy in the early 20$^{th}$ century, Schumpeter (1942) argued that larger (monopolistic) firms are the source of innovation. Based on Schumpeter's writings, two hypotheses, labeled as the Schumpeterian hypotheses, were developed in the subsequent years and research was conducted to test them. The first hypothesis predicts a positive relationship between firm size and innovation. What is more, the second hypothesis assumes a positive relationship between market structure and innovation and indicates that market power supports innovative activities.

Starting with a short overview of the research related to the second Schumpeterian hypothesis shows that various empirical studies were conducted to analyze the relationship between market structure and R&D. Despite this, the results are heterogeneous. Already Markham (1965) points out that Schumpeter never claimed a continuous relationship between market power or firm size and R&D, just that the incentive to invest in risky projects does not exist up to a certain firm size. Acs and Audretsch (1987) e.g. show that large firms have an advantage in capital-intensive and concentrated industries whereas small firms draw a benefit from highly innovative industries. Gilbert (2006) argues that different types of market structure lead to different results. Thus, neglecting the influence of market structure on innovation would be problematic. An improved measure of the market structure can be found e.g. in Nickell (1996) or in Artés (2009). The latter argues that market structure effects only the long-run R&D decisions. Considering different types of markets Audretsch et al. (2014a) show that not all new firms spend money on R&D. They show that markets with a high level of uncertainty will reduce the probability of innovating or patenting for young and small firms. Cohen (2010) summarize that most studies indicate a positive relationship between R&D and market structure.

More research was conducted in the last 70 years related to the innovative activities of firms depending on their size, but the results are also heterogeneous. Scherer (1965a; b) e.g. shows that the inventive output (patents) increases with firm sales, but there is no proportional growth, hence he doubts that large firms are that perfect as Schumpeter once believed. This question was investigated by further authors like e.g. Villard (1958) who argues that innovation varies with size and size might be an advantage, or Schmookler (1959) and Worley (1961) who argue that there is not necessarily a general or systematic relationship between innovation and firm size. Contrary to that, Soete (1979) argues that there is neither clear evidence for a strong increase nor for a strong decrease in innovative activities based on size, hence to be big is not per se a disadvantage. Vernon and Gusen (1974) e.g. show that the elasticity of technological change with respect to the size of the firm increases with size. They conclude that larger firms have decisive advantages in innovation compared to smaller firms. In addition to the studies of the effects of firm size on firm level R&D there are studies about the relationship between firm size and R&D on business unit level. The results of Cohen et al. (1987) show that the performed R&D intensity is not affected by the size of the business unit, but the probability that R&D is conducted is mainly influenced by the size of the business unit rather than by the size of the whole firm. This is also shown in Cohen and Klepper (1996a).

Markham (1965) criticizes that the research of that time only analyzed the statistical relationship between some kind of Schumpeterian variables and the firm variables, which are at least only a measure of monopoly. Most of the studies are concentrated on R&D expenditures and patents in general. Cohen (2010) notes that the selection of the firms in the examined samples (only the large manufacturing firms; firms reporting no R&D are excluded), as well as the use of the control variables, should be seen critically in the early studies. Furthermore, Fisher & Temin (1973) argue that empirical tests of the Schumpeter hypothesis in the literature have little in common with the hypothesis itself. They point out that the only appropriate way to test the Schumpeterian hypothesis is to analyze the relationship between firm size and innovative output. In line with Fisher and Termin (1973) also Cohen and Levin (1989) argue that rather primitive econometric techniques were used to analyze data, which often was inadequate to study the questions and also the equations were only loosely specified. They further argue that the dependent variable should be a measure of innovative output rather than the innovative input (R&D expenditures, R&D staff) to the innovation process. Thus, the relationship between innovative output and firm size should be examined. The weaknesses of measuring innovative activities via input variables are also indicated in later articles (Kleinknecht 2000; Kemp et al. 2003).

The results of more recent studies are still heterogeneous. Malerba and Orsenigo (1996b) e.g. argue that across technological classes the innovative activities are systematically different. Analyzing the economies of scale and scope in the pharmaceutical industry, Henderson and Cockburn (1996) find that large firms in this industry benefit more from conducting research. Shefer and Frenkel (2005) show that in the plastics and metal industry size does not influence the rate of investment in R&D. In line with the criticism about the right choice of data, Stock et al. (2002) show that in a dynamic sense smaller firms are more technologically innovative. They point out that innovation is an ongoing process and thus the relationship between firm size and innovation should be investigated over time. What is more, recent studies by Hashi and Stojčić (2013) support the findings about the firm size and innovation relationship. Based on two settings (mature and transition economies) the drivers of the innovation process are examined. In both settings similar processes work as triggers of the innovation process, e.g. the characteristics of a firm, the institutional background and also industry specific factors. The results support earlier findings, namely that the probability to engage in innovation increases with firm size but the output of the innovation process decreases with size. Research related to the Schumpeterian hypothesis mostly examined the influence from firm size on the innovative activity. In the last years the focus of research shifted away from analyzing the pure role of firm size on innovation. The influence of networks on innovation e.g. gained attention in the last decade. Rogers (2004) shows that networking has a positive influence on the innovative activities, especially for small firms. Linking innovative activities with survival shows that innovative firms dependent on age and firm size survive longer (Cefis and Marsili 2006). A current topic is the subsidization of firms and it became more important in the last years. Herrera and Sánchez-González (2013) study the effects of R&D subsidies in relation to firm size. In general, firms with R&D experience are more likely to be subsidized and subsidies in general stimulate investments in R&D. Comparing subsidized with non-subsidized firms indicates that only small and large subsidized firms can benefit from this and increase their economic returns to innovation, whereas small firms improved sales of products new for the firms and large firms improved sales of products new to the market. Furthermore, the private R&D intensity of small and medium firms is stimulated by R&D subsidies. Back to the pure investigation of the innovation and firm size relationship shows that in the last years various researchers also started to analyze the reverse relationship, the impact of innovation for firm growth. In doing so, innovation is seen as one element triggering growth, but also further firm characteristics are relevant for firm growth, e.g. the patenting activities and also the persistence of them (Coad and Rao 2008; Demirel and Mazzucato 2012). The results of Deschryvere (2014) indicate that large and continuous innovative firms have mutually interdependent relationships between sales growth and R&D growth. New and small firms are most likely to become high growth firms. But, the propensity for growth is in fact higher for firms which invest in R&D (Segarra and Teruel 2014).

Cohen (2010) suggests that more complete models are needed to fully evaluate the Schumpeterian hypotheses and to understand in a better way the relationship between firm size and R&D and also between R&D intensity and market concentration.

## 3 Theoretical background

One approach in the literature that considers these claims and empirical findings is the cost spreading model of Cohen and Klepper (1996a). From the enormous body of literature about innovation, R&D and firm size relationship, the authors compiled four robust empirical patterns, based on the empirical observations. Firstly, the likelihood that a firm conducts R&D increases with its size (stylized fact 1). Secondly, in most industries there is a positive and close relationship between firm size and R&D inside the industry (stylized fact 2). Thirdly, a proportionate rise of R&D and firm size is observable in most industries (stylized fact 3). Fourthly, while the firm grows, the number of innovations which is generated per dollar decreases. Thus, small firms have compared to large firms more patents relative to their size (stylized fact 4). A model, based on the idea of cost spreading was developed by Cohen and Klepper, dealing with firm size, innovation and R&D productivity to explain the empirical patterns found about the R&D firm-size relationship. It is based on the idea that larger firms have an advantage of size given that they can spread their fixed costs of innovation over a larger amount of output and through this the returns on R&D will also increase with the level of output. A big advantage of this model is that it is also possible to distinguish between product and process innovations while testing the theory (Cohen and Klepper 1996a).

### 3.1 The cost spreading model

In the model of Cohen and Klepper (1996a) it is assumed that firms are more likely to use their innovation internally rather than selling it and, furthermore that firms anticipate that the benefits of innovation are a reduction of their average costs, and not an expectation of growth. The second assumption is based on the idea that innovations can be imitated very quickly, in fact in one up to three years. The profits from innovation are limited. Firms can only benefit from innovation by increasing the price-cost-margin. Because of this, the incentives to invest in R&D increase with the ex-ante output. In the model, size is not defined by the overall firm size but by the size of the business unit. Large firms might be less productive in R&D than small firms, but due to the fact that large firms have a greater output, they can spread the average costs over more output and can benefit from this. The larger a firm is the more it can draw advantages from these circumstances (Cohen and Klepper 1996a).

The amount of R&D spending and the returns on R&D are at all times determined by the size of the firm. The variable $\mathbf{r_i}$ describes the amount of R&D a firm undertakes in a certain product market. The fixed costs of R&D are zero and the price of R&D is normalized to equal one. Given this, the total costs of R&D are equal to $\mathbf{r_i}$. The price-cost margin is described as $\mathbf{pc_i\ (r_i)}$[1], and it can be increased by R&D. It is assumed that all firms face the same price-cost schedule. Considering the technical opportunities it is assumed that the productivity of the R&D effort can be described as $\mathbf{pc_i{'}(r_i) = f/r_i}$, where **f** is the index of the technical opportunities in an industry. All firms have the same technical opportunities. The variable $\mathbf{q_i}$ describes the output of the firm at the time when it conducts the R&D. The output, for which the innovation will be used, will be similar to the size of the output at the time when the R&D was conducted, given the fact that innovations can be copied quickly. Given a proportionate relationship between $\mathbf{q_i}$ and the output on which the innovation is applied, the output level with the innovation is $\mathbf{gq_i}$. The variable **g** is a factor of proportionality, which is used to describe in which amount q changes between the conduction and the commercialization of the innovation. The firms are constructed as price takers that make their R&D decisions independently. They choose $\mathbf{r_i}$ to maximize their profits from R&D described as $\mathbf{pc_i(r_i)gq_i\text{-}r_i}$. The profit maximizing value of $\mathbf{r_i}$ must satisfy the following condition:

$$\boldsymbol{r_i = aq_i\,,} \tag{1}$$

where $\mathbf{a \equiv fg}$ (Cohen and Klepper 1996a).

Cohen and Klepper (1996a) show that equation 1 can be used to capture most of the stylized facts. The spending on R&D in an industry is determined entirely by the size of the firm. This explains the strong relationship between firm size and R&D (stylized fact 2). The equation also implies that R&D will vary proportionally with firm size. Therefore, it also explains the variation in R&D depending on firm size (stylized fact 3). Equation 1 shows that the expenditures for R&D are determined by the size of the firm (level of firm output). The equation 1 can also help to elucidate the patterns presented in stylized fact 4, namely that the number of innovations or patents generated per each dollar of R&D

[1] The first derivative of the price-cost margin is $pc_i{'}(r_i) > 0$ and shows the positive influence of further R&D efforts. But the second derivative is $pc_i{''}(r_i) < 0$ and shows diminishing returns from R&D (Cohen and Klepper, 1996a).

decreases with the size of the firm. To examine this, it must be assumed that $\boldsymbol{pc_i(r_i)}$ is a measure of the number of patents or innovations of a firm. Since the second derivate is <0, $\boldsymbol{pc_i(r_i)/r_i}$ must be a decreasing function of $r_i$. The larger a firm, the fewer innovations or patents will be generated per dollar. Because $r_i$ and $q_i$ are directly related, this leads to a decline of the innovative output per each dollar of R&D spending while the spending increases. Because of this, small firms have more innovations or patents relative to their size compared to large firms. Due to their size (larger amount of output), large firms can profit more from their innovations. Thus, larger firms will undertake more R&D projects at the margin compared to small firms, but small firms will have more patents than proportional to the firm size. Cost spreading can be used as an explanation for the indicated lower productivity of larger firms mentioned in the first part of the stylized fact 4 (Cohen and Klepper 1996a). Up to here most of the stylized facts can be explained by the cost spreading model.

