Auburn University
Department of Economics
Working Paper Series

The Impact of R&D Cooperation on Drug Variety Offered on the Market. Evidence from the Pharmaceutical Industry

Tannista Banerjee[†] and Ralph Siebert[*]

[†]*Auburn University* and [*]*Purdue University*

AUWP 2013-21

The Impact of R&D Cooperation on Drug Variety Offered on the Market. Evidence from the Pharmaceutical Industry*

Tannista Banerjee† Ralph Siebert‡

August, 2013

Abstract

This study shows that R&D cooperations can be used as an instrument to coordinate drug development portfolios among participating firms, which has crucial implications on the number of drugs offered on the market. Our study puts special attention to the fact that R&D cooperations, formed at different stages throughout the drug development process, have different impacts on the technology and product markets. Using a comprehensive dataset on the pharmaceutical industry, our results show that R&D cooperations formed at the early stages increase the number of R&D projects and the number of drugs launched on the product market. Late stage R&D cooperations, however, have a positive impact on the drug development process and drug variety only in the short run. In the long run, late stage cooperations provoke that firms re-optimize their drug development portfolios which reduces the number of drugs offered on the market.

JEL: L24, L25, L65, D22.

Keywords: Drug development, Dynamics, Co-development, Pharmaceutical industry, Product variety, Product market competition, Research and Development cooperations.

*We thank seminar participants at Purdue University. We especially thank Jack Barron, Jeremiah Harris, Alberto Galasso and Stephen Martin for their helpful discussions and valuable suggestions. All errors are our own.

†Auburn University, Department of Economics, Address: 0326 Haley Center, Auburn, AL 36849, United States, Email: tzb0018@auburn.edu.

‡Purdue University, Department of Economics, Krannert School of Management, Address: 403 West State Street, West Lafayette, IN 47907-2056, United States, Email: rsiebert@purdue.edu.

1 Introduction

In many research intensive industries, the success of innovation determines the number of products offered on the market and market performance. For example, in the pharmaceutical industry, the treatment of diseases such as Alzheimers, Cancer, HIV, and Parkinsons hinges on the successful completion of drug development processes. A larger number of drugs on the market increases the likelihood of curing symptoms and diseases. This in turn has a significant impact on quality of life and life expectancy.

In the pharmaceutical industry, the drug development process is characterized by many obstacles, such as financial constraints and uncertainties. Research and development (R&D) cooperations (e.g., joint ventures, licensing agreements etc.) are an important instrument to overcome those innovation impediments. They allow firms to exploit synergy effects and to share R&D costs, which are considered as a fixed cost to enter product markets[1]. Previous literature has shown that lower fixed costs increase the number of products offered on the market, see e.g., Salop (1979) and Lancaster (1979). It is reasonable to conjecture that the benefits in R&D caused by R&D cooperations imply a larger number of products offered on the market.

Policy debates often raise concerns about potential anti-competitive collusive implications of R&D cooperations[2]. In fact, policy makers devote special attention to the question if R&D cooperations result in price fixing agreements[3]. Seminal empirical studies, e.g., Goeree and Helland (2012), Roeller et al. (2013) as well as Suetens (2008), investigate if

[1]Prominent empirical studies on research cooperations focus on the impact of R&D cooperations on R&D investments, see for example Sakakibara (2002), Cassiman and Veugelers (2002), Irwin and Klenow (1996), and Roller, Siebert and Tombak (2007), Nicholson et al. (2003), Higgins and Rodriguez(2006), Arora et al. (2009), Nicholson et al. (2003) and Grabowski and Margaret (2012) among others. The majority of studies finds that R&D cooperations are an appropriate instrument to overcome financial impediments and have pro-competitive effects on R&D investments. These findings encouraged many countries to establish research programs that support the formation of research cooperations, mainly to keep up with increased international competition in R&D. Examples are the National Cooperative Research Act enacted by the U.S., the MITI by Japan, as well as EU framework programs, ESPRIT and EUREKA in Europe, among many others.

[2]See also Jorde and Teece (1990) and Shapiro and Willig (1990).

[3]Many cartels were formed in the past in the pharmaceutical industry. Perhaps the best known price fixing case in the pharmaceutical industry involved Hoffmann-La Roche Ltd and BASF which plead guilty in 1999. In a worldwide conspiracy the parties fixed prices and allocated sales. The companies paid a fine of close to 1 billion U.S. dollars. See also http://www.justice.gov/opa/pr/1999/May/196at.htm for further information on this case.

R&D cooperations facilitate price fixing behavior in the product market[4].

However, price fixing is one example where cooperating firms in R&D cause a potentially harmful impact on the product market, and there are other alternatives that deserve empirical attention. For instance, R&D collaborators may re-optimize their drug development portfolio and discontinue some of their R&D projects, either willingly or due to the pressure from cooperators. One well known example involves Tanox, Novartis and Genentech. In 1996 Tanox, Novartis and Genentech reached a R&D cooperation agreement to select and develop some anti-Immunoglobulin E antibodies previously identified and synthesized by the parties. As a result of the collaboration they identified XOLAIR as the potential drug for the treatment of peanut allergy and discontinued the clinical trials for other anti-Immunoglobulin E antibodies including CGP 56901, a potential drug for the same disease. The collaborators decided to discontinue the development of other anti-Immunoglobulin E antibodies but Tanox wanted to develop CGP 56901 independently because Tanox thought that it was a good potential molecule. Novartis and Genentech forced Tanox to cease the development of CGP 56901 through court orders arguing that the development violated their R&D cooperation agreement. In 2003, FDA approved XOLAIR for treating severe allergy and asthma[5]. Moreover, in response to jointly developed drugs cooperating firms have the opportunity to re-optimize their drug development portfolio, alter the objectives, and terminate some of their R&D projects, which has an impact on the drug variety launched on the market. Firms have an interest to discontinue the development of their existing drugs, which are close substitutes to the newly developed new drug, as they cannibalize the potential sales of the jointly developed drug. To avoid this anticipated business stealing effect, firms re-optimize their drug development portfolios and terminate closely related drugs in development. Consequently, the total number of drugs offered on the market declines. Hence, an improved drug development process can be offset by firms re-optimizing their drug development portfolio at the later development stages. Therefore, conclusive inferences about the ultimate impact of R&D

[4]Martin (1995) and Cabral (2000) provide theoretical evidence of tacit product market collusion arising from R&D cooperations.

[5]Please see Sammi (2006) for a detailed discussion.

cooperations on the drug variety offered on the market have to account for the fact that firms respond to jointly developed drugs and re-optimize their drug development pipeline.

At present, very little is known about the impact of R&D cooperations on drug development projects and the variety of drugs offered on the market. The aim of this study is to investigate the impact of R&D cooperations on the drug development process and drugs offered on the market. Our study contributes to the question if R&D cooperations have potential anti-competitive implications for product market level of activity.

Our study accounts for an inherently dynamic R&D process, as R&D projects may take several months or years to be completed. The drug development process is characterized by several phases which need to be successfully completed before a drug receives permission by the Food and Drug Administration to be launched on the market.

It is common in the pharmaceutical industry to separate the drug development process into two distinct stages, i.e., the early (Discovery, Lead Molecule or Pre-Clinical stage) and the late stage (development and testing), see e.g., DiMasi et al. (1991 and 2003)[6]. The requirements, aims and scope differ for early and late research stages. Early stage projects focus on rather basic research questions, e.g., finding a new molecule structure[7]. Late stage projects concentrate on empirical testing of the drug on different populations (applied research), involving more human subject management, drug safety testing, and passing clinical trials in order to launch new drugs on the markets.

Since different strengths and expertise are required throughout different research stages, firms incentives to join R&D cooperations at the early stages differ from those at the late stages. Note, that R&D cooperations can be performed at the early stages and/or at the later stages[8]. Consequently, it is reasonable to assume that R&D cooperations formed at different stages also have different impacts on technology and product markets. As of now, few theoretical studies consider the fact that research cooperations can be formed

[6]Table 1 provides a list of different research phases.

[7]See Table 1 in the Appendix and Section 3 for a more detailed description of the industry and the research phases.

[8]For example, R&D cooperations can be formed only at the late stages, after the R&D projects passed the early stages with individual research efforts. Table 3 shows the number of R&D cooperations formed at different stages.

at different stages of the development process, see e.g., Grossman and Shapiro (1986 and 1987). We are not aware of any empirical studies which elaborate on the incentives to form R&D cooperations at different stages throughout the R&D process and, more importantly, to evaluate the different impacts on technology and product markets.

Our study focuses on the pharmaceutical industry as it characterizes a natural object to focus on the dynamic R&D process and to investigate the impact of R&D cooperations on product variety for several reasons. First, drugs must successfully pass multiple clinical or research stages before being launched on the market. This enables us to track the dynamics of the R&D process and the timing of engaging into R&D cooperations. Second, new drug applications must be approved by the Federal Drug Administration. Therefore, the number of new products introduced into the market is publicly documented, which allows us to establish a reliable database on the number of products in the market. The study uses firm-level information on R&D cooperations, firms' research agendas and their drugs offered on the market from 1993-2011. In evaluating the impact of R&D cooperations, we face a missing data problem, since we only observe firms that either cooperated or did not cooperate, but not both. To overcome this problem, we apply a switching regression framework. Hence, we estimate the 'counterfactual' effect of how R&D cooperations impact a firms activity in the technology and product market, compared to performing R&D individually, and vice versa. We estimate a heterogeneous treatment effect model, adopting the estimator suggested by Heckman, Urzua, and Vytlacil (2006)[9].