Relaxing the assumptions of the model through the introduction of fixed costs **F**, needed for a formal R&D program, and the opportunity for firms having different technological possibilities, considered in the model through firms featuring different R&D productivities **n**, allows to explain the remaining stylized facts (1 and part of 3). The heterogeneity is captured by assuming that all firms have the same R&D project opportunities, but they differ concerning the R&D effort productivity. This is included in the model by describing the price-cost margin as $\boldsymbol{pc_i(r_i)=n_i pc(r_i)}$, where $n_i>0$, where across firms $\boldsymbol{n_i}$ is not related to $\boldsymbol{q_i}$.

The maximum profit of R&D can be described as an increasing function of $\boldsymbol{n_i}$ and $\boldsymbol{q_i}$. Considering a certain degree of productivity **n**, the probability that a firm will conduct R&D will be higher, the larger $\boldsymbol{q_i}$ is. The introduction of fixed costs and firm heterogeneity (differences in R&D productivity) leads to

$$\boldsymbol{r_i = a n_i q_i\,.} \qquad (2)$$

A firm will only conduct R&D when

$$\boldsymbol{n_i pc(a n_i q_i) - a n_i q_i - F > 0.} \qquad (3)$$

This shows that the only change is that the R&D spending of a firm is determined by its R&D productivity. The larger a firm is, the larger the probability of conducting R&D (Cohen & Klepper 1996a).

By relaxing the assumption of diminishing returns to R&D, certain departures from proportionality can be explained by the model. It is considered that in different industries the R&D effort can increase more (less) than in proportion to the size of the firm. The marginal productivity of the firm is modified by introducing b>0, which adjusts the point at which the marginal return to R&D decreases, so if $pc_i'(r_i)=f/r_i^{1/b}$, then

$$\mathbf{r}_i = \mathbf{a}^b \mathbf{n}_i^b \boldsymbol{q}_i^b. \tag{4}$$

This shows, that the amount of R&D a firm undertakes within an industry will feature a more than proportional growth, if b>1 and a less proportional growth, if b<1. Cohen and Klepper point out, that proportionality should not be overestimated given that sometimes firms do not even know their productivity schedule and the rather simple methods used might lead to a picture of proportionality in an industry (Cohen and Klepper 1996a).

An important implication of the cost spreading mechanism is that it conditions the rate as well as the composition of R&D (Cohen 2010). The composition of R&D refers to the distinction between product and process innovations. The quantity of R&D which is undertaken by a firm is directly related to the output of the firm, at the time the firm conducts the R&D. But process R&D and product R&D differ in the extent to which they are dependent on the ex-ante output. Product R&D is assumed to lead to new products or can improve the quality of the product and boost the price of it. Because of this, the returns are assumed to be more independent of the ex-ante output. Process R&D, on the other hand, is not so easy to sale and it is assumed that this kind of R&D aims at lowering the average costs of production. Thus, the returns on process R&D will be affected by the ex-ante output. The larger a firm is, the more it can spread its average costs over the output of a certain product (Cohen and Klepper 1996a; b).

### 3.2 Firm size and the allocation of the R&D effort

Not only the total R&D effort is important, also the allocation of the R&D effort concerning product and process innovation plays an important role (Cohen and Klepper 1996b). Building on the model of cost spreading Cohen and Klepper (1996b) set up and test a model to determine to what extent the allocation of product and process innovation is influenced by the size of the firm. The underlying assumptions of the model are that the incentives to conduct R&D depend on the ex-ante output level of a given product. Only the firm size on business unit level matters and the firms are price takers. Furthermore R&D increases the firm´s price-cost margin (Cohen and Klepper 1996b).

The variable **q** describes the output at the time the firm conducts R&D and ***a*** labels the time until the innovation can be imitated. A certain part of the existing byers ***h*** will consume the new product and through sales to new buyers and licensing, the firm can earn further rents through additional output called ***K.*** With $\boldsymbol{r_1}$ the spending on process R&D of the firm **i** is described. The decrease of the average cost of process R&D are captured by $pc_1(r_1) - r_1$. For this it is assumed that $pc'_1(r_1) > 0$ and $pc''_1(r_1) < 0$ for all $r_1 \geq 0$. With $\boldsymbol{r_2}$ the spending on product R&D of the firm **i** is pictured. It is assumed that process R&D leads to a greater reduction of the manufacturing costs, but this happens at a decreasing rate. By $pc_2(r_2) - r_2$ the earnings of the price-cost margin related to the new product are represented. For product R&D it is also assumed that $pc'_2(r_2) > 0$ and $pc''_2(r_2) < 0$ for all $r_2 \geq 0$. In the model the returns to process (5) and product (6) R&D are described as

$$\pi_1 = a_1 q p c_1(r_1) - r_1 \ , \quad (5)$$

$$\pi_2 = a_2(hq + K) p c_2(r_2) - r_2. \quad (6)$$

The returns on process R&D are proportional to the ex-ante output of the firm (5), while the returns on product R&D do not exhibit a proportional rise to the size of the firm (6). In doing so, the equation 5 and 6 can explain why the shares of process R&D will increase more while the firm is growing (Cohen & Klepper 1996b).

Cohen and Klepper (1996b) argue that the model can be structured further to derive different propositions about the relationship between q and p. Firstly, the share of process R&D (p) of a firm can be described as an increasing function of the ex-ante output of the firm. Secondly, p should increase with q at a decreasing rate. Thirdly, depending on the balance between product and process R&D, δp/δq should vary across industries. Fourthly, the effect from output on process R&D will be greater in industries where the sale opportunities of product innovations are higher or can lead to a faster growth. The model indicates that the advantage of cost spreading is more important for process rather than for product R&D (Cohen and Klepper 1996b).

### 3.3 Testing of the cost spreading model in the literature

The cost spreading model or parts of it were tested in recent years by various authors e.g. Fritsch and Meschede (2001) and Tsai and Wang (2005). The assumptions about the advantages of size attributable to cost spreading can be confirmed in most parts. Fritsch and Meschede (2001) test one assumption of the model, namely that large firms invest more in R&D related to process R&D than small firms. The analysis is based on 1800 manufacturing firms in three German regions, but it lacks information about the business unit size. Their results indicate that the R&D expenditures rise less than proportional with firm size. Analyzing the differences between product and process innovation related to firm size indicates that the spending on process R&D increases slightly more than the spending on product R&D with firm size. Fritsch and Meschede (2001) argue that this is a confirmation of the Cohen and Klepper hypothesis. Furthermore, Tsai and Wang (2005) use R&D output elasticity as a measure of firm R&D performance to analyze parts of the cost spreading model at the example of 126 manufacturing firms in Taiwan, which are listed on the Taiwan Stock Exchange. Their results indicate that there is an approximating U-type relationship between firm size and R&D productivity. The R&D productivity is higher for small and for large firms compared to medium sized firms. Additionally, these firms have a cost spreading advantage which is related to economies of scale. The authors argue that their results support the Schumpeterian hypotheses in that way that the R&D performance can be seen as an increasing function of the size of the firm.

**3.4 Hypotheses and the new approach to test the cost spreading model**

Different aspects of the cost spreading model and its application have been analyzed and tested. Fisher and Temin (1973) and, in line with this, also Cohen and Levin (1989) already claimed that an appropriate testing of the Schumpeterian hypothesis, namely the relationship between firm size and innovation, requires that innovative output should be used rather than innovative input (R&D expenditures, R&D employees), which is purely a consequence of economies of scale. The data analyzed was mostly past data, e.g. input data like R&D expenditures or R&D employees and output data like patents or sales. Using planned output data as a measure for the growth plans of the firms is new. In this article, the first step is to test whether cost spreading is a possible explanation for the innovative activities of the solar cell producing firms in Germany by using patents as an innovation output as well as the output as a measure of firm size on business unit level. However, firms do not act only in the past - firms plan and they have a future. The cost spreading model states that larger firms have a competitive advantage because they can spread their costs over a larger amount of output. But is cost spreading also a part of the strategic behavior of firms? Do firms that plan to grow invest more in product or process innovations because they are aware of the cost spreading mechanism and that they will gain an advantage through this? In doing so, the contribution of this article to the literature is the investigation of the planned output of the firms.

Cohen and Klepper (1996a) tested the model with data for several industries. In this article the model will be tested only for one industry and furthermore, this dataset lacks information about R&D expenditures. I will focus on the firm size, innovation and growth plan relationship. Thus, the hypotheses will be adjusted for this case. Those patterns investigating the relationship between R&D investments and innovation cannot be analyzed with this data. Based on the model, different hypotheses are developed to test whether cost spreading is observable in the photovoltaic industry and whether the model explains the patterns of innovative firm behavior in this industry.

In the model, the first pattern is that the likelihood to perform R&D increases with firm size (Cohen and Klepper 1996a). To investigate whether this pattern is also observable for the firms in the photovoltaic industry the hypothesis is adopted:

**H1:** *The probability for patents increases with the size of the firm.*

The second pattern indicates that firm size and R&D are positively related in industries (Cohen and Klepper, 1996a). To test whether the pattern is also observable in the photovoltaic industry the following hypothesis is developed:

**H2:** *Large firms have more innovations.*

Large firms tend to have relatively more process innovations. Cohen and Klepper (1996a, b) argue that a second possibility to test the cost spreading argument is to differentiate between certain types of R&D, with the underlying assumption being that firm size also has an influence on the composition of the R&D. If the firm grows, the output will be larger and accordingly the firm will spend more money on innovative activities concerning process R&D. The following hypotheses are built concerning the relationship of firm size and the different types of R&D, namely process R&D:

**H3a**: *Large firms have more process innovations.*

**H3b**: *The number of process innovations increases with size at a decreasing rate.*

In the model it is assumed that the amount of R&D is determined by the size of the output while the firm conducts the R&D because the firm does not expect to grow that fast. So if a firm plans to grow, it anticipates that the output over which it can spread the costs will be larger. Due to this, if firms anticipate that they can spread their costs over a larger output because they plan to grow, they will probably invest more in R&D. It will be analyzed whether cost spreading can be seen in the strategic decisions. In detail, if it is observable in the planned data:

**H4**: *Firms which plan to grow will have more innovative activities than other firms.*

**H5:** *Firms which plan to grow will have more process innovations.*

## 4 Data and econometric specification

### 4.1 Sample

For the analysis different variables are used to test the cost spreading model and the further implications. Using output and planned output as a measure of firm size is new and distinguishes the testing from the way and the data Cohen and Klepper (1996a) used to test the cost spreading argument. Cohen and Klepper (1996a) use industry data (FTC data). The dataset of them includes 75 industries with 10 or more observations for every single industry. They have information about R&D, sales, business unit size and R&D expenditures, which can furthermore be classified as expenditures for product or process innovation.