Previous literature on estimating treatment effects pointed out two associated potential biases, i.e., the pre-treatment heterogeneity bias or selection bias, and the heterogeneous treatment effect bias[10].

The pre-treatment heterogeneity refers to firms self-selecting themselves into early/late stage R&D cooperations based on firm-level attributes. In our case, it relates to the fact that observed firm-level attributes, such as technology and product market level of

[9]Estimators such as fixed effects and difference in difference eliminate the selection bias, but not the treatment effect heterogeneity bias (Angrist and Krueger, 1999).

[10]See also Angrist and Krueger (1999), Heckman, Urzua and Vytlacil (2006), Morgan and Christopher (2007), Dehejia and Wahba (2002), and Brand and Xie (2010), Brand and Thomas (2012) and Pais (2011).

activity, determine firms' decision to form R&D cooperations. In addition, we account for unobserved firm-level attributes such as managerial ability, absorptive capacity, etc., which affect firms' decision to participate in R&D cooperations. Finally, we account for the fact that firms' participation in R&D cooperations also has an expected impact on future positions in the technology and product markets. Ignoring this self-selection incentive results in a potential selection bias. For example, R&D cooperations seemingly appears to increase technology and product market level of activity, simply because cooperating firms were already more research and production intensive than non-cooperating firms, before they selected themselves into R&D cooperations. Hence, we apply an identification strategy based on instrumental variables to control for a potential self selection bias[11].

The heterogeneous treatment effect bias relates to the fact that firms sort themselves on an expected profit gain to improve their positions in the technology and product markets. For example, firms with a strong R&D market presence are able to benefit more from additional knowledge due to higher absorptive capacity. Moreover, production intensive firms might achieve higher gains in launching new drugs. Hence, we allow the impact of R&D cooperations to vary across firms, applying the heterogeneous treatment effects model by Heckman, Urzua, and Vytlacil (2006).

Our results show that early and late stage R&D cooperations differ in their impact on the technology and product markets. More specifically, early stage cooperations allow firms to benefit from synergy effects in their R&D activity, which increases by 21%. The gains in the technology markets will be transmitted to the product market and increases the number of drugs offered on the market. We also find that the impacts on technology and product markets increase over time. Turning to the results on the impact of late stage R&D cooperations, we find only a minor improvement in firms' activity in the technology market, which is not sustainable over time. Most importantly, our study shows that firms re-optimize their drug development portfolios and eliminate existing drug development projects which reduces the number of drugs offered on the product markets. Our results

[11]It is important to account for firms incentives to collaborate depending on their technology and product market positions in order to thoroughly investigate the impact of R&D collaborations on technology and product market level of activity, see e.g. Roller, Siebert and Tombak (2007).

also confirm that firms select themselves into early and late stage cooperations depending on their level of activity in the R&D and the product market. While early stage R&D cooperations are formed among less research intensive firms, late stage R&D cooperations are more common for firms, which are more active in technology and product markets.

The remainder of the paper is organized as follows. The next section surveys the relevant literature. In Section 3, we provide an industry description. Section 4 presents a description of the data sources and the variable definitions. Section 5 discusses the empirical model and the estimation procedure. Section 6 presents the empirical results. We conclude in Section 7.

2 Literature overview

Research cooperations have frequently been analyzed in the literature and are considered to have socially beneficial impacts such as internalizing research externalities or avoiding wasteful duplications in research. In the pharmaceutical industry, several studies showed that R&D cooperations increase the likelihood of developing new drugs. For example, Arora et al. (2009), Danzon et al. (2005) and Nicholson et al. (2003) show a positive effect of R&D collaborations on clinical trial successes. Powell et al.(1999) shows a positive effect of R&D collaboration on R&D and the production of patents. Lerner and Tsai (2000) show that R&D collaborations generate more approved drugs under more favorable financing conditions. In a related context, Ornaghi (2009) studies the effects of mergers in the pharmaceutical industry on firms' R&D activity and finds that merged companies reduce their R&D activity.

As mentioned above, drug development projects underly an inherently dynamic process. A new drug must successfully pass multiple phases before it receives permission by the Food and Drug Administration to be launched on the product market. Throughout the research process, companies invest enormous amounts in scientific knowledge. Empirical studies in the pharmaceutical industry highlight the relevance of economies of scope and scale throughout different stages of the R&D process, see e.g., Henderson and Cock-

burn (1996 and 2001) and Nesta and Saviotti (2005). Economies of scale are important features in the drug discovery or early stages of the R&D process. In contrast, scope economies are relevant in the drug development or late stages, see Henderson and Cockburn (1996 and 2001). To date, only a few empirical studies account for the timing of forming research cooperations throughout the R&D process. For example, Siebert and von Graevenitz (2012) investigate firms incentives to form a licensing agreement at the beginning of the research process or to exchange their inventions after the research has been completed unilaterally. Oxley and Sampson (2004) and Erkal and Minehart (2008) have shown that firms should be hesitant to form R&D cooperations at the late stages. Lerner and Merges (1998) focus on partner matches between firms in forming research cooperations. They find that smaller firms frequently have incentives to form cooperations with larger companies at the early stages of the drug development process. Small biotechnology companies frequently lack resources to complete the entire R&D process for the innovation of a new drug. They develop a new drug molecule and search for sponsors to complete the development process (Powell and Brantley, 1992).

In a related context, the theoretical work by D'Aspremont and Jacquemin (1988) shows that Research Joint Ventures (RJVs) spend more on R&D and generate higher profits if technological spillovers are sufficiently high; see also Spence (1984) and Katz (1986) for similar results. The study by Kamien, Muller, and Zang (1992) shows that firms' investments and their incentives to engage in research cooperations depend on the degree of spillovers and their product relatedness. For example, if firms face high spillovers in the technology market, but operate in different product markets (i.e., their products are totally differentiated), cooperations lead to an increase in R&D investments. The reason is that information spilling towards a rival does not encompass any competitive externality on a firms profits as firms operate in different product markets. Hence, firms do not have to be concerned that rivals gain a free ride on their own investments which results in lower profits via business stealing. In contrast, if firms offer similar products on the market, i.e., if the degree of product differentiation is low, firms exert negative externalities o

neach other via business stealing which lowers their profits. Gugler and Siebert (2007) consider the interdependence in technology and product markets to explain the formation and impact of mergers on firm and market performance in the semiconductor industry.

3 Industry description

The pharmaceutical industry is a research intensive industry. Pharmaceutical companies face a permanent pressure to discover new drugs for improving the quality of human life and to target diseases such as AIDS, cancer, Alzheimers etc. The associated returns from launching a new drug are frequently extraordinarily high, which increases firms incentives to receive a patent on a specific drug. Moreover, an increased global competition puts more pressure on firms to develop drugs in a shorter time period. As a result, firms invest an enormous amount of money in research and development.

In order to launch a new drug on the market, the drug must successfully pass seven stages before it is granted an approval by the Food and Drug Administration (FDA), see Table 1 in the Appendix. The "Discovery" and "Lead Molecule or Pre-clinical" stages, involve research on identifying potential molecules. The new compound is also extensively tested for toxicity in animals. In "Phase I" the compound is tested for safety on healthy human volunteers; in "Phase II" the drug is tested on a small group of patients to establish efficacy; and in "Phase III" the compound is tested on a larger number of more diverse patients to establish both safety and efficacy.

The literature separates the drug development process into two distinct stages, i.e., the early and the late stage[12]. The requirements, expertise, aim and risks of the development process are fundamentally different between early and late stages. The aim of the drug development at the early stage is to explore a compound that demonstrates some desirable effect on either an animal or chemical screen. For example, firms search for compounds that can make obese rats thinner or that can block the action of an enzyme that is known

[12]See also the description of the drug development guidelines provided by FDA http://www.fda.gov/Drugs/DevelopmentApprovalProcess/default.htm.

to regulate metabolic activity. In contrast, the aim of the late stage drug development process is to explore the degree to which a particularly promising compound is safe and effective for humans.

After completing the development stages, the firm submits its new drug application (NDA) to the Food and Drug Administration (FDA), which eventually decides on approval of the application. The FDA requires evidence of a drug's effectiveness, through thoroughly-controlled clinical investigations. Pharmaceutical firms receive Intellectual Property Rights in the form of patents for the drugs introduced into the product market; the length of patent protection lasts on average 15 years. Table 2 presents the number of new drug approvals between 1993 and May 2011. NDA approvals vary a lot between different years and declined in the last few years.

The pharmaceutical industry is divided into 18 therapeutic areas (technological markets), see Table 3 in the Appendix. The disease areas (product markets) are also divided into same 18 different markets. The classifications on the therapeutic and disease areas are commonly used in the health literature, the Food and Drug Administration and BioPharm Insight. Table 4 shows that the maximum number of technology ($Tech$) and disease ($Product$) market areas in which pharmaceutical companies are active are 18 and 14, respectively. The technology market level of activity (TMA) measures a firm's research activities in different therapeutic areas[13]. Similarly, the product market level of activity (PMA) represents a firm's strength in different disease areas.

Firms frequently face innovation impediments due to lacking specific research capabilities and knowledge. R&D cooperations are considered an appropriate instrument for firms that enables them to overcome such innovation impediments. Our dataset shows that 3,756 R&D cooperations took place between 1993 and May 2011, see Table 3 in the Appendix[14]. The table shows the sharp discrepancies in the number cooperations between technology areas. For example, the technology area cancer has the largest number of co-

[13]The variable definitions are explained in more detail in Section 4.

[14]The information is taken from our established database, which contains firm-level information on R&D cooperations, drugs under development and drugs offered on the market. All information is distinguished by technology and disease areas. Details on the data sources can be found in Section 4

operations (781), followed by infectious diseases (404), which includes HIV. One reason for finding a large variation in cooperations between technology areas is that the areas are characterized by different financial requirements, different technological spillovers, and different relationships between firms in the disease areas.

Firms engage in R&D cooperations at various stages of the drug development process. As an example the Takeda Pharmaceutical Company Limited signed an early stage R&D cooperation with XOMA Ltd., a biotechnology company from California, for monoclonal antibody development in November, 2011. From this cooperation XOMA Ltd. overcame the financial impediments and both XOMA Ltd. and Takeda Pharmaceutical Company benefited from the expansion of product variety in the oncology market[15]. In formulating firms' decisions on cooperating in R&D, it is important to note that the R&D process in the pharmaceutical industry is characterized by a high degree of uncertainty and is difficult to foresee. For example, several drug development projects such as Viagra, Cialis and Aspirin had a different original goal, and the uncertainty in R&D eventually changed the application of the drugs. Moreover, the uncertainty in R&D, as well as the different aims and requirements of research projects are different throughout the research process (DiMasi et al., 2010). Hence, R&D cooperations are rather formed independently throughout the research process, i.e., firms make separate decisions whether to form an early stage or late stage R&D cooperation. For example, after firms individually invested effort in a specific research project at the early stage of the development process, they might decide at the late stage of the development process to form a (late stage) R&D cooperation. Hence, in this case a late stage cooperation does not necessarily require the cooperation at the early stages.

Table 5 shows the descriptive statistics of firm characteristics that participated in early and late stage R&D cooperations. It is frequently claimed that R&D cooperations are associated with enormous transactions costs, which involve specifically designed contracts, managerial expertise, lawyers, partner matching etc. Since early and late stage R&D cooperations are different in nature, they are also characterized by specific organizational,

[15]More examples of research cooperations formed at different stages are shown in the Appendix.

contractual and managerial efforts. Hence, the associated transaction costs are specific to the types of cooperations. We account for the fact that transaction cost decline as firms accumulate more experience from participating in early and late stage R&D cooperations, see also Siebert and von Graevenitz (2010).

4 Data sources and variable definitions

Since the main interest of our empirical study is to estimate the impact of early and late stage R&D cooperations on firms' level of activity in the technology (therapeutic) and product market (disease) areas, we use detailed firm-level data on the Biotechnology and Pharmaceutical industry over time. Our database is constructed from a variety of different data sources, of which the main part is provided by BioPharm Insight. In the following, we provide a thorough description of the data sources and the definitions of variables we use in our empirical model[16].

4.1 Early/late stage R&D cooperations: d_{irt}^{Y}

Our study evaluates the associated impact of R&D cooperations on the drug development process as well as the number of drugs offered on the market. In our context, a proper evaluation requires a distinction between different types of R&D cooperations, and to concentrate on those cooperations that allow for potential synergy effects in research. Therefore, we concentrate our study on (ex ante) R&D cooperations, in which firms jointly invest and work in R&D[17]. Consequently, we exclude R&D cooperations in which only one company discovers a new drug molecule without gaining any synergy effects from collaborators, such as (ex post) R&D cooperations (already discovered inventions are transferred to other firms) and product development deals (R&D activities are outsourced to other firms).

[16]Further information on the data can also be found on BioPharm Insight's website, see http://www.biopharminsight.com/biopharm_insight.html.

[17]In (ex ante) R&D cooperations firms decide on joining R&D cooperations before the invention has been made.

Our database on R&D cooperations is provided by the BioPharm Insight research cooperation, which collected the original information from the U.S. Securities and Exchange Commission filings, a global network of journalists and expert industry research analysts. As mentioned above, we account for dynamics in the drug development process and firms' decision when to join R&D cooperations. Firms decide whether to engage into an early stage R&D cooperation, a late stage R&D cooperation, or not to engage in either one. Our database contains information on which firms formed R&D cooperations at what time, at which research stag [18]. The R&D cooperations are classified into 18 different drug development (or therapeutic) areas and disease areas.

We formulate dummy variables $[d^{Y}_{irt}, (Y = Early, Late)]$, which take on a value of 1 if firm i formed a R&D cooperation at the early or late stage of the drug development process, in a certain therapeutic area (r), in time period t. Otherwise, the dummy takes on a value of zero.

Early/late stage experience: EXP^{Y}_{irt}

As mentioned above, we account for transaction costs associated to early and late stage cooperations. We use a proxy, the number of past early and late stage R&D cooperations, to account for firms' experience[19]. Hence, firm i's experience in signing early and late stage cooperation agreements in therapeutic market r in year t is defined as,

$$EXP^{Y}_{irt} = \sum_{s=1}^{t-1} Y_{irs};\ ,\ (\text{Y=Early, Late})$$

where Y_{irs} is the number of early stage or late stage cooperations a firm signed in year s.

4.2 Technological market level of activity: TMA_{irt}

Firms' strength of research capabilities vary across different therapeutic areas. We account for firms' expertise in different therapeutic areas and establish a measure, technological

[18] Changes in company names over time via mergers and takeovers are corrected for.

[19] See also Siebert and von Graevenitz (2010).

market level of activity, that evaluates how a firm's technology or drug market portfolio evolves over time,

$$TMA_{irt} = \widetilde{z}_{irt} + \sum_{s=1}^{t-1}(1-\delta)^s \widetilde{z}_{irt}.$$

The measure accounts for the number of research projects ($\widetilde{z}_{irt}$) undertaken by firm i in therapeutic area (r) in period t. Since research in the pharmaceutical industry is a highly paced process, the value of research knowledge quickly depreciates over time. To account for the depreciation of a firm's knowledge base over time, we discount a firm's drug portfolio over time, and apply a perpetual inventory method with a depreciation rate of $\delta = 0.15\%$, see also Griliches and Mairesse (1984), Hall (1993), etc.

4.3 Product market level of activity: PMA_{irt}

Similar to the therapeutic market level of activity, pharmaceutical firms have different strengths and expertise in launching new drugs in different disease or product market areas. We construct a measure called product market level of activity to account for firms' expertise in different disease areas (r) over time. The measure refers to the number of approved drugs a firm launched in different disease areas, and is specified as follows[20]:

$$PMA_{irt} = \widetilde{p}_{irt} + \sum_{s=1}^{t-1}(1-\delta)^s \widetilde{p}_{irt}.$$

The total number of approved drugs for firm i in disease area (r) in period t is denoted as the product market portfolio $\widetilde{p}_{irt}$. Since a drug is offered on the market for several years, its market value as well as the experience associated with launching a new drug depreciates over time. Again, we apply the perpetual inventory method with a depreciation rate of $\delta = 0.15\%$.

[20]Information about the name of the approved drug, the approval date, the company name and the disease area is released by the FDA on the Drugs@FDA webpage, see http://www.accessdata.fda.gov/scripts/cder/drugsatfda/index.cfm. BioPharm Insight collects this drug approval data from the FDA. See also Table 3 for the number of drug approvals by year.

4.4 Diversification in product and technology markets: $Product_{it}$ and $Tech_{it}$

Firms have the opportunity to absorb new knowledge from forming R&D cooperations. Firms being active in multiple therapeutic areas have the opportunity to transfer their knowledge gained in one therapeutic area to another therapeutic area, and benefit from a larger scope in R&D, which increases their incentives to form R&D cooperations. We control for the scope in R&D and construct a variable ($Tech_{it}$), which refers to the number of different therapeutic areas in which firm i operates. According to the definition of the therapeutic areas in Table 2, the variable can take on a value between 1 and 18. A higher value refers to a larger scope in R&D and higher potential technological spillovers.

We apply the same rationale to the product market. Firms operating in several disease areas have a higher incentive to engage in R&D cooperations as the innovations can be applied to drug testing in different disease areas. We establish a variable that accounts for firms' multimarket character. The variable ($Product_{it}$) indicates the number of different disease areas in which a firm operates[21].

5 The model and estimation

In this section, we introduce the theoretical underlying framework to estimate the impact of early and late stage R&D cooperations on technology and product market level of activity. We begin with firms' decisions to join an early stage R&D cooperation, or not[22].

[21]Finally, note that financial information taken from firms' balance sheet information is a potential variable that could have been included in our empirical study. The drawback of including this information, however, is that we had to condition our empirical study on the fact that firms are public and we would lose the majority of observations. We therefore control for firm-level fixed effects to capture the financial aspects in our model.