For the analysis in this article an unbalanced panel data set of the solar cell producing firms of the photovoltaic industry in Germany is used. The dataset is based on the annual cell production overview of *Photon*, which has been published since 2001. *Photon* is the most popular trade publication in the photovoltaic industry. The unbalanced panel data set covers the time period between 2000 and 2011. It includes output data for most of the solar cell producing firms - in total 37 firms. The production overview lists all producing facilities in Germany. As some firms are active in more than one country, the output is listed separately for each country. This article focuses on the output (Megawatt per year) for all facilities producing in Germany. Output can be used as a direct measure of the size of the business unit. In doing so, output directly reflects the size and it is not weakened by e.g. prices. Two firms are missing, because no output information is available for them. The production overview contains information[2] on the firm level. In detail it contains realized output[3] for the time period between 2000 and 2011, and furthermore planned output for the years 2005-2011. Given the different time periods the number of observations in the estimation approaches will differ. Entry dates of all firms are identified from the first appearance in the output overview. The data for the years before 2000 was collected from books (c.f. Räuber 2005). Missing data was collected through an internet based search (Blankenberg and Dewald 2013). Entry dates refer to the point in time when a firm entered the solar cell market, not necessarily the founding date of the firm.

Beside the output data, the dataset also includes information about the innovative activities of the firms measured via patent data. In the literature patents are commonly used as indicator for innovative activities (Acs and Audretsch 1989; Griliches 1990; Acs et al. 2002; Nagaoka et al. 2010). Nevertheless, patents as an indicator of innovative activity also feature some problems. Problems are e.g. that patents are not applied for all inventions, some inventions are not patentable and also the patents vary greatly in quality. Griliches (1990) indicates that some of the problems can be handled by just limiting the analysis to only one industry. However, patents will be used as an indicator of innovative activities in this article. The patents were identified through a named based search of all solar cell producing firms Germany in DEPATISnet[4]. DEPATISnet allows for different search levels. The "Beginner's search" allows for entering just the name[5]. The results can be limited to various aspects, e.g. it is possible to remove family members. Family members, i.e. all patents related to the priority patent, have been deleted. Only DE, EP and WO publications are included in the analysis. Following previous research (e.g. Breschi et al. 2003; Ma and Lee. 2008) and to get as close as possible an

[2] To a certain extent the data is an estimation by Photon.
[3] Output describes what the firms have produced. Planned output gives an overview of what the firms want to produce in the following year.
[4] service offered by the "German Patent and Trademark Office (DPMA)"
[5] Search example: entering the name "Q-Cells" leads to the search query: (((Q (L) Cells)/PA) OR ((Q (L) Cells)/IN))

overview of the innovative activities, all types of patents, from patent application to granted patents, have been considered. Based on the assumption that if a firm applies for a patent it has invested money up front and it has tried to innovate, regardless of whether the patent is granted or not. However, each patent is only counted once. In DEPATISnet it is possible to download the result list. After doing this, all patents were stepwise analyzed and rechecked by hand, whether they are really related to the photovoltaic industry. All patents not related to the photovoltaic industry were deleted from the dataset (Bauckloh 2012). Given the interesting time period, only the information about patent data in the years between 2000 and 2011 has been used. In total, 32 of the 37 firms have 699 patents in this time period. The patents of the firms before 2000 are only considered as a dummy (99 patents). Hagedoorn and Cloodt (2003) summarize that raw patent counts are a straight forward quantitative measure, which is in most areas of the economic literature widely accepted as a measure for comparing the innovative activities of firms or can be seen as a performance indicator. This article uses only patent counts for the econometric analysis.

For the model it is important to differentiate between product and process innovation. A schema was developed to classify the patents of the dataset concerning product and process patents. The main class of the patent is relevant here. All patents have been analyzed and it is observable that most of the patents belong to a certain number of IPC classes (most important are the classes: H01L[6], G01N[7], C23C[8]). These classes are based on technological and functional aspects. Looking at the patents it is apparent that most of the patents in one subclass or one group belong to the same innovation type. What is more, this is also rechecked and confirmed by looking at the name of the subclass or group. Hence, the assignment to one kind of innovation is based on the attribute "IPC class" (Bauckloh 2012). On the basis of IPC classes a differentiation is made between product innovations, process innovations and patents that raise the degree of efficiency, given that the type of innovation is also reflected in the name of the subclasses. For the analysis, all patents which raise the degree of efficiency are counted as product innovations. The reason for this is that the increase of the efficiency degree is not a cost reduction but it can be seen as an improvement of an existing product.

Given that DEPATISnet is a database that covers all patents from application to granted patents, this database is assumed to be complete. While no information[9] about the output or the planned output

---

[6] H01L: "Semiconductor devices; electric solid state devices not otherwise provided for"

[7] G01L: "Measuring force; stress; torque; work; mechanical power; mechanical efficiency; or fluid pressure"

[8] C23C: "Coating metallic material; coating material with metallic material; surface treatment of metallic material by diffusion into the surface, by chemical conversion or substitution; coating by vacuum evaporation, by sputtering, by ion implantation or by chemical vapour deposition, in general"

[9] The data collection of Photon is done by questionnaire, which is sent to the company annually by Photon (telephone conversation with Photon on the 05/February/2013).

is a missing, the lack of patents is information. This means that if a firm has no patent in a certain year, this is information about the innovative activities of the firm and not missing information.
Thus, 37 firms have been considered, with 32 firms having 699 patents for the relevant years.

### 4.2 Variables

A set of variables is used to capture the structure of the solar cell producing firms in the photovoltaic industry in Germany.

Dependent variables: To capture the innovative activities the information about patents is used as a proxy for R&D. The dependent variable for the econometric analysis is PATENTS. The dependent variable is a count variable, i.e. the number of patents each firm has in each year between 2000 and 2011. The range is from zero per firm to a maximum of 56 in total. The second dependent variable is PROCESS PATENTS. Like patents, it is also a count variable, i.e. all process patents a firm has in each year. Here too, the range is from zero to a maximum of 27 in total. As a third dependent variable a dummy variable for the probability of patents (D_PATENTS) is used. This variable is coded as 1 if a firm has patents in a year, i.e. whether there is any innovative activity and is coded as 0 otherwise.

Independent variables: The first main independent variable is **OUTPUT** (realized output). Output gives the value of the number of solar cells in Megawatt per year (Mw/p.a.) (based on Photon 2001-2013). It captures the size of the firm on the business unit level. For this industry it is plausible to use output as a proxy for size because of a very narrow definition of the firm population in one industry – the cell producers in the photovoltaic industry. All firms produce the same product – solar cells. For this reason and because in the pertinent literature the business unit is of particular relevance, the output of a firm is a good measure of the size of a firm in a certain industry.
For the second part of the analysis a variable is required which captures the growth plans of the firms. Thus, compared to the output variable, a growth rate is used. The variable for the planned output is used to analyze whether growth plans in general influence the innovative activities. Therefore, the growth plans itself are of interest and not the size of the firm. Using only the growth rates reflects directly growth plans and ignores the size of the firm. Thus, the second main variable is **G_PLAN**. This variable illustrates the growth of the planned output of each firm:

$$g\_plan = planned\ output_t - output_{t-1} \quad (7)$$

Control variables: To control for that the results are driven by further effects, I control for different firm characteristics. Some solar cell producers are not completely independent because they belong to a group (of companies). The **GROUP** variable controls for possible effects that can arise through the fact that a firm is part of a larger group which might influence the innovative process, e.g. those firms possibly invest more in R&D than a single firm because the group provides security (Frenkel et al. 2001; Shefer and Frenkel 2005; Segarra and Teruel 2014). This variable is coded as 1 if a firm belongs to a group and is coded as 0 otherwise. The variable **PATENTS BEFORE ENTRY** is used to control for pre-entry experience of the firms which probably determines the performance of the firms. Some solar cell producers started research (and patenting) about solar cells 1 to 5 years before they entered the market. This variable is coded as 1 if a firm has patents before entry and is coded as 0 otherwise. Some firms have additional production facilities in other countries. As a third control the variable **GLOBAL** is used. This variable shows whether a firm is not only active in Germany, but has also production facilities in other countries. So I can control for whether this influences the innovative behavior in the production facilities in Germany. This variable is coded as 1 if the firm has also production facilities in other countries and is coded as 0 otherwise.

### 4.3 Descriptive statistics

Table 1 gives an overview of the summary statistics of the main variables considering all observations with output data (no information is a missing) for the time period between 2000 and 2011. Table 2 reports the overview considering all observations with planned output data (no information is a missing) for the years 2005-2011. The smaller number of observations concerning planned output is due to the fact that Photon only started reporting planned output in 2005 while reporting realized output started in 2000.

Table 3 reports the pairwise correlations between the dependent and independent variables for the time period 2000-2011. Furthermore, Table 4 reports the correlation coefficients between the dependent and independent variables for the smaller dataset covering the time period between 2005 and 2011. This is important for the analysis of the planned data. Both tables show a high correlation between output and planned output. Both the dependent variables patents and process patents are also highly correlated. This shows an expected strong relationship between the realized and planned activities of the firms. But in general the relationships between the variables are mostly weak. Hence, the results should not be disturbed by multi-collinearity problems.

### 4.4 Econometric models

To estimate the relationship between firm size, innovation and firm growth different methods are used. The panel dataset offers annual firm observations of the solar cell producing firms in Germany over a time period of 12 years. The panel is an unbalanced panel because not every firm is active in all periods, some firms entered the market later and other firms dropped out of the market.

To test hypothesis 1 the probability of having patents is analyzed. Thus, an adequate method here is to use probit models. In a first step, the panel structure is ignored and a pooled probit model is estimated. Subsequently, the second step of the analysis will be to apply random effect (RE) models[10] to consider the panel structure and the firm specific effects (Hausman et al. 1984).