[22]The same rationale applies to a firm's decision to engage in a late stage R&D cooperation. Remember, that the two types of cooperations are characterized by different objectives and therefore are not considered to be substitutes for one another. Moreover, as argued above, it is reasonable to assume that the decisions to form early and late stage cooperations are independent.

Let V_{irt}^{*} be the present value of firm $i = 1, ..., N$ in market r in period t:

$$V_{irt}^{*} = \widetilde{\alpha} PMA_{irt} + \widetilde{\beta} TMA_{irt} + \epsilon_{irt},$$

where PMA_{irt} and TMA_{irt} refers to the product and technology market level of activity, respectively, and ϵ_{irt} is an i.i.d. normally distributed error term with mean zero. The present value of a firm that joined an early stage R&D cooperation (indexed by superscript 1), is given by:

$$V_{irt}^{1*} = \alpha_1 PMA_{irt}^{1} + \beta_1 TMA_{irt}^{1} + \epsilon_{irt}^{1}.$$

The present value of a noncooperating firm (indexed by superscript 0), is:

$$V_{irt}^{0*} = \alpha_0 PMA_{irt}^{0} + \beta_0 TMA_{irt}^{0} + \epsilon_{irt}^{0}.$$

We can infer that a firm joins an early stage R&D cooperation, if $V_{irt}^{1*} > V_{irt}^{0*}$. Note, that we don't observe V_{irt}^{1*} or V_{irt}^{0*}, but we observe if a firm forms an early or late stage R&D cooperation. Hence, a firm's cooperation decision is based on the following equation:

$$V_{irt}^{*} = V_{irt}^{1*} - V_{irt}^{0*} = \alpha(PMA_{irt}^{1} - PMA_{irt}^{0}) + \beta(TMA_{irt}^{1} - TMA_{irt}^{0}) + \epsilon_{irt}^{*} \tag{1}$$

where $\epsilon_{irt}^{*} = \epsilon_{irt}^{1} - \epsilon_{irt}^{0}$. As shown in the equation, firms make their decisions to form a R&D cooperation based on their impact in the technology and product markets. Ideally, we would be interested in estimating the hypothetical heterogeneous treatment effect for firm i on the product market:

$$\alpha_i = PMA_{irt}^{1} - PMA_{irt}^{0},$$

and on the technology market:

$$\beta_i = TMA^1_{irt} - TMA^0_{irt}.$$

Note, that we observe at most only one of the two outcomes, PMA^1_i or PMA^0_i, but not both, i.e., the fundamental problem of causal inference, see Holland (1986)[23]. Hence, identifying a direct treatment effect is beset with a missing data problem and deriving causal inferences is not feasible at the individual level.

We proceed with performing a comparison at the group level, and decompose the PMA^1 into its mean $\mu^1(X)$ given regressors X, and its deviation from the mean u^1.

$$PMA^1 = \mu^1(X) + u^1.$$

A similar decomposition is applied to PMA^0:

$$PMA^0 = \mu^0(X) + u^0.$$

We assign a treatment variable denoted by the binary variable d^Y_i, Y=Early (Late), which takes on a value of one if firm i joined an early (late) stage R&D cooperation, respectively. Since firms are observed only if $d^Y_i = 1$ or $d^Y_i = 0$, we use a switching regression framework, and obtain for the observed outcome

$$PMA = d^Y PMA^1 + (1 - d^Y)PMA^0.$$

Substituting the above equations, yields

$$PMA = PMA^0 + d^Y(PMA^1 - PMA^0) = \mu^0 + d^Y[\mu^1(X) - \mu^0(X) + u^1 - u^0] + u^0,$$

where the second term expresses the benefit of participation. The first component $\mu^1(X) -$

[23]The following outline refers to the product market level of activity, PMA. Note, that they also apply to the technology market level of activity TMA.

$\mu^0(X)$ measures the average gain to a firm described by characteristics X. The second component $u^1 - u^0$ is the individual-specific benefit. Replacing for $\mu^0 = g_0(X) + e_0$ and $\mu^1 = g_1(X) + e_1$ and using regression notation, we get

$$PMA_i = \gamma X_i + \alpha d_i^Y + g_0(X_i) + d_i^Y(X_i - \mu^x)b + \epsilon_i. \tag{2}$$

where $\gamma X = \mu^0(X)$, and $\alpha = (PMA^1 - PMA^0) = \mu^1(X) - \mu^0(X) + e^1 - e^0$ and $\epsilon = e^0$. The parameter γ is a regression coefficient measuring the changes in PMA associated with the changes in the firm characteristics X. The coefficient α represents the change in PMA associated with the R&D cooperation. Even though our goal is to estimate a heterogeneous treatment, α_i, for simplicity, we briefly consider a homogeneous treatment effect, i.e., $\alpha = PMA^1 - PMA^0$ is the same for all firms. In this restricted case, a least square regression (a mean difference between cooperating and noncooperating firms) is subject to a potential selection bias or pretreatment heterogeneity bias. This bias is due to a non-zero correlation between d_i^Y and ϵ_i, see also Griliches (1977). It is reasonable to assume that stronger positions in the technology and/or product markets will directly affect future technology and product market positions, as well as participation in R&D cooperations. Therefore, we have to account that firms self-select themselves into early or late stage R&D cooperations dependent on potential outcomes in the therapeutic and product market. Regarding the self-selection problem, we also need to account for the fact that firms are characterized by different attributes, some of which are unobserved. Hence, conditional on observed covariates there are unobserved factors that are associated with the participation and potential outcomes on the therapeutic and product markets[24]. Since the ignorability assumption ignores heterogeneity due to unobserved variables, we apply an identification strategy based on instrumental variables.

Finally, even after correcting for the selection bias or pretreatment heterogeneity bias,

[24]This implies that we cannot keep the conditional independence assumption or the ignorability assumption, which states that outcomes are uncorrelated with treatment status(or engaging into R&D cooperations), conditional on observed covariates. For applications based on the conditional independence assumption using propensity score estimation methods, see e.g., Brand and Xie (2010), Brand and Thomas (2012) and Pais (2011)

we have to be aware of a potential treatment heterogeneity effect bias, which is based on $Cov(\alpha, d) \neq 0$, see also Heckman et al. (2006). This is a crucial concern in our application as firms are sorting themselves on the gain to improve their positions in the technology and product markets. Our study relaxes the homogeneity assumption and allows the responses of the treatment $\alpha_i = PMA_i^1 - PMA_i^0$ to vary across firms (heterogeneity) applying the heterogeneous treatment effects model, as will be further discussed below.

5.1 Empirical model specification

Our empirical model is based on equations (1) and (2). The regression equations of main interest measure the impact of early and late stage R&D cooperations (d_{irt}^Y) on the technology and product market. Turning to the product market,

$$PMA_{irt+j} = \alpha_1 + \alpha_2 d_{irt}^Y + \alpha_3 PMA_{irt} + \alpha_4 TMA_{irt} + \alpha_5 d_{irt}^Y * (PMA_{irt} - mean(PMA_{irt}))$$

$$+\alpha_6 CT_{it}^1 + \alpha_7 CT_{it}^0 + \sum_{i=8}^{27} \alpha_t Time_{it} + \nu_{irt} \quad , \tag{3}$$

where $j = 1, 2$; $Y = Early, Late$. Note, that we account for different time lengths $j = 1, 2$ and also control for time-specific fixed effects using time dummies ($Time$).

An equivalent equation is estimated to evaluate the impact on the technology market:

$$TMA_{irt+j} = \beta_1 + \beta_2 d_{irt}^Y + \beta_3 PMA_{irt} + \beta_4 TMA_{irt} + \beta_5 d_{irt}^Y * (TMA_{irt} - mean(TMA_{irt}))$$

$$+\beta_6 CT_{it}^1 + \beta_7 CT_{it}^0 + \sum_{i=8}^{27} \beta_t Time_{irt} + \xi_{irt} \ . \tag{4}$$

The coefficients of main interest are α_2 and β_2 in the outcome equations (3) and (4) for cooperation-signing and no- cooperation signing companies, while controlling for heterogeneity, i.e., α_5 and β_5.

We apply a heterogeneous treatment effects model (see also Heckman, Urzua, and Vytlacil (2006)), to control for the two potential biases, i.e., (i) the pre-treatment heterogeneity bias or selection bias , and (ii) the heterogeneous treatment effect bias[25]. As firms select themselves into R&D cooperations based on anticipated gains and costs of participation, an ordinary least square regression may result in an upward bias on α and β. We correct the outcome equations using selection terms, CT_1 and CT_0, which account for firms' self-selection[26]. They account for the fact that unobserved firm-level attributes might strengthen firms positions in the technology and product markets, and also increase the likelihood to form R&D cooperations[27]. We derive the selection terms by estimating the likelihood that a firm i will sign an early/late stage R&D cooperation in market r in year t:

$$d_{irt}^{Y*} = \gamma_1 PMA_{irt} + \gamma_2 TMA_{irt} + \gamma_3 EXP_{irt}^{Early} + \gamma_4 EXP_{irt}^{Late} + \gamma_5 Tech_{it} + \gamma_6 Product_{it}$$

$$+\gamma_7 mean(PMA_{irt}) \; + \gamma_8 mean(TMA_{irt}) + \sum_{t=1}^{19} \gamma_{9t} Time_{it} + u_{irt}^{Y} + c_{ir}^{Y}, \tag{5}$$

where the latent variable $d_{irt}^{Y*} > 0$, iff the dummy variable d_{irt} takes on a value of 1, and $d_{irt}^{Y*} < 0$, iff the dummy variable $d_{irt} = 0$. In estimating the selection equation, we allow for unobserved heterogeneity which enters the equation above as follows: $\varepsilon_{irt}^{Y} = c_{ir}^{Y} + u_{irt}^{Y}$,

[25]The pre-treatment heterogeneity refers to firms self-selecting themselves into early/late stage R&D cooperations due to firm-level heterogeneities such as managerial ability and absorptive capacity.