The estimation of the remaining hypotheses requires a different estimation approach. Since the dependent variables patents and process patents are nonnegative integers (count data), this class of models[11] will be used for the analysis. The dependent variable PATENTS is strongly skewed to the right as is the dependent variable PROCESS PATENTS. The variance is nearly 14 times larger than the mean – showing that there is overdispersion in the data. The large chi-square value in the poisson goodness of fit also indicates that a poisson model is inappropriate. A general model which deals with overdispersion is the negative binomial model (Hilbe 2011). Again, as a first step I estimate pooled negative binomial models. The likelihood-ratio tests in the model indicate whether it is the right choice to use the negative binomial model compared to the poisson model. Unobserved heterogeneity may be present in the data due to unobserved factors (e.g. the R&D expenditures, financial equipment of the firms, laws, performance of the firm and distance to other firms or universities, etc.). Given this, the second step of the analysis will be to apply random effects (RE) models and fixed effects (FE) models[12] to consider the firm specific effects (Hausman et al. 1984). Technically the Hausman test will be an indicator, to decide whether FE or RE is the best model and it indicates that RE is the right model. But, given that it is very likely that the error term is correlated with the explanatory variables the FE model will be also considered and presented as a robustness check in all estimation approaches. The FE specification features a reduced number of groups and observations given that this model specification is a conditional estimation and all observations of a group are dropped, when there is only one observation in one group or when all outcomes are zeroin one group. In addition to the coefficients also the average marginal effects are illustrated for each model, given that through this it is possible to quantify the relationship between the variables.

## 5 Results

This article is based on a dataset which offers the opportunity to test the cost spreading model (and through this also the Schumpeterian hypothesis) in an appropriate way. The data is used to analyze whether large firms have an advantage of size which is attributable to the cost spreading mechanism.

[10] FE model is not available for a panel data probit model.

[11] The aim of count models is to explain the number of different events or occurrences. It is assumed that the counts are right skewed, heteroscedasticity is intrinsic and with the mean of the distribution also the variance will increase. The basic model for count data is the Poisson regression model, where the heterogeneity parameter has the value zero and equidispersion is assumed (Hilbe 2011).

[12] The negative binomial is not a "real" fixed-effects model because of a lack of control for all covariates (Hilbe 2011; Allison and Waterman 2002).

To do this, the influence from the realized output of the firms as a measure of firm size on the business unit level is regressed on the innovative output, namely raw patent counts (2.5.1). Additionally, in the second step the relationship between the growth plans (measured via the planned output in proportion to the realized output of the firms) and the number of patents is analyzed (2.5.2).

**5.1 The relationship between firm size and innovation**

In the cost spreading model various assumptions about firm size and innovative activities are made. Large and small firms differ in their innovative output. The theory assumes that the rate and the composition of R&D are conditioned by cost spreading. Cost spreading can lead to economies of scale in innovation. The data allows an estimation of the relationship between innovation and firm size on the business unit level. If the cost spreading argument is valid, a close relationship between innovation and firm size should be found. Considering the criticism in the literature about the variables used an advantage of this empirical study is that patents are used as an innovative output and furthermore output as a measure of the firm size on business unit level. The level of analysis in this study is the firm level. The main independent variable in hypothesis 1 up to 3 is the size of the firm measured via output. I start with estimating the impact of output on the probability for patents (Table 5a). What is more, I investigate the effects between output and innovation or process innovation (Table 6a-8a). To interpret the results, Tables 5b-8b report the average marginal effects.

**Hypothesis 1** predicts that the probability for patents increases with the size of the firm. Cohen and Klepper (1996a) show that the likelihood to conduct R&D increases with firm size. To test the first hypothesis, following the approach of Cohen and Klepper (1996a) a dummy variable for patents is introduced and used as the dependent variable in these estimations. The results of the probit estimation (1a-1e) are presented in Table 5a and b. Different control variables were considered stepwise in the RE approach (1a-d) considering the panel structure as well as in the pooled[13] (1e) approach. The main independent variable output is positive and significant to highly significant in all estimation approaches (Table 5a, models 1a-1e). The average marginal effects are reported in Table 5b and it reflects that the probability for a patent will increase by 0.0017 (1a) when output in Mw/p.a. increases for one unit (+1 Mw/p.a.). In all estimations the coefficients are positive and highly significant and this indicates that an increase of the output increases the predicted probability for patents. All estimations in Table 5a include a set of industry dummies to control for firm effects. Considered here are the dummies whether a firm is active globally and whether it belongs to a

[13] The assumptions behind probit and logit models are little different, given that the probit model is not an advanced model but just an alternative to the logit model. The results of the logit estimation are very similar and can be provided upon request.

group. The results show that having global output as well as being part of a group does not influence the number of patents. Adding one or both control variables in the model only leads to a slightly smaller significant output coefficient. This shows that the results concerning the influence of the firm size are quite robust over the different estimation approaches. Including the control variable, the average marginal effects still indicates that the probability for a patent will increase by 0.0015 for each additional unit of output (in Mw/p.a.). In hypothesis 1, presented in Table 5a and b the main interest is on the effects from output on the probability for patents. The results thus indicate that the size of the firm increases the probability of having patents. Hypothesis 1 can be supported.

**Hypothesis 2** predicts that large firms have more innovations. Thus, that R&D and firm size show a positive relation. The results of the negative binomial estimation (2a-e, 3, 4) are presented in Table 6a and b. Column 2a-e presents the results of the RE estimation. The fixed effects estimation is presented in Column 3 and the pooled estimation in 4. In this part of the econometric analysis the count of patents is used as the dependent variable. The different estimation specifications, as well as the stepwise inclusion of control variables, lead to a clear picture presented in Table 6a. The coefficients of the main independent variable output are positive and highly significant in all estimation approaches. This means that for a one unit change in the output (measure for firm size; in Mw/p.a.), the difference in the expected counts of the patents is expected to change by 0.005 in the RE (2a-2e) and the FE (3a) negative binomial estimation. Which is a small decrease compared to the pooled estimation (4a, coefficient = 0.006). The average marginal effect also indicates this. It can be interpreted as that the number of patents will increase by 0.015 (Table 6b, Column 2a) respectively by 0.0123 up to 0.0145 (Table 6b, Column 2b -4) for every additional unit of output (Mw/p.a.). Thus, innovation and firm size have a close and positive relationship. The larger a firm is the more innovations it will have. Including the control variables shows that there is a positive effect on innovation when firms start patenting before entering the market, but this does not influence the coefficient of the output. The coefficient of the main variable is very similar in all estimation approaches what again indicates that the results are quite robust, thus it shows that firm size has a strong impact on the innovative activities. The results about the firm size and innovation relationship support the findings of the literature about the Schumpeterian hypothesis.

What is more, researchers often also find a proportional relationship between innovation and firm size. Analyzing this here for the data is difficult given that the dataset lacks information about innovative input. Probably there is a proportional relationship between the innovation input factors and size in this industry, but it is not possible to examine this. Analyzing a possible proportional relationship between innovative output and size indicates that there is none. Patents and output have no proportional relationship. It is only possible to show that there is a positive relationship. Nevertheless there is empirical support for Hypothesis 2.

**Hypothesis 3a** states that large firms have more process innovations. The distinction between product and process innovations plays a central role in the model of Cohen and Klepper (1996a; 1996b). At a certain point related to size the investment in process innovations will be cost efficient. So it is assumed that larger firms have more process innovations. The results of the negative binomial estimation (5a-e, 6, 7) are presented in Table 7a and b. Column 5a-e shows the results of the RE negative binomial estimation. The FE negative binomial estimation is presented in column 6a and the pooled estimation in 7. In this part of the econometric analysis the count of process patents is used as the dependent variable. The different estimation specifications, as well as the stepwise inclusion of control variables, lead to a clear picture presented in Table 7a and b. All models (5a-e, 6a, 7) in their specification show a highly significant influence from the main independent variable output on process innovation. This means that for a one unit change in the output (measure for firm size; in Mw/p.a.), the difference in the expected counts of the patent variable is expected to change by 0.005 in the full RE (Table 7a; 5e) and FE (Table 7a; 6) negative binomial estimation or 0.006 in the pooled negative (Table 7a; 7) binomial estimation. Because the FE specification is used the number of observations dropped from 181 to 143, but still output has a positive and highly significant influence on process patents. The average marginal effect of output (Table 7b) reflects the increase in process patents by 0.087 (Table 7b, model 5a) for every additional unit of output (Mw/p.a.). Thus, an increase in output leads to an increase in the number of patents. Again the control variables are added stepwise. The coefficients for global output and being part of a group are insignificant. Having patents before entry is relevant for the number of process patents a firm has. The coefficient of the main variable output is very similar in all estimation approaches what again indicates that the results are quite robust. The results indicate that there is empirical support for Hypothesis 3 that larger firms have more innovative activities.

**Hypothesis 3b** predicts that the number of process innovations increases with size at a decreasing rate. This hypothesis is based on Cohen and Klepper (1996b). The results of the negative binomial estimation (8a-e, 9a, 10a) are presented in Table 8a and b. Column 8a-e displays the results of the RE negative binomial estimation. The FE negative binomial estimation is presented in column 9a and the pooled estimation in 10a. In this estimation approach the count of process patents is again used as the dependent variable. To test whether there is a non-linear, (inverted) U-shaped relation between firm size and innovation the squared output is used as a second main independent variable, additionally to the main independent variable output. All specifications lead to a clear picture. The coefficient of the main independent variable output is always positive and highly significant. Again, the number of observations dropped from 181 to 143 in the FE model. Nevertheless, output has still a positive and highly significant influence on process patents. Using the squared form of output shows that the coefficients are systematically negative. The effect itself is consistent, but not significant. Given this observation, it can be assumed that a larger sample size could probably lead to significant results. But nevertheless, for this data it is not possible to show that the number of process innovations increases with size at a decreasing rate. Thus, Hypothesis 3b cannot be supported.

**5.2 The relationship between growth plans and innovation**

The previous results show that the cost spreading mechanism is a possible explanation for the innovative activities in the photovoltaic industry in Germany. The findings are quite robust. To show that cost spreading is an important advantage of size is and it has not been shown for the photovoltaic industry before. Moving to the second part of the econometric results, the following hypotheses analyze whether cost spreading is observable in the planned data (planned output) – the relationship between growth plans and innovation will be examined. Given that the unbalanced panel data set only includes planned data for the years 2005-2011 the results are based on a reduced dataset. The main independent variable in hypothesis 4 up to 5 is the planned size of the firm measured via the growth of the planned output (g_plan). In Tables 9a-10a I investigate the effects between planned output and innovation or process innovation. The marginal effects are reported in Tables 9b-10b.

**Hypothesis 4** analyzes whether firms which plan to grow will have more innovative activities than other firms. The influence from the planned growth on innovations (patents in general) is analyzed.
The results of the negative binomial estimation (Column 11a-f, 12a,b, 13a,b) are presented in Table 9a and b. Column 11a-f displays the results of the RE negative binomial estimation. The FE negative binomial estimation is presented in column 12a,b and the pooled estimation in 13a,b. The dependent variable in this estimation approach is the count of patents. The main independent variable is g_plan. To control for level effects the lagged $output_{t-1}$ (Column 11f, 12b, 13b) was used and furthermore the control variables are introduced stepwise. The different estimation specifications, as well as the stepwise inclusion of control variables, lead to a clear picture presented in Table 9a and b. The coefficients of the main independent variable are positive and highly significant in all estimation approaches. This means that for a one unit change in the firm size (planned output growth is measured in (Mw/p.a.), the difference in the expected counts of the patent variable is expected to change by 0.008 in the full RE (11e,f) negative binomial model. The coefficients only change little in

the different estimation approaches. But in all models a positive and significant influence is observable. Thus, the larger the planned growth is the more innovations a firm has. The number of observations dropped from 115(113) to 93(91) in the FE estimation, but still output has a positive and on a 5% level (12a) significant influence on process patents. Also the number of observations decreases in the model, the variety in the standard errors is small. Considering the level effects, the coefficient of output in the FE model is only significant on a 10% level. The level effects itself is insignificant. It is only significant on a 10% level in the pooled negative binomial model. But, the size of the sample must be considered by interpreting the results. The coefficients are positive and highly significant over all model specifications. The average marginal effect here can be interpreted as that the number of patents will increase by 0.0325 (Table 9b, Column 11a) for every additional unit of megawatt/p.a. of planned output growth. Thus, a positive and significant correlation between the planned growth and the innovative activities can be seen. In addition, global output has an influence on the innovative activity in this estimation. Moreover the effect of the innovative activity before entry is significant on the 10% level. The stepwise consideration of the control variables does not really affect the size of the coefficients, what indicates that the results are quite robust – Hypothesis 4 can be supported.

**Hypothesis 5** predicts that firms that plan to grow will have more process innovations. The influence from the growth of the planned output on process innovation (process patents) is analyzed. The results of the negative binomial estimation (14a-f; 15a,b; 16a,b) are presented in Table 10a and b. Column 14a-f displays the results of the RE negative binomial estimation. The FE negative binomial estimation is presented in column 15a,b and the pooled estimation in 16a,b. The dependent variable in this estimation approach is the count of process patents. The main independent variable is g_plan. To control for level effects the lagged $output_{t-1}$ (Column 14f, 15b, 16b) was used and furthermore the control variables are introduced stepwise. The coefficients are highly significant and positive. The results can be interpreted in that way that when the firm size increases for one unit (planned output growth is measured in Mw/p.a.), the difference in the expected counts of the patent variable is expected to change by 0.010 in the full RE (14e) negative binomial model and to change by 0.009 in the full RE (14f) negative binomial model controlling for the level effect. The coefficients are smaller in the FE and the pooled estimation, but still positive and significant. However, there is a positive and significant relationship between planned output growth and innovative activities, concerning process innovations. The average marginal effect of planned output indicates that the number of process patents will increase by 0.0198 (Table 10b, model 14a) for each unit of additional planned output growth (Mw/p.a.). Including the control variables shows that there is a positive effect on innovation when firms start patenting before entering the market. The coefficients of the main independent variable g_plan are only little influenced by the introduction of the control variables, indicating robust results. The variable for the planned output growth of the firm is positive and significant in all models, indicating that there is a positive correlation between planned output growth and innovative activities. The results indicate that firms that plan to grow will have more process innovations. Thus, there is empirical support for Hypothesis 5.

## 6 Conclusions

This paper provides evidence that the cost spreading mechanism can be used as an explanation for the innovative activities in the photovoltaic industry in Germany. Going back to Schumpeter, a large amount of research has been undertaken in the past 70 years to investigate the relationship between R&D, innovation and firm size. The cost spreading model of Cohen and Klepper (1996a) considers the relationship between R&D effort and productivity as well as size. This paper deals with the question whether cost spreading can explain the innovative behavior in the photovoltaic industry in Germany and whether cost spreading can also be found in the planning data of the firms. Output as a measure for firm size was used in the econometric analysis, as it was claimed before by different scholars (e.g. Fisher and Temin 1973, Cohen and Levin 1989). The use of the planned output of the firms to examine the relationship between the innovative activities and the growth plans of the firms is an original contribution to the literature.

The results show that the probability for patents increases with firm size. The results of earlier research indicate that the probability to engage in innovation increases with firm size but the output of the innovation process decreases with size. Due to the lack of information about the input of the innovations process, it is not possible to say something about the decreasing rate, but the results support the finding, that the probability for innovation increases with firm size. Furthermore a close and positive relationship between firm size on business unit level and innovation is observable. The results indicate that because of their size large firms have an advantage in innovation in the photovoltaic industry in Germany, which is attributable to cost spreading. The results related to the planned firm data indicate that the firms are aware of the advantage of cost spreading leading to higher innovative activities. Thus, the incentives to invest in innovation are also triggered by growth plans of the firms. Firms that plan to grow will have more innovations and also more process innovations. Schumpeter argued that large firms have an advantage in innovation. The results of this article support this. The explanation for this is the cost spreading mechanism, which was modeled by Cohen and Klepper whose objective was to find an explanation for the findings in the literature related to the Schumpeterian hypothesis.

A number of questions concerning the evolution of the photovoltaic industry in Germany are still left. The data used to test the model in this article has some limitations, e.g. the dataset does not include any information about R&D expenditures. Cohen und Klepper (1996a) assume that innovations are generated all the time, but while the firm size increases the number of innovations which is generated per dollar decreases. So this assumption could not be tested. A further limitation of this article is that the results show a high correlation, but it is not possible to show causality between firm size and innovative behavior and growth plans and innovation. What is more, various articles indicate that there is a proportional relationship between innovation and firm size. Firstly, this study lacks information about innovative input factors like e.g. R&D staff or R&D expenditures. Thus, it is only possible to examine whether there is a proportional relationship between the innovative output and firm size. This is not found here. Nevertheless, Cohen and Klepper (1996a) argue that a proportional growth can be found in some industries and it should not be overvalued. They point out that sometimes firms don´t even know their own productivity schedule, because the firms use simple methods which eventually can lead to a picture of proportionality in an industry. However, also there is no proportional relationship in the photovoltaic industry in Germany, the results support the findings of the earlier studies which analyze the relationship between firm size and innovation. Furthermore, the results of this paper are limited to one industry. Further research will therefore be needed to check whether these results are also applicable for other industries.

Nevertheless, the results support the argumentation of Cohen and Klepper (1996). The cost spreading model can explain the firm size, innovation and R&D relationship and furthermore this relationship is also observable in the planned data of the solar cell producing firms of the photovoltaic industry in Germany. The implications of the cost-spreading model leading to economies of scale can be confirmed – they are a possible explanation for the innovation and firm size relationship in the photovoltaic industry in Germany.

## Tables

**Table 1: Descriptive statistics (N=37, observations = 181, year 2000-2011)**

| Variable | Obs | Mean | Std. Dev. | Min | Max |
|---|---|---|---|---|---|
| id | 181 | 20.18232 | 11.39615 | 1 | 37 |
| year | 181 | 2007.271 | 3.034739 | 2000 | 2011 |
| patents | 181 | 3.071823 | 7.011921 | 0 | 56 |
| process patents | 181 | 1.657459 | 3.893558 | 0 | 27 |
| output | 181 | 59.31257 | 103.9242 | 0.001 | 570 |
| planned output | 119 | 93.66319 | 136.0545 | 0.1 | 780 |
| $output_{t-1}$ | 140 | 57.75539 | 103.2762 | 0.001 | 570 |
| global output | 181 | 0.1491713 | 0.3572454 | 0 | 1 |
| g_plan | 102 | 32.95598 | 49.6525 | -151 | 212 |
| group | 181 | 0.6906077 | 0.4635253 | 0 | 1 |
| patents before entry | 181 | 0.4475138 | 0.4986169 | 0 | 1 |

**Table 2: Descriptive statistics (N=37, observations = 138, year 2005-2011)**

| Variable | Obs | Mean | Std. Dev. | Min | Max |
|---|---|---|---|---|---|
| id | 138 | 20.05797 | 11.29613 | 1 | 37 |
| year | 138 | 2008.717 | 1.695805 | 2005 | 2011 |
| patents | 138 | 3.65942 | 7.85126 | 0 | 56 |
| process patents | 138 | 1.956522 | 4.34946 | 0 | 27 |
| output | 119 | 81.75387 | 121.151 | 0.1 | 570 |
| planned output | 138 | 83.22471 | 129.2704 | 0.01 | 780 |
| $output_{t-1}$ | 113 | 68.22859 | 112.0476 | 0.001 | 570 |
| global output | 138 | 0.1521739 | 0.3604979 | 0 | 1 |
| g_plan | 113 | 30.99689 | 47.78964 | -151 | 212 |
| group | 138 | 0.6956522 | 0.4618069 | 0 | 1 |
| patents before entry | 138 | 0.4057971 | 0.4928345 | 0 | 1 |

**Table 3: Correlation coefficients for dependent and independent variables** –
unbalanced panel, N=37, 181 observations, time refers to period between 2000 and 2011 (no information about output = missing)

| | id | year | patents | process patents | output | planned output | $output_{t-1}$ | global output | g_plan | group |
|---|---|---|---|---|---|---|---|---|---|---|
| id | 1.0000 | | | | | | | | | |
| year | -0.1058 | 1.0000 | | | | | | | | |
| patents | 0.0012 | 0.1764 | 1.0000 | | | | | | | |
| process patents | 0.0202 | 0.1626 | 0.9191 | 1.0000 | | | | | | |
| output | -0.0662 | 0.3086 | 0.5526 | 0.5213 | 1.0000 | | | | | |
| planned output | -0.0501 | 0.2300 | 0.5699 | 0.5507 | 0.9480 | 1.0000 | | | | |
| $output_{t-1}$ | -0.0538 | 0.3259 | 0.5084 | 0.4706 | 0.9266 | 0.9460 | 1.0000 | | | |
| global output | -0.0340 | 0.2136 | 0.4614 | 0.4044 | 0.3887 | 0.4773 | 0.4696 | 1.0000 | | |
| g_plan | -0.0598 | 0.0486 | 0.4818 | 0.4126 | 0.5532 | 0.6506 | 0.3693 | 0.2473 | 1.0000 | |
| group | -0.0797 | 0.1191 | 0.2000 | 0.1995 | 0.2838 | 0.3156 | 0.2905 | 0.2803 | 0.2800 | 1.0000 |
| patents before entry | -0.1386 | -0.1099 | 0.1846 | 0.1939 | 0.0034 | 0.0297 | -0.0211 | 0.0910 | 0.0200 | 0.0736 |

**Table 4: Correlation coefficients for dependent and independent variables -**
unbalanced panel, N=37, 138 observations, time refers to period between 2005 and 2011
(no information about pl_output = missing)

| | id | year | patents | process patents | output | planned output | $output_{t-1}$ | global output | g_plan | group |
|---|---|---|---|---|---|---|---|---|---|---|
| id | 1.0000 | | | | | | | | | |
| year | -0.0849 | 1.0000 | | | | | | | | |
| patents | -0.0140 | 0.1281 | 1.0000 | | | | | | | |
| process patents | 0.0155 | 0.1082 | 0.9153 | 1.0000 | | | | | | |
| output | -0.0701 | 0.1799 | 0.5368 | 0.5079 | 1.0000 | | | | | |
| planned output | -0.0499 | 0.2370 | 0.5763 | 0.5566 | 0.9480 | 1.0000 | | | | |
| $output_{t-1}$ | -0.0475 | 0.2117 | 0.4982 | 0.4616 | 0.9215 | 0.9473 | 1.0000 | | | |
| global output | -0.0434 | 0.1544 | 0.5291 | 0.4698 | 0.4583 | 0.4624 | 0.4801 | 1.0000 | | |
| g_plan | -0.0424 | 0.0387 | 0.4895 | 0.5193 | 0.5510 | 0.6602 | 0.3848 | 0.2407 | 1.0000 | |
| group | -0.1967 | 0.1224 | 0.1967 | 0.1932 | 0.3102 | 0.3145 | 0.2839 | 0.2801 | 0.2644 | 1.0000 |
| patents before entry | -0.1091 | -0.0015 | 0.2152 | 0.2194 | 0.0494 | 0.0090 | 0.0058 | 0.1018 | 0.0116 | -0.0307 |