[26]The terms, CT_1 and CT_0, are also comnonly referred to as inverse Mill's ratio.

[27]Another problem is that the error term and firms' positions in the markets might be correlated, causing a potential endogeneity problem. We therefore use different lags of the technology and product market variables as instruments, which is common practice in panel data estimations.

$Y = Early, Late$, where c_{ir}^{Y} is the unobserved heterogeneity and $u_{irt}^{Y} \approx N(0,1), Y = Early, Late$, are the idiosyncratic error terms. Following Wooldridge (2002), we include the time averages of the technology, $[mean(TMA)]$, and product markets, $[mean(PMA)]$ to control for unobserved firm-specific heterogeneity and estimate equation (5) using a probit estimator.

Note, the selection equation is dependent on experience (EXP_{irt}^{Early} and EXP_{irt}^{Late}) and the diversification variables ($Tech_{it}$ and $Product_{it}$), which serve as the exclusion restriction to identify the model. Regarding the experience variables, we build on the transaction cost argument, i.e., that the specific R&D cooperations are associated with significant organizational, contractual and administrative efforts that diminish the more expertise a firm collected. We therefore use the number of previous early and late stage cooperations as an instrument which has an impact on the likelihood of forming a cooperation, but does not directly impact the position in the therapeutic and product market. They rather indirectly impact the outcome equations via R&D cooperations. We use the diversification variables as additional instruments, since they reflect different potential learning and synergy effects associated to engaging into early and late stage R&D cooperations.

Following Heckman, Urzua, and Vytlacil (2006), we apply a heterogeneous treatment effects model. As described above, our empirical model consists of a selection and an outcome equation. The selection equation (5) describes firms' endogenous selection into early/late stage R&D cooperations, and is estimated using a probit estimation to account for unobserved heterogeneity. We then formulate the correction terms, CT_1 and, CT_0 and estimate the outcome equations (4) and (3), which evaluate the impact of early/late stage R&D cooperations on the therapeutic market and the product market, respectively[28].

6 Empirical results

We first report the results from estimating the selection into early/late stage R&D cooperations, and then turn to the impact of early/late stage R&D cooperations on the

[28]IVTREATREG command in STATA has been used (Cerulli (2012)).

product and the therapeutic market.

6.1 Selection into early/late stage R&D cooperations

We estimate firms' selection into early/late stage R&D cooperations using a probit model. Based on the estimation results, we test for the presence of unobserved heterogeneity by estimating $\rho = \frac{\sigma_c^2}{(1+\sigma_c^2)}$, where ρ is the proportion of total variance contributed by the panel-level variance component. Our test returns significant estimates of ρ of 0.342 and 0.388 with p values of 0.00, which confirms the presence of unobserved heterogeneity. Following Wooldrige (2002), we control for unobserved heterogeneity by including additional time-invariant firm-specific regressors $[mean(TMA)\}]$ and $[mean(PMA)\}]$ in our model specification and apply a probit estimation procedure.

The estimation results are shown in Table 6. Columns 2 and 3 report the marginal effects for signing early and late stage R&D cooperations, respectively. As shown in column 2, a 10 percent decrease in the therapeutic market activity (TMA) significantly increases the likelihood of forming early stage R&D cooperations by 5 percent. Hence, early stage R&D cooperations are more relevant among firms which are less active in therapeutic areas. Research projects at the early stages, which involve the invention of new drug molecules, are costly and firms face impediments in financing those projects. Early stage R&D cooperations are formed among more financially constrained firms and allow firms to share costs and to overcome financial constraints. This finding is consistent with the results found in previous studies.

Interestingly, a stronger presence in the product market (PMA) has an insignificant impact on forming early stage cooperations. One explanation is that active firms in the product market do not rely on sharing R&D expenses. This explanation is supported by the fact that active firms in the product market benefit from higher revenues earned from their marketed drugs. Hence, they are less constrained in performing their basic research projects. Another explanation is that active firms are more hesitant to reveal information on their early stage research projects to competitors. Therefore, more established firms

in the product market do not have an incentive to form R&D cooperations at the early stage.

Turning to firms' experience of engaging into early stage R&D cooperations (EXP^{Early}), our results show that one further past engagement in early stage R&D cooperations increases the probability of engaging into an early stage R&D cooperation by 5 percent. More experience in joining early stage R&D cooperations lowers firms' transaction and contracting costs due to improved organizational and management skills. Note that experience in late stage R&D cooperations lowers the probability of engaging into cooperations at the early stage. This result indicates that transaction costs are specific to the type of cooperation and also supports the relevance to distinguish between cooperations formed at the early and late stages. Moreover, this result also confirms that experience has a direct impact on forming R&D cooperations, a result that supports our identification argument.

Firms' presence in different technology and product markets ($Tech$ and $Product$) turns out to have an insignificant impact on forming early stage cooperations. Firms are not able to gain higher benefits via spillovers and transferring their knowledge from one therapeutic area to other therapeutic areas. Hence, knowledge is highly area-specific and not easily transferable. This result confirms that therapeutic area-specific expertise is important for the types of R&D cooperations.

Finally, firm-level heterogeneity in the technology market [$mean(TMA)$] explains firms' choice to engage in early stage R&D cooperations. Omitting the firm-level heterogeneity would result in an overestimated impact of technology market level of activity on forming early stage cooperations. In contrast, accounting for firm-level heterogeneity in the product market [$mean(PMA)$] is not significant. This result is consistent with the previous finding that firms' presence in the product market (PMA) has an insignificant impact on explaining early stage cooperations. Hence, firm-specific factors associated with product markets are less important than firm-specific factors related to therapeutic markets.

Turning to the estimation results for late stage R&D cooperations (see Table 6, column

3), it is interesting to note that a 10 percent increase in a firm's therapeutic market level of activity (TMA) increases the likelihood of forming late stage R&D cooperations by 9.2 percent. Firms search for partners with similar research agendas and a strong experience in applied drug testing to increase the success probability of launching new drugs. Hence, more research active firms select themselves into late stage R&D cooperations. This result stands in contrast to early stage R&D cooperations, which are rather formed among less research active firms. In general, the results show that self-selection into early and late stage cooperations occurs depending on their level of activity in the technological market.

Moreover, a strong product market presence (PMA) has a positive impact on forming late stage cooperations. An increase in a firm's product market level of activity (PMA) by 10 percent increases the likelihood of forming a late stage R&D cooperation by 23 percent. The tendency to cooperate with close product market competitors at the final stages of the research process emphasizes the fact that scope economies, human subject management, and drug safety testing are important drivers for participating in late stage R&D cooperations. Moreover, since a drug becomes more likely to be launched on the market once it passed the final stages of the research process, firms are more effective in pooling their R&D efforts, as well as re-optimizing their drug development portfolios to avoid closely related substitute drug production. Hence, stronger product market competitors self-select into late stage cooperations once the drug development process approaches the final stages.

The experience in late stage cooperations increases the likelihood of engaging in further late stage cooperations. This result again shows that transaction costs are rather specific to the type of R&D cooperations. It also emphasizes the fact that the objectives and requirements change throughout the R&D process and firms specialize in participating in early or late stage cooperations. The estimates for the technology market ($Tech$) and product market ($Product$) portfolios carry a positively and negatively significant estimate, respectively. Firms, diversified in technology markets, benefit from late stage cooperations managing large number of late stage clinical trials and research projects. Firms, diversi-

fied in product markets, have more in-house expertise and need less external knowledge. Finally, the firm-level heterogeneity, $[mean(TMA)\ ]$ and $[mean(PMA)]$, resembles the previous finding that firms' presence in the technology and product market has a significant impact on explaining late stage cooperations. Hence, firm-specific factors associated with both technology and product markets are important in explaining the engagement into late stage R&D cooperations.

To summarize, the probit results provide evidence that firms select themselves into early and late stage cooperations, depending on their presence in the technological and product markets. Early stage R&D cooperations are used as an instrument for less research active firms to overcome financial impediments. Late stage R&D cooperations are performed among active firms in therapeutic and product markets. Moreover, the results show that previous experience in engaging into specific types of R&D cooperations has a significant impact, confirming our identification argument.

6.2 Impact of early/late stage cooperations on technology and product markets

In this section, we report on the impact of early/late stage R&D cooperations on the technology and product market. We apply an estimation procedure which accounts for selection on observables and unobservables and also accounts for post selection heterogeneity. Using the estimates from the probit selection equation (5), we construct the selection terms for the early and late stage R&D cooperations and estimate the impact on the product and technology market according to equations (3) and (4), respectively[29].