**Table 5a: Effects of output on innovation (dummy for patents) in different model approaches**

| | Random-effects probit regression | | | | Probit Regression (pooled) |
|---|---|---|---|---|---|
| | 1a | 1b | 1c | 1d | 1e |
| output | 0.005***<br>(0.001) | 0.004***<br>(0.001) | 0.004***<br>(0.001) | 0.004***<br>(0.001) | 0.003**<br>(0.001) |
| **Control variables** | | | | | |
| global output | | | 0.260<br>(0.396) | 0.204<br>(0.398) | 0.318<br>(0.309) |
| group | | 0.293<br>(0.307) | | 0.265<br>(0.308) | 0.280<br>(0.218) |
| _cons | -0.171<br>(0.157) | -0.357<br>(0.252) | -0.193<br>(0.158) | -0.356<br>(0.248) | -0.318*<br>(0.171) |
| Log likelihood | -114.66973 | -114.21686 | -114.45676 | -114.08567 | -116.9938 |
| LR chi² | | | | | 16.26*** |
| Pseudo R² | | | | | 0.0650 |
| Wald statistics (all | 10.70*** | 11.59*** | 10.81*** | 11.64*** | |
| Number of observations | 181 | 181 | 181 | 181 | 181 |
| Groups | 37 | 37 | 37 | 37 | |

Note: Standard Errors in parentheses. Dept. Var. = dummy patents
Note: FE model not available for probit
*p<0.1; **p<0.05; ***p<0.01.

**Table 5b: Average marginal effects**

| | Average marginal effects (AME) | | | | |
|---|---|---|---|---|---|
| | 1a | 1b | 1c | 1d | 1e |
| output | 0.0017***<br>(0.0047) | 0.0015***<br>(0.0005) | 0.0016***<br>(0.0005) | 0.0015***<br>(0.0005) | 0.0011***<br>(0.0004) |
| **Control variables** | | | | | |
| global output | | | 0.0959<br>(0.1449) | 0.0747<br>(0.1446) | 0.1179<br>(0.1134) |
| group | | 0.1074<br>(0.1103) | | 0.0970<br>(0.1107) | 0.1037<br>(0.0796) |

Note: Standard Errors in parentheses. Dept. Var. = dummy patents
*p<0.1; **p<0.05; ***p<0.01.

**Table 6a: Effects of output on innovation (patents) in different model approaches**

| | Random-effects negative binomial regression | | | | | Conditional FE negative binomial regression | Negative binomial regression |
|---|---|---|---|---|---|---|---|
| | 2a | 2b | 2c | 2d | 2e | 3 | 4 |
| output | 0.005*** | 0.005*** | 0.005*** | 0.005*** | 0.005*** | 0.005*** | 0.006*** |
| | (0.001) | (0.001) | (0.001) | (0.001) | (0.001) | (0.001) | (0.001) |
| **Control variables** | | | | | | | |
| global output | | 0.526* | | | 0.467* | 0.195 | 0.650* |
| | | (0.278) | | | (0.276) | (0.341) | (0.350) |
| patents before entry | | | 0.522** | | 0.461* | 0.526 | 0.591** |
| | | | (0.260) | | (0.250) | (0.459) | (0.238) |
| group | | | | 0.244 | 0.167 | 0.118 | 0.364 |
| | | | | (0.299) | (0.289) | (0.450) | (0.281) |
| _cons | -0.665*** | -0.756*** | -0.974*** | -0.845*** | -1.155*** | -1.034** | -0.334 |
| | (0.218) | (0.221) | (0.261) | (0.308) | (0.332) | (0.414) | (0.241) |
| Log likelihood | -334.73756 | -333.05079 | -332.79126 | -334.39949 | -331.2042 | -220.47503 | -335.68014 |
| LR chi² | | | | | | | 64.05*** |
| Pseudo R² | | | | | | | 0.0871 |
| Wald statistics (all parameters) | 50.32*** | 58.35*** | 53.36*** | 50.36*** | 61.16*** | 33.09*** | |
| Number of observations | 181 | 181 | 181 | 181 | 181 | 167 | 181 |
| Groups | 37 | 37 | 37 | 37 | 37 | 31 | |

Note: Standard Errors in parentheses. Dept. Var. = patents
*p<0.1; **p<0.05; ***p<0.01.

**Table 6b: Average marginal effects**

| | Average marginal effects (AME) | | | | | | |
|---|---|---|---|---|---|---|---|
| | 2a | 2b | 2c | 2d | 2e | 3 | 4 |
| output | 0.0151*** | 0.0129*** | 0.0144*** | 0.0145*** | 0.0125*** | 0.0040*** | 0.0172** |
| | (0.0050) | (0.0042) | (0.0044) | (0.0049) | (0.0037) | (0.0015) | (0.0070) |
| **Control variables** | | | | | | | |
| global output | | 1.4427* | | | 1.2355 | 0.1695 | 2.0217* |
| | | (0.8300) | | | (0.7778) | (0.2985) | (1.1996) |
| patents before entry | | | 1.4021* | | 1.2215* | 0.4569 | 1.8391** |
| | | | (0.7730) | | (0.7026) | (0.4086) | (0.7920) |
| group | | | | 0.6878 | 0.4412 | 0.1023 | 1.1316 |
| | | | | (0.8624) | (0.7713) | (0.3916) | (0.8925) |

Note: Standard Errors in parentheses. Dept. Var. = patents
*p<0.1; **p<0.05; ***p<0.01.

**Table 7a: Effects of firm size (output) on process innovation (process patents) in different model approaches**

| | Random-effects negative binomial regression | | | | | Conditional FE negative binomial regression | Negative binomial regression |
|---|---|---|---|---|---|---|---|
| | 5a | 5b | 5c | 5d | 5e | 6 | 7 |
| output | 0.005***<br>(0.001) | 0.005***<br>(0.001) | 0.006***<br>(0.001) | 0.005***<br>(0.001) | 0.005***<br>(0.001) | 0.005***<br>(0.001) | 0.006***<br>(0.001) |
| **Control variables** | | | | | | | |
| global output | | 0.210<br>(0.335) | | | 0.150<br>(0.328) | -0.054<br>(0.388) | 0.373<br>(0.383) |
| patents before entry | | | 0.748**<br>(0.326) | | 0.713**<br>(0.321) | 0.627<br>(0.609) | 0.819***<br>(0.267) |
| group | | | | 0.356<br>(0.381) | 0.303<br>(0.372) | 0.113<br>(0.555) | 0.516<br>(0.327) |
| _cons | -0.580**<br>(0.278) | -0.622**<br>(0.287) | -1.127***<br>(0.353) | -0.862**<br>(0.410) | -1.381***<br>(0.454) | -1.113**<br>(0.568) | -1.150***<br>(0.295) |
| Log likelihood | -251.10351 | -250.91017 | -248.67543 | -250.66114 | -248.22667 | -155.75212 | -255.66428 |
| LR chi² | | | | | | | 53.81*** |
| Pseudo R² | | | | | | | 0.0952 |
| Wald statistics (all parameters) | 41.61*** | 42.48*** | 46.38*** | 42.11*** | 47.76*** | 28.56*** | |
| Number of observations | 181 | 181 | 181 | 181 | 181 | 143 | 181 |
| Groups | 37 | 37 | 37 | 37 | 37 | 27 | |

Note: Standard Errors in parentheses. Dept. Var. = process patents
*p<0.1; **p<0.05; ***p<0.01.

**Table 7b: Average marginal effects**

| | Average marginal effects (AME) | | | | | | |
|---|---|---|---|---|---|---|---|
| | 5a | 5b | 5c | 5d | 5e | 6 | 7 |
| output | 0.0087**<br>(0.0036) | 0.0081**<br>(0.0034) | 0.0082***<br>(0.0029) | 0.0083**<br>(0.0034) | 0.0076***<br>(0.0027) | 0.0045**<br>(0.0019) | 0.0097**<br>(0.0039) |
| **Control variables** | | | | | | | |
| global output | | 0.3327<br>(0.5354) | | | 0.2203<br>(0.4824) | -0.0479<br>(0.3459) | 0.6202<br>(0.6554) |
| patents before entry | | | 1.1143*<br>(0.5920) | | 1.0482*<br>(0.5580) | 0.5583<br>(0.5306) | 1.3628***<br>(0.5182) |
| group | | | | 0.5731<br>(0.6455) | 0.4455<br>(0.5635) | 0.1006<br>(0.4977) | 0.8583<br>(0.5642) |

Note: Standard Errors in parentheses. Dept. Var. = process patents
*p<0.1; **p<0.05; ***p<0.01.

**Table 8a: Effects of firm size (output) on process innovation (process patents) in different model approaches**

| | Random-effects negative binomial regression | | | | | Conditional FE negative binomial regression | Negative binomial regression |
|---|---|---|---|---|---|---|---|
| | 8a | 8b | 8c | 8d | 8e | 9 | 10 |
| output | 0.007***<br>(0.002) | 0.007***<br>(0.002) | 0.007***<br>(0.002) | 0.007***<br>(0.002) | 0.006***<br>(0.002) | 0.006**<br>(0.003) | 0.009***<br>(0.003) |
| output² | -4.13e-06<br>(4.30e-06) | -4.25e-06<br>(4.30e-06) | -3.16e-06<br>( 4.53e-06) | -3.51e-06<br>(4.38e-06) | -2.66e-06<br>(4.55e-06 ) | -2.33e-06<br>(5.00e-06 ) | -7.89e-06<br>(6.61e-06) |
| **Control variables** | | | | | | | |
| global output | | 0.218<br>(0.327) | | | 0.155<br>(0.326) | -0.056<br>(0.383) | 0.383<br>(0.387) |
| patents before entry | | | 0.714**<br>(0.328) | | 0.685**<br>(0.323) | 0.615<br>(0.609) | 0.749***<br>(0.272) |
| group | | | | 0.300<br>(0.387) | 0.266<br>(0.377) | 0.078<br>(0.557) | 0.458<br>(0.327) |
| _cons | -0.661**<br>(0.293) | -0.708**<br>(0.303) | -1.158***<br>(0.355) | -0.885**<br>(0.412) | -1.380***<br>(0.454) | -1.124**<br>(0.569) | -1.161***<br>(0.292) |
| Log likelihood | -250.64264 | -250.42204 | -248.43126 | -250.33994 | -248.05521 | -155.64295 | -255.01433 |
| LR chi² | | | | | | | 55.11*** |
| Pseudo R² | | | | | | | 0.0975 |
| Wald statistics (all parameters) | 41.05*** | 42.12*** | 46.40*** | 41.49*** | 47.77*** | 28.23*** | |
| Number of observations | 181 | 181 | 181 | 181 | 181 | 143 | 181 |
| Groups | 37 | 37 | 37 | 37 | 37 | 27 | |

Note: Standard Errors in parentheses. Dept. Var. = process patents
*p<0.1; **p<0.05; ***p<0.01.