Table 7 shows the estimation results of early and late stage R&D cooperations, also referred to as the average treatment effect on the treated (ATET). The upper left panel of Table 7 shows the impact of early stage cooperations on the technology market. Moving along the columns shows how the impact evolves over time. The results show that early

[29]We estimate the model accounting for an impact of one and two years after cooperations have been formed. Due to data constraints, we are not able to evaluate longer time horizons.

stage cooperations increase the technology market level of activity. Accounting for the fact that a firm participates on average in 5.3 early stage R&D projects[30], we can predict that early stage R&D cooperations increase the level of activity in the therapeutic market by 21% (30%), one (two) years after formation. The results clearly indicate that firms are able to transfer the knowledge they gained in early stage R&D cooperations to initiate new individual research projects. Hence, knowledge spillovers encourage firms to initiate further individual projects over time. Considering the previous result that less research active firms engage in early stage cooperations, which also help firms overcome their financial constraints, we can conclude that firms were able to successfully overcome their financial constraints[31].

The upper right panel of Table 7 shows the impact of early stage R&D cooperations on the product markets, see column 5. The results show that early stage R&D cooperations increase firms' level of activity in the product market by 22% (34%), one (two) years after formation. It is important to note that early stage R&D cooperations further increase firms' level of activity over time. Hence, the early stage R&D cooperations have long lasting synergy effects and allow firms to increase the number of drugs offered on the market.

It is worth comparing the impact of early stage R&D cooperations on the therapeutic market activity with the impact on the product market activity, see Table 7, columns 3 and 5. The short-run impact (one year after the cooperation) resulted in a comparable increase of around 22% in firms' level of activity in the therapeutic and product market area. The long-run impact (two years after the cooperation) on the number of drugs offered on the market increased by 4% more than the R&D projects. At first glance, this seems to be surprising as research projects result in new drugs launched on the market. Therefore, we would rather expect a lower impact on drugs marketed than on new research activity. This argument is even more pronounced if we take into consideration that some research activities may also become obsolete due to a high degree of uncertainty in R&D.

[30]This number is taken from Table 5.

[31]Remember that Table 6, column 2 shows a negative impact of TMA on the likelihood of forming early stage cooperations.

Hence, the numbers clearly illustrate two effects: first, early stage cooperations enable firms to overcome financial constraints and to increase their research activities. Second, the number of drugs increases significantly in the long run and even overturns the increase in research projects. This result illustrates that firms are able to transfer their newly gained synergy effects towards their individual projects which increases the number of drugs launched on the market.

The lower left panel of Table 7 (column 3) shows that late stage cooperations increase the therapeutic market level of activity by 12% in the short-run[32]. This impact is lower than in early stage cooperations, emphasizing the notion that synergy effects are lower in late stage cooperations than in early stage cooperations. This seems plausible as technological spillovers are higher at the early basic research stages than in the later stages in which drugs are mostly clinical trials. Note that the short-run gains are not sustainable in the long run.

The lower right panel of Table 7 (column 5) shows the impact of late stage cooperations on the product market. Focusing on the short-run effect, firms are able to increase their number of drugs launched on the market by 6% one year after a late stage cooperation has been formed. This impact is significantly smaller than the short-run effect of early stage R&D cooperations on product markets, which confirms the fact that synergy effects are prevalent at the early stages of the research process rather than at the late stages. Very surprisingly, the short-run impact of late stage cooperations on the therapeutic market is 12%, while the same impact on the product market is only 6%. This finding has two alternative explanations: either many research projects that were initiated after the R&D cooperation become unsuccessful, or late stage R&D cooperations provide an opportunity for firms to coordinate their drug development portfolios as to avoid wasteful duplication and to avoid highly substitutable drugs being launched on the market. To summarize, in the short run, the number of R&D projects drastically increases while the number of new drugs offered on the product market only slightly increases. Most impressive is

[32]This number is evaluated at the average number of late stage projects (6.89 late stage R&D projects) that firms are involved in.

the finding that the long-run impact of late stage R&D cooperations on the product market is negative, i.e., the number of drugs decreases by 49% two years after a late stage R&D cooperation has been formed. Part of the reduction is explained by the non-significant impact of late stage R&D cooperations on the therapeutic market. However, this would only explain a non-increase in the number of drugs. The drastic reduction is rather explained by the fact that R&D cooperations represent an opportunity for firms to re-optimize their drug portfolios and to avoid close substitutes being offered on the market. Moreover, this explanation is consistent with the previous finding that active firms in research and production engage in late stage R&D cooperations (see Table 6), and active firms in the product market are more likely to cannibalize each other's sales by introducing similar drugs into the market. Hence, the gain to re-optimize drug portfolios and avoiding the introduction of closely related drugs is higher for those firms.

We now turn to discuss the counterfactual impact on the therapeutic and product market level of activity, if firms which did not engage in early/late stage R&D would have formed one of those cooperations, see Table 8[33]. The upper panel of Table 8, column 3, shows that early stage cooperations do not significantly impact firms' technology market level of activity. This result is plausible as those firms did not self-select themselves into early stage R&D cooperations. Consequently, non-selected firms would not find the same benefit as those who self-selected themselves into early R&D cooperations, i.e., overcoming financial constraints. Consequently, firms would have not increased their research activity. Interesting is the following result: if noncooperating firms would have selected themselves into early stage cooperations, product market level of activity would have increased by 32% one year later. This result emphasizes the fact that the product market gains and efficiency gains in clinical trials are not constrained to self-selected firms, but could have been achieved by other firms as well.

Finally, the lower panel of Table 8 shows that late stage cooperations increase technology (product) market level of activity by 12% (8%) in the short run. This positive effect, however, is not sustainable in the long run and disappears. Comparing this result to the

[33]Also commonly referred to as the treatment effect on the nontreated (ATENT).

sample of firms that self-selected themselves into late stage cooperations (see lower panel of Table 7, column 5), we realize that the impact on product market is higher (see Table 8). Hence, firms which did not self-select into late stage R&D cooperations would have not reduced the number of drugs. Firms' benefit in re-optimizing their drug portfolios is also supported by the previous finding that more active firms in research and production engage in late stage R&D cooperations, and these firms benefit the most from coordinating and re-optimizing drug portfolios. Finally, the results for the average treatment effects of early and late stage R&D cooperations, as shown in Table 9, resemble the results in Tables 7 and 8.

Robustness checks The technology and product market level of activity variables, TMA and PMA, represent the most relevant variables in our study. Since the results might be sensitive to the specific variable definition, we apply robustness checks with respect to different definitions of those variables. Hence, we weighted TMA and PMA by the Herfindahl index of the technology market and the product market areas. This definition accounts for the level of activity in the technology and product markets relative to other technology and product markets. TMA is defined as:

$$TMS_{irt} = \frac{\widetilde{z}_{irt}}{\sum_{i=1}^{N} \widetilde{z}_{irt}^2}.$$

$$TMA_{irt} = TMS_{irt} + \sum_{s=1}^{t-1} (1-\delta)^s TMS_{irt}.$$

And PMA is defined as:

$$PMS_{irt} = \frac{\widetilde{p}_{irt}}{\sum_{i=1}^{N} \widetilde{p}_{irt}^2}.$$

$$PMA_{irt} = PMS_{irt} + \sum_{s=1}^{t-1} (1-\delta)^s PMS_{irt}.$$

The results for the first stage probit and the heterogeneous treatment effect do not change significantly[34]. We run the probit without the portfolio diversification variables ($Tech$ and $Product$). All coefficient estimates obtained are of same sign and significance level, except PMA for the prediction of the likelihood of signing early stage cooperations. The sign was negative at the 10% significance level. Additionally, we run the model with time trend instead of time dummies without any significant change in the result. Finally, we also apply an alternative estimation method, i.e., a propensity score estimator to check for the associated conditional independence assumption. Given the data availability, we conclude with three stratas in our propensity estimation method, and find that the basic results are unchanged.

7 Conclusion

This study analyzes the impact of R&D cooperations formed at different stages on technology and product market competition in the pharmaceutical industry. Using a rich dataset on firms' activities in therapeutic and disease areas over time, we estimate a heterogeneous treatment effect model and account for firms selecting themselves into early and late stage R&D cooperations (pretreatment heterogeneity). Moreover, we explicitly allow firms having a specific impact on the therapeutic and product market after having selected themselves into specific types of R&D cooperations (heterogeneous treatment effect).

Our main results show that early stage R&D cooperations increase firms' level of activity in the therapeutic and product markets. Most interestingly, we find that late stage R&D cooperations significantly reduce the number of drugs launched on the market. This result highlights the fact that firms re-optimize their drug development portfolio to avoid wasteful duplication and cannibalizing the sales of the jointly developed drug in R&D cooperations.