**Table 8b: Average marginal effects**

| | Average marginal effects (AME) | | | | | | |
|---|---|---|---|---|---|---|---|
| | 8a | 8b | 8c | 8d | 8e | 9 | 10 |
| output | 0.0120**<br>(0.0057) | 0.0113**<br>(0.0054) | 0.0105**<br>(0.0048) | 0.0111**<br>(0.0055) | 0.0095**<br>(0.0046) | 0.0056*<br>(0.0031) | 0.0141**<br>(0.0058) |
| output² | -6.65e-06<br>(7.27e-06) | -6.69e-06<br>(7.07e-06) | -4.72e-06<br>6.95e-06 | -5.62e-06<br>(7.25e-06) | -3.93e-06<br>(6.84e-06) | -2.08e-06<br>(4.51e-06) | -0.0000<br>(0.0000) |
| **Control variables** | | | | | | | |
| global output | | 0.3432<br>(0.5166) | | | 0.2286<br>(0.4807) | -0.0499<br>(0.3435) | 0.6051<br>(0.6258) |
| patents before entry | | | 1.0685*<br>(0.5872) | | 1.0118*<br>(0.5570) | 0.5502<br>(0.5325) | 1.1831**<br>(0.4976) |
| group | | | | 0.4799<br>(0.6421) | 0.3933<br>(0.5686) | 0.0699<br>(0.5004) | 0.7226<br>(0.5342) |

Note: Standard Errors in parentheses. Dept. Var. = process patents
*p<0.1; **p<0.05; ***p<0.01.

### Table 9a: Effects of planned output growth on innovation (patents) in different model approaches

| | Random-effects negative binomial regression | | | | | | Conditional FE negative binomial regression | | Negative binomial regression | |
|---|---|---|---|---|---|---|---|---|---|---|
| | 11a | 11b | 11c | 11d | 11e | 11f | 12a | 12b | 13a | 13b |
| g_plan | 0.010*** | 0.007*** | 0.010*** | 0.009*** | 0.008*** | 0.008*** | 0.005** | 0.004* | 0.006** | 0.008** |
| | (0.002) | (0.002) | (0.002) | (0.003) | (0.002) | (0.003) | (0.002) | (0.003) | (0.002) | (0.003) |
| $output_{t-1}$ | | | | | | 0.001 | | 0.000 | | 0.005*** |
| | | | | | | (0.001) | | (0.001) | | (0.002) |
| **Control variables** | | | | | | | | | | |
| global output | | 1.120*** | | | 1.069*** | 1.031*** | 0.711* | 0.580 | 1.632*** | 0.957** |
| | | (0.330) | | | (0.322) | (0.340) | (0.420) | (0.522) | (0.386) | (0.400) |
| patents before entry | | | 0.537* | | 0.483* | 0.508 | 0.590 | 0.475 | 0.502* | 0.547* |
| | | | (0.315) | | (0.273) | (0.319) | (0.559) | (0.592) | (0.279) | (0.294) |
| group | | | | 0.363 | 0.205 | 0.238 | 0.576 | 0.784 | 0.597 | 0.229 |
| | | | | (0.383) | (0.348) | (0.347) | (0.596) | (0.633) | (0.371) | (0.384) |
| _cons | -0.715** | -0.925*** | -0.959*** | -0.956** | -1.359*** | -1.584*** | -1.230 | -1.339** | -0.248 | -0.367 |
| | (0.311) | (0.322) | (0.334) | (0.399) | (0.459) | (0.613) | (0.552) | (0.572) | (0.324) | (0.318) |
| Log likelihood | -234.34311 | -229.63849 | -232.90847 | -233.87604 | -228.06011 | -223.00865 | -129.61551 | -125.08481 | -232.73801 | -223.33971 |
| LR chi² | | | | | | | | | 42.04 | 50.30 |
| Pseudo R² | | | | | | | | | 0.0828 | 0.1012 |
| Wald statistics (all parameters) | 16.02*** | 33.10*** | 21.91*** | 16.70*** | 35.26*** | 27.45*** | 12.43** | 10.24* | | |
| Number of observations | 115 | 115 | 115 | 115 | 115 | 113 | 93 | 91 | 115 | 113 |
| Groups | 36 | 36 | 36 | 36 | 36 | 36 | 24 | 24 | | |

Note: Standard Errors in parentheses. Dept. Var. = patents
*p<0.1; **p<0.05; ***p<0.01.

### Table 9b: Average marginal effects

| | Average marginal effects (AME) | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|
| | 11a | 11b | 11c | 11d | 11e | 11f | 12a | 12b | 13a | 13b |
| g_plan | 0.0325** | 0.0236** | 0.0326** | 0.0305** | 0.0261** | 0.0268* | 0.0046 | 0.0039 | 0.0230* | 0.0435 |
| | (0.0153) | (0.0098) | (0.0138) | (0.0151) | (0.0107) | (0.0146) | (0.0029) | (0.0027) | (0.0132) | (0.0340) |
| $output_{t-1}$ | | | | | | 0.0033 | | 0.0002 | | 0.0264 |
| | | | | | | (0.0044) | | (0.0014) | | (0.0188) |
| **Control variables** | | | | | | | | | | |
| global output | | 3.5676** | | | 3.4839** | 3.6489* | 0.6836 | 0.5327 | 6.5889** | 5.4019 |
| | | (1.4726) | | | (1.4405) | (1.8952) | (0.4743) | (0.5126) | (2.8718) | (3.5620) |
| patents before entry | | | 1.7464 | | 1.5758 | 1.7976 | 0.5675 | 0.4361 | 2.0268 | 3.0881 |
| | | | (1.2490) | | (0.9916) | (1.4694) | (0.5970) | (0.5798) | (1.3380) | (1.9647) |
| group | | | | 1.2402 | 0.6695 | 0.8427 | 0.5536 | 0.7201 | 2.4117 | 1.2941 |
| | | | | (1.4562) | (1.1445) | (1.2536) | (0.6002) | (0.6309) | (1.5778) | (2.0698) |

Note: Standard Errors in parentheses. Dept. Var. = patents
*p<0.1; **p<0.05; ***p<0.01.

**Table 10a: Effects of planned output growth on process innovation (process patents) in different model approaches**

| | Random-effects negative binomial regression | | | | | | Conditional FE negative binomial regression | | Negative binomial regression | |
|---|---|---|---|---|---|---|---|---|---|---|
| | 14a | 14b | 14c | 14d | 14e | 14f | 15a | 15b[+] | 16a | 16b |
| g_plan | 0.011*** | 0.009*** | 0.012*** | 0.009*** | 0.010*** | 0.009*** | 0.006** | 0.005* | 0.006** | 0.009*** |
| | (0.003) | (0.003) | (0.002) | (0.003) | (0.002) | (0.002) | 0.003 | (0.003) | (0.003) | (0.003) |
| $output_{t-1}$ | | | | | | 0.001 | | 0.001 | | 0.005*** |
| | | | | | | (0.001) | | (0.002) | | (0.002) |
| **Control variables** | | | | | | | | | | |
| global output | | 0.773* | | | 0.705* | 0.669* | 0.402 | 0.177 | 1.288*** | 0.545 |
| | | (0.403) | | | (0.377) | (0.355) | (0.465) | (0.584) | (0.409) | (0.440) |
| patents before entry | | | 0.889** | | 0.838*** | 0.839*** | 0.864 | 0.603 | 0.666** | 0.782** |
| | | | (0.362) | | (0.324) | (0.291) | (0.738) | (0.776) | (0.324) | (0.325) |
| concern | | | | 0.679 | 0.579 | 0.598 | 1.027 | 1.329 | 1.000** | 0.622 |
| | | | | (0.473) | (0.433) | (0.430) | (0.797) | (0.914) | (0.436) | (0.457) |
| _cons | -0.704* | -0.849** | -1.190*** | -1.192** | -1.836*** | -2.198*** | -1.771** | -2.003** | -1.198*** | -1.355*** |
| | (0.395) | (0.430) | (0.420) | (0.515) | (0.582) | (0.517) | (0.803) | (0.893) | (0.404) | (0.403) |
| Log likelihood | -180.8308 | -179.15016 | -177.86481 | -179.72815 | -175.10107 | -170.29801 | -94.937906 | -91.089651 | -180.80134 | -172.19904 |
| LR chi² | | | | | | | | | 36.62 | 44.30 |
| Pseudo R² | | | | | | | | | 0.0920 | 0.1140 |
| Wald statistics (all parameters) | 15.99*** | 21.65*** | 28.89*** | 17.98*** | 39.65*** | 70.62*** | 10.59** | 9.21 | | |
| Number of observations | 115 | 115 | 115 | 115 | 115 | 113 | 79 | 77 | 115 | 113 |
| Groups | 36 | 36 | 36 | 36 | 36 | 36 | 20 | 20 | | |

Note: Standard Errors in parentheses. Dept. Var. = process patents
*p<0.1; **p<0.05; ***p<0.01.
[+]Note: In 15b the prob>chi2 = 0.1011

**Table 10b: Average marginal effects**

| | Average marginal effects (AME) | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|
| | 14a | 14b | 14c | 14d | 14e | 14f | 15a | 15b[+] | 16a | 16b |
| g_plan | 0.0198** | 0.0158** | 0.0207** | 0.0179** | 0.0181*** | 0.0184*** | 0.0066 | 0.0049 | 0.0134* | 0.0247 |
| | (0.0097) | (0.0071) | (0.0081) | (0.0089) | (0.0070) | (0.0063) | (0.0044) | (0.0037) | (0.0072) | (0.0181) |
| $output_{t-1}$ | | | | | | 0.0029 | | 0.0008 | | 0.0137 |
| | | | | | | (0.0026) | | (0.0015) | | (0.0095) |
| **Control variables** | | | | | | | | | | |
| global output | | 1.3804 | | | 1.2985 | 1.3661* | 0.4169 | 0.1607 | 2.7558** | 1.5703 |
| | | (0.8675) | | | (0.8224) | (0.8233) | (0.5390) | (0.5411) | (1.3571) | (1.4800) |
| patents before entry | | | 1.5846* | | 1.5440** | 1.7131** | 0.8955 | 0.5475 | 1.4251* | 2.2527* |
| | | | (0.8879) | | (0.7155) | (0.7243) | (0.9614) | (0.7721) | (0.8188) | (1.2759) |
| group | | | | 1.2823 | 1.0656 | 1.2206 | 1.0640 | 1.2063 | 2.1412** | 1.7920 |
| | | | | (1.0944) | (0.8377) | (0.9148) | (0.9450) | (0.9631) | (1.0642) | (1.3239) |

Note: Standard Errors in parentheses. Dept. Var. = process patents
*p<0.1; **p<0.05; ***p<0.01.