We also find that early stage R&D cooperations are formed among less research-active

[34]The results are available from the authors upon request.

companies, which supports previous findings that less research active firms use early stage R&D cooperations as an instrument to overcome financial constraints. On the contrary, late stage R&D cooperations are formed among more research-active and production-oriented firms. The result reflects that product market competitors avoid disclosing their research pipeline to their competitors at the early stages. At the later stages, however, their incentive to form R&D cooperations with close product market competitors increases. In line with previous findings, partnering with closer product market competitors emphasizes the fact that scope economies are important at the later stages of the research process.

To conclude, our study suggests that antitrust authorities should pay special attention to late stage cooperations, as those have the potential to lower the number of drugs offered on the market, causing a potential harm to product market competition and consumer welfare. Even though, we believe that this study provides an important insight into different types of R&D cooperations and their ultimate impact on product variety, we also would like to emphasize that this study is one of the firs studies in this area, and further research is warranted to derive stronger policy statements. For example, it would be interesting to analyze to what extent different types of R&D cooperations impact the prices of the developed drugs. These questions, however, are beyond the scope of our study and our data availability.

References

Angrist, J. and A. Krueger, 1999, "Empirical strategies in labor economics," Handbook of Labor Economics, in: O. Ashenfelter and D. Card (ed.), edition 1, 3(23), pp. 1277-1366, Elsevier.

Arora, A., A. Gambardella, L. Magazzini and F. Pammolli, 2009, "A Breath of Fresh Air? Firm Type, Scale, Scope, and Selection Effects in Drug Development," Management Science, INFORMS, 55(10), pp. 1638-1653.

Brand, J.E. and J.S. Thomas, 2012, "Casual Effect Heterogenity,"
http://papers.ccpr.ucla.edu/papers/PWP-CCPR-2011-012/PWP-CCPR-2011-012.pdf.

Brand J.E. and Y. Xie, 2010, "Who benefits most from college? Evidence for negative selection in heterogeneous economic returns to higher education," American Sociological Review, 75, pp. 273-302.

Cabral, L., 2000, "R&D cooperation and product market competition," International Journal of Industrial Organization, 18, pp. 10331047.

Cassiman, B. and R. Veugelers, 2002, "Spillovers and R&D Cooperation: Some Evidence from Belgium," American Economic Review, 92(4), pp.1169-1184.

Cerulli G., 2012, "ivtreatreg: a new STATA routine for estimating binary treatment models with heterogeneous response to treatment under observable and unobservable selection, " CNR-Ceris Working Papers, No. 03/12.

D'Aspremont, C. and A. Jacquemin, 1988, "Cooperative and Noncooperative R&D in Duopoly with Spillovers," American Economic Review, 78, pp. 1133-1137.

Danzon, P., S. Nicholson and N.S. Pereira, 2005, "Productivity in pharmaceutical-biotechnology R&D: the role of experience and alliances," Journal of Health Economics, 24(2), pp.31739.

Dehejia, H.R. and S. Wahba, 2002, "Propensity score-matching methods for nonexperimental causal studies," The Review of Economics and Statistics, MIT Press, 84(1), pp. 151-161.

DiMasi, J.A., R.W. Hansen, H.G. Grabowski and L. Lasagna, 1991, "Cost of innovation in the pharmaceutical industry," Journal of Health Economics, 10, pp. 107-142.

DiMasi, J.A., R.W. Hansen and H.G. Grabowski, 2003, "The price of innovation: new estimates of drug development costs," Journal of Health Economics, 22, pp. 151-185

DiMasi, J.A., L. Feldman, A. Seckler and A. Wilson, 2010, "Trends in risks associated with new drug development: success rates for investigational drugs," Clinical Pharmacology & Therapeutics, 87, pp. 272277.

Erkal, N. and D. Minehart, 2008, "Optimal Sharing Strategies in Dynamic Games of Research and Development," Available at SSRN: http://ssrn.com/abstract=997336.

Goeree, M.S. and E.A. Helland, 2010, "Do Research Joint Ventures Serve a Collusive Function?," http://www.econ.uzh.ch/faculty/mgoeree/research/rjvmay18.pdf.

Griliches, Z. and J. Mairesse, 1984, "Productivity and R&D at the firm level, in Z. Griliches (ed.), R&D, Patents, and Productivity," The University of Chicago Press and NBER: Chicago.

Griliches, Z. 1977, "Estimating theReturns to Schooling:Some Econometric Problems," Econometrica, 45, pp. 122.

Grabowski G. H. and K. Margaret, 2012, "Mergers, Accquisitions, and Alliances, in The Oxford Handbook of The Economics of the Biopharmaceutical Industry, " edited by Patricia M. Danzon and Sean Nicholson, pp. 552-577, Oxford Press.

Grossman, G.M. and C. Shapiro, 1986, "Optimal Dynamic R&D Programs," RAND Journal of Economics, The RAND Corporation, 17(4), pp. 581-593.

Grossman, G.M. and C. Shapiro, 1987, "Dynamic R&D Competition," Economic Journal, 97(386), pp. 372-87.

Gugler K. and R. Siebert, 2007,"Market Power versus Efficiency Effects of Mergers and Research Joint Ventures: Evidence from the Semiconductor Industry," Review of Economics and Statistics, 89(4), pp. 645-659.

Hall, B.H., 1993, "The stock market's valuation of R&D investment during the 1980's," American Economic Review, 83, pp. 259-264.

Heckman J., S. Urzua and E. Vytlacil, 2006, "Understanding Instrumental Variables in Models with Essential Heterogeneity," The Review of Economics and Statistics, 88, pp. 389-432.

Henderson, R. and I. Cockburn, 1996, "Scale, Scope, and Spillovers: The Determinants of Research Productivity in Drug Discovery, " RAND Journal of Economics, 27(1), pp. 32-59.

Henderson, R and I. Cockburn, 2001, "Scale and Scope in Drug Development: Unpacking the Advantages of Size in Pharmaceutical Research," Journal of Health Economics, 20(6), pp. 1033-1057.

Higgins, M.J. and D. Rodriguez, 2006, "The Outsourcing of R&D through Acquisition in the Pharmaceutical Industry," Journal of Financial Economics, 80.

Holland, P., 1986, "Statistics and Causal Inference," Journal of the American Statistical Association, 81, pp. 945-970.

Irwin, D. and P. Klenow, 1996, "High Tech R&D Subsidies: Estimating the effects of SEMATECH," Journal of International Economics, 40,(3-4), pp. 323-344.

Jorde, T. and J. T. David, 1990, "Innovation and Cooperation: Implications for Competition and Antitrust," Journal of Economic Perspectives, 4(3), pp. 75-96.

Kamien, M., E. Muller and I. Zang, 1992, "Research Joint Ventures and R&D Cartels," American Economic Review, 82, (5), pp. 1293-1306.

Katz, L.M., 1986, "An Analysis of Cooperative Research and Development," The RAND Journal of Economics, 17(4), pp. 527-543.

Lancaster, K., 1979, "Variety, Equity and Efficiency," New York, Columbia University.

Lerner, J., and A. Tsai, 2000,"Do Equity Financing Cycles Matter? Evidence From Biotechnology Alliances, " Journal of Financial Economics, 67,(3), pp. 411446.

Lerner, J. and R.P. Merges, 1998, "The control of technology alliances: an empirical analysis of the biotechnology industry," Journal of Industrial Economics, 46, pp. 125-156.

Martin, S., 1995,"R&D joint ventures and tacit product market collusion," European Journal of Political Economy, 11, pp. 73341.

Morgan, S. and W. Christopher, 2007 "Counterfactuals and Causal Inference: Methods and Principles for Social Research," Cambridge, UK: Cambridge University Press.

Nesta, L. and P.P. Saviotti, 2005, "Coherence of the knowledge base and the firms' innovation performance: Evidence from the U.S. Pharmaceutical industry," Journal of Industrial Economics, Wiley Blackwell, 53(1), pp. 123-142.

Nicholson, S., P.M. Danzon, and M. Jeffrey, 2003, "Biotech-pharmaceutical alliances as a signal of asset and firm quality, " Journal of Business.

Ornaghi, C., 2009, "Mergers and innovation in big pharma," International Journal of Industrial Organization, 27, pp. 7079.

Oxley, J.E. and R.C. Sampson, 2004, "The scope and governance of international R&D alliances," Strategic Management Journal, 25, pp. 723-750.

Pais, J., 2011, "Socioeconomic background and racial earnings inequality: A propensity score analysis," Social Science Research, 40, pp. 37-49.

Powell, W.W. and P.Brantley, 1992, "Competitive cooperation in biotechnology: learning through networks? In: Nohria, N., Eccles, R.G. (Eds.), Network and Organizations: Structure, Form and Action," Harvard Business School Press, Boston, pp. 366394.

Powell, W.W., K.W. Koput, L. Smith-Doerr and J. Owen-Smith, 1999, "Network position and firm performance: organizational returns to cooperation in the biotechnology industry, " In: Andrews, S., Knoke, D. (Eds.), Networks in and Around Organizations. JAI Press, Greenich, CT.

Roeller, L.H., R. Siebert and M. Tombak, 2007, "Why Firms Form (or don't Form) RJVs," Economic Journal, 117(522), pp. 1122-1144.

Roeller, L.H., D. Tomaso and J. Seldeslachts, 2013, "Collusion Through Joint R&D: An Empirical Assessment," The Review of Economics and Statistics.