## References


Acs Z J, Audretsch DB (1987) INNOVATION, MARKET STRUCTURE, AND FIRM SIZE. The Review of Economics and Statistics 69(4): 567-574

Acs Z J, Audretsch DB (1989) Patents as a Measure of Innovative Activity. Kyklos 42(2): 171-180

Acs ZJ, Anselin L, Varga A (2002) Patents and innovation counts as measures of regional production of new knowledge. Res Pol 31: 1069–1085

Allison PD, Waterman R (2002) Fixed effects negative binomial regression models. Sociological Methodology, 32: 247-265

Anderson P, Tushman ML (1990) Technological discontinuities and dominant designs – a cyclical model of technological change. Administrative Science Quarterly 35: 604–633

Artés J (2009) Long-run versus short-run decisions: R&D and market structure in Spanish firms. Res Policy 38: 120-132

Audretsch DB (1991) New-firm survival and the technological regime. Review of Economics and Statistics 73: 441–450

Audretsch DB, Segarra A, Teruel M (2014) Why don't all young firms invest in R&D?. Small Bus Econ DOI 10.1007/s11187-014-9561-9

Bauckloh T (2012) Firmengröße und FuE-Aktivität – eine Analyse anhand von Patentdaten für die Photovoltaikbranche. Diplomarbeit I Univ. Kassel (not published).

Blankenberg A, Dewald U (2013) Public Policy and Industry Evolution: The Evolution of the Photovoltaic Industry in Germany. DRUID Conference Paper

Breschi S, Lissoni F, Malerba F (2003) Knowledge-relatedness in firm technological diversification. Res Policy 32(1): 69-87

Cefis E, Marsili O (2006) Survivor: The role of innovation in firms' survival. Research Policy 35: 626–641

Cohen WM, Klepper S (1996a) A reprise of size and R&D. The Economic Journal 106: 925-951

Cohen WM, Klepper S (1996b) Firm size and the nature of innovation within industries: The case of process and product R&D. The Review of Economics and statistics 78 (2): 232-243

Cohen WM, Levin RC, Mowery, DC (1987) Firm size and R&D intensity: A re-examination. J of Industrial Economics 35: 543–563

Cohen W, Levin R (1989) Empirical studies of innovation and market structure. Ch. 18. In: Schmalensee R, Willig R (eds) Handbook of industrial organization, II. North-Holland

Cohen WM (2010) Fifty Years of Empirical Studies of Innovative Activity and Performance. In: Hall BH, Rosenberg N (ed) Handbook of the Economics of Innovation Vol. I, Elsevier, Amsterdam, pp 129-213

Crépon B, Duguet E, Mairesse J (1998) Research, Innovation and Productivity: An Econometric Analysis At The Firm Level. Economics of Innovation and New Technology 7 (2): 115-158

Fisher FM, Temin P (1973) Returns to scale in research and development: What does the Schumpeterian hypothesis imply?, Journal of Political Economy 81: 56–70

Frenkel A, Shefer D, Koschatzky K, Walter GH (2001) Firm Characteristics, Location and Regional Innovation: A Comparison Between Israeli and German Industrial Firms. Regional Studies 35(5): 415-429

Fritsch M, Meschede M (2001) Product Innovation, Process Innovation, and Size. Review of Industrial Organization 19: 335-350

Gilbert R (2006) Looking for Mr. Schumpeter: Where are we in the competition-innovation debate?, Innovation Policy and the Economy 6: 159-215

Griliches Z (1990) Patent statistics as economic indicators: a survey. Journal of Economic Literature, 28: 1661-707

Hagedoorn J, Cloodt M (2003) Measuring innovative performance: is there an advantage in using multiple indicators?. Res Policy 32(8): 1365-1379

Hashi I, Stojčić N (2013) The impact of innovation activities on firm performance using a multi-stage model: Evidence from the Community Innovation Survey 4. Research Policy 42(2): 353-366

Hausman J, Hall BH, Griliches Z (1984) Economentric Models for Count Data with an Application to the Patents-R&D Relationship. Econometrica 52(4): 909-938

Henderson RM, Cockburn I (1996) Scale, scope, and spillovers: Determinants of research productivity in the pharmaceutical industry. RAND Journal of Economics 27: 32–59

Herrera L, Sánchez-González G (2013) Firm size and innovation policy. International Small Business Journal 31(2) 137-155, first published on January 27, 2012

Hilbe JM (2011) Negative Binomial Regression [e-book]. Cambridge University Press Textbooks, Available from: MyiLibrary http://www.myilibrary.com?ID=301603. Accessed 6 November 2013

Jovanovic B, MacDonald GM (1994) The life-cycle of a competitive industry. Journal of Political Economy 102: 322–347

Kemp RGM, Folkeringa M, de Jong JPJ, Wubben EFM (2003) Innovation and firm performance. Scales research reports. Zoetermeer: EIM business and policy research (downloaded from http://www.ondernemerschap.nl/pdf-ez/H200207.pdfon 27 December 2014).

Kleinknecht A (2000) Indicators of manufacturing and service innovation: their strengths and weaknesses. In: Metcalf JS; Miles I (ed) Innovation system and the service economy, Kluwer AP: Boston, pp 169-186

Klepper S (1996) Entry, Exit, Growth, and Innovation over the Product Life Cycle. Am Econ Rev 86(3): 562-583

Klepper S, Graddy E (1990) The evolution of new industries and the determinants of market structure. Rand Journal of Economics 21: 27–44

Klepper S, Simons KL (2005) Industry shakeouts and technological change. International Journal of Industrial Organization 23: 23– 43

Ma Z, Lee Y (2008) Patent application and technological collaboration in inventive activities: 1980–2005. Technovation 28(6): 379-390

Malerba F, Orsenigo L (1996) Schumpeterian patterns of innovation are technology-specific. Res Pol 25(3): 451–478

Markham JW (1965) Market structure, business conduct, and innovation. American Economic Review 55: 323–332

Nagaoka S, Motohashi K, Goto A (2010) Patent Statistics as an Innovation Indicator. In: Hall BH, Rosenberg N (ed) Handbook of Economics of Innovation Vol. II, Elsevier, Amsterdam, pp 1083 - 1127

Nickell SJ (1996) Competition and Corporate Performance. Journal of Political Economy 104(4): 724-746

Peltoniemi M (2011) Reviewing Industry Life-cycle Theory: Avenues for Future Research. International Journal of Management Reviews 13: 349-375

Photon (2001) Marktwachstum gewinnt an Fahrt - Solarzellenproduktion im Jahr 2000 stark gestiegen. Photon - Solarstrom Magazin 4: 12-14.

Photon (2002) Noch besser als erwartet - Der Weltmarkt für Solarzellen überrascht sogar Optimisten. Photon - Solarstrom Magazin, 4: 32-34.

Photon (2003) Ein ermutigendes Jahr - Marktübersicht zur weltweiten Solarzellenproduktion 2002. Photon - Solarstrom Magazin 4: 42-45.

Photon (2004) Sharp und der Rest der Welt - Markterhebung zur weltweiten Solarzellenproduktion 2003. Photon - Solarstrom Magazin 4: 18-22.

Photon (2005) Wachstum um die Wette - Die weltweite Solarzellenproduktion 2004 betrug über 1 Gigawatt. Photon - Solarstrom Magazin 4: 28-36.

Photon (2006) Siliziummangel – na und? - Marktanalyse der Solarzellenproduktion 2005. Photon - Solarstrom Magazin 4: 18-30.

Photon (2007) Die neue Maßeinheit heißt Gigawatt 2006 wurden 2,54 Gigawatt Solarzellen produziert. Photon - Solarstrom Magazin 4: 52-65.

Photon (2008) Und raus bist du Q-Cells hat Sharp als größten Zellhersteller abgelöst – die größte Herstellernation aber heißt China. Photon - Solarstrom Magazin 4: 24-38.

Photon (2009) Verhaltenes Lächeln auf langen Gesichtern - Die weltweite Zellproduktion hat sich 2008 auf 7,9 Gigawatt nahezu verdoppelt, doch die Zukunft ist ungewiss. Photon - Solarstrom Magazin 4: 54-71.

Photon (2010) Von wegen Krise - Solarzellenproduktion im Jahr 2009 wächst um 56 Prozent und übersteigt Zwölf-Gigawatt-Marke. Photon - Solarstrom Magazin 4: 38-65.

Photon (2011) Das Jahr des Tigers - 2010 wurden mehr Solarzellen produziert als in den vier vorherigen Jahren zusammen. Photon - Solarstrom Magazin 4: 38-71.

Photon (2012) Das Jahr des Drachen - Die Produktion von Solarzellen ist 2011 auf 37 Gigawatt gestiegen – und die Dominanz Chinas wächst", Photon - Solarstrom Magazin 4: 42-75.

Räuber A (2005) Photovoltaik in Deutschland. Eine wechselvolle Erfolgsgeschichte. In: Jansen S (ed) 30 Jahre DGS. Auf dem Weg in die solare Zukunft, Deutsche Gesellschaft für Sonnenenergie, München, pp 151-170

Rogers M (2004) Networks, firm size and innovation. Small Business Economics 22(2): 141-153

Scherer FM (1965a) Size of firm, oligopoly and research: A comment. Canadian Journal of Economics and Political Science 31(2): 256-266

Scherer FM (1965b) Firm Size, Market Structure, Opportunity, and the Output of Patented Inventions. The American Economic Review 55(5): 1097-1125

Schmookler J (1959) Bigness, Fewness, and Research. Journal of Political Economy 67: 628-635

Schumpeter JA (1942) Capitalism, Socialism and Democracy (2.rd ed.). Harper & Brothers Publishers, New York.

Segarra A, Teruel M (2014) High-growth firms and innovation: An empirical analysis for Spanish firms. Small Business Economics. doi:10.1007/s11187-014-9563-7

Shefer D, Frenkel A (2005) R&D, firm size and innovation: an empirical analysis. Technovation 25: 25–32

Soete LLG (1979) Firm size and inventive activity – The Evidence Reconsidered. European Economic Review 12: 319-340

Stock GN, Greis NP, Fischer WA (2002) Firm size and dynamic technological innovation. Technovation 22:537–549.

Suarez FF, Utterback JM (1995) Dominant designs and the survival of firms. Strategic Management Journal 16: 415–430

Tsai KH, Wang JC (2005) Does R&D performance decline with size? – A re-examination in terms of elasticity. Res Policy 34: 966-976

Utterback JM, Suárez FF (1993) Innovation, competition, and industry structure. Research Policy 22: 1-21

Vernon J, Gusen P (1974) Technical Change and Firm Size: The Pharmaceutical Industry. Review of Economics and Statistics 56(3): 294-302

Villard H (1958) Competition, Oligopoly and Research. Journal of Political Economy 66: 483-97

Worley JS (1961) Industrial Research and the New Competition. Journal of Political Economy 69: 183-86