Saami Z., 2006, "Suppression Of Innovation Or Collaborative Efficiencies?: An Antitrust Analysis Of A Research & Development Collaboration That Led To The Shelving Of A Promising Drug," 5 J. Marshall Rev. Intell. Prop. L. 348.

Salop, S.C., 1979, "Monopolistic Competition with Outside Goods," Bell Journal of Economics, 10, pp. 141-156.

Sakakibara, M., 2002, "Formation of R&D Consortia:Industry and Company Effects," Strategic Management Journal, 23(11), pp. 1033-1050.

Shapiro C. and D. R. Willig, 1990, "On the Antitrust Treatment of Production Joint Ventures," Journal of Economic Perspectives.

Siebert, R. and G. V. Graevenitz, 2010, "Jostling for Advantage or Not: Choosing Between Patent Portfolio Races and Ex Ante Licensing," Journal of Economic Behavior & Organization, 73, pp. 225-245.

Siebert, R. and G. V. Graevenitz, 2012, "Are Licensing Agreements Appropriate Instruments to Cut Through the Patent Thicket?,"

http://www.law.northwestern.edu/searlecenter/papers/Siebert_Licensing_Agreements.pdf

.

Spence, M., 1984, "Cost Reduction, Competition, and Industry Performance," Econometrica, Econometric Society, 52(1), pp. 101-21.

Suetens S., 2008, "Does R&D cooperation facilitate price collusion? An experiment," Journal of Economic Behavior and Organization, 66, pp. 822-836.

Wooldridge, J. M., 2002, "Econometric analysis of cross section and panel data," MIT Press, Cambridge, Mass.

8 Appendix

8.1 Examples of Early stage R&D cooperation

ANGIOGENE signed a contract at Discovery stage with MAXCYTE in 2002. The therapeutic area was cardiovascular. Atgiogene did not have any approved drugs at that time but Maxyte was operating in cardiovascular area.

ASTRAZENECA signed a research contract with MORPHOCHEM AG in 2002. The therapeutic area was Blood & Hematopoietic. ASTRAZENECA was research active in Gastroenterology, Neurology, and Oncology and had approved NDAs in these areas but not in the Blood & Hematopoietic area. Morphochem was research active in the Blood & Hematopoietic area.

8.2 Examples of Late stage R&D cooperation

GLAXOSMITHKLINE signed a Phase III R&D cooperation with Roche in Dec 1, 2001. Got the license and signed a co-development agreement for Ibandronate for osteoporosis. At that time GLAXOSMITHKLINE was operating in metabolic disorder and had NDAs in Endocrinology(1), Hematology(1), Pulmonary/Respiratory Diseases(1). Roche had NDAs in Immunology/Infectious Diseases(1), Oncology, Opthalmology and Pulmonary/Respiratory Diseases(1).

Novartis and CELGENE signed Development, License cooperations in April 2001 at Phase III for a Central Nervous System drug. Novartis had approved NDA in Dermatology/Plastic Surgery(1), Endocrinology(2), Hematology(1), Immunology/Infectious Diseases(1), Neurology(3), Obstetrics/Gynecology(1), Pharmacology/Toxicology(1), Celgene had Hematology(1), Neurology(1).

8.3 Technological areas

Cancer; Cardiovascular; Central Nervous System; Dermatology; Diagnostic/Imaging agent/Delivery; Eye and ear; Gastrointestinal; Genitourinary; Hematological; HIV; infection; Hormonal system; Immune system; Infectious disease; Musculoskeletal; Nephrology; Pain; Respiratory.

Table 1: Drug development stages

Stage	Description
Discovery	Target identification, biochemical mechanism.
Formulation	Identifying drug's stability.
Lead Molecule	Identifying the lead molecule for the development.
Preclinical	On animals to find out various parameters.
Phase I	Small-scale, identify tolerance, repeated-dose studies.
	Healthy volunteers. Initial single-dose, dose increase.
Phase II	Small-scale, preliminary efficacy on patients.
Phase III	Large-scale clinical trials, safety and efficacy
	Large scale patients, Preparation for NDA.

Table 1 presents different drug development stages in the pharmaceutical industry. Source: www.pacificbiolabs.com/drug_stages.asp

Table 2: Total number of approved NDA

Year	Total drugs	Year	Total drugs
1993	70	2002	78
1994	62	2003	72
1995	82	2004	119
1996	131	2005	80
1997	121	2006	101
1998	90	2007	78
1999	83	2008	89
2000	98	2009	90
2001	66	2010	93
		2011	59

Table 2 presents the total number of NDA approved by FDA. Dataset is provided by Biopharm Insight (approved drug database).

Table 3: Number of alliances in different technological areas

Technological area	Number of Alliances
Cancer	781
Cardiovascular	226
Central Nervous System	380
Dermatology	86
Diagnostic/Delivery	36
Eye and Ear	110
Gastrointestinal	205
Genitourinary	90
HIV Infections	60
Hematological	100
Hormonal Systems	256
Immune System	198
Infectious Diseases	404
Miscellaneous	328
Musculoskeletal	162
Nephrology	38
Pain	200
Respiratory	96
Total	3,756
Phases	**Number of Alliances**
Discovery	561
Formulation/Lead Molecule	123
Pre-Clinical	1087
Phase I	538
Phase II	798
Phase III	575
Regulatory Filing	74

Table 3 presents the number of alliances in different technological areas, as well as alliances by phases at the time of the deal. Dataset is provided by Biopharm Insight.

Table 4: Descriptive statistics for all companies

Variable	Observations	Mean	Std. Dev.	Min	Max
Treatment variables					
$Early$	7320	0.2262	0.4184	0	1
$Late$	7320	0.2868	0.4523	0	1
Explanatory variables					
TMA	7320	0.6382	0.0338	0	8.37
PMA	7320	0.0650	0.0057	0	17
$Tech$	7320	0.7170	2.3068	0	18
$Product$	7320	0.3453	1.4038	0	14
EXP^{Early}	7320	2.2394	5.4591	0	49
EXP^{Late}	7320	3.4284	8.1471	0	60

Table 4 presents summary statistics using data from 1993 until 2011 for all the cooperation-signing companies. Dataset is provided by Biopharm Insight. TMA_{irt} refers to the technology market level of activity and PMA_{irt} refers to the product market level of activity.

Table 5: Descriptive statistics by early and late stage partners

Early stage partners					Late stage partners			
Variable	Mean	Std. Dev.	Min	Max	Mean	Std. Dev.	Min	Max
TMA	1.2951	0.0473	0	8.37	1.2005	0457	0	7.9853
PMA	1.1754	0.0078	0	15.00	1.324	0.0080	0	17.00
$Tech$	1.3503	3.1817	0	18	1.2667	2.9699	0	18
$Product$	0.5100	1.9284	0	14	0.4328	1.5416	0	14
EXP^{Early}	5.3405	7.7078	0	49	3.5142	6.2862	0	44
EXP^{Late}	6.2983	11.2530	0	60	6.8861	9.6734	0	60

Table 5 presents summary statistics using data from 1993 until 2011 for for early and late stage cooperation-signing companies separately. Dataset is provided by Biopharm Insight. TMA_{irt} refers to the technology market level of activity and PMA_{irt} refers to the product market level of activity.

Table 6: Decision to form early and late stage R&D cooperation

VARIABLES	Dependent variable: Early stage cooperation	Dependent variable: Late stage cooperation
TMA_{t-1}	-0.0045*	0.0092***
	(0.003)	(0.002)
PMA_{t-1}	0.0006	0.0228**
	(0.015)	(0.011)
EXP^{early}	0.0479***	-0.0450***
	(0.003)	(0.003)
EXP^{late}	-0.0280***	0.0412***
	(0.002)	(0.002)
$Tech$	-0.0039	0.0240***
	(0.005)	(0.004)
$Product$	0.0092	-0.0626***
	(0.009)	(0.007)
$mean(TMA)$	0.0136**	0.0043**
	(0.006)	(0.004)
$mean(PMA)$	-0.0781*	0.1033***
	(0.042)	(0.031)
$Time\ dummies$	Yes	Yes
$Observations$	7,320	7,320
$Loglikelihood$	-1911	-3714

Table 6 presents the marginal effects of probit estimation for equation 5. Dependent variables are the early/late stage R&D cooperation. Explanatory variables are the lagged technology market level of activity, lagged product market level of activity, early/late stage experience, diversification in technology and product markets, the time averages of the technology and product market, and time dummies. Potential endogeneity and unobserved heterogeneity are controlled. Standard errors are shown in the parenthesis. *** and ** denote 99% and 95% level of confidence.

Table 7: Average treatment effect on treated: impact of early/late stage R&D cooperations on technology and product markets

Technological and product market level of activity (ATET(x))				
Early stage cooperations	TMA_{t+1}	0.0383**	PMA_{t+1}	0.0412*
		(0.148)		(0.026)
	TMA_{t+2}	0.0553**	PMA_{t+2}	0.0636**
		(0.032)		(0.087)
Late stage cooperations	TMA_{t+1}	0.0171**	PMA_{t+1}	0.0093*
		(0.005)		(0.005)
	TMA_{t+2}	0.0155	PMA_{t+2}	-0.0713**
		(0.015)		(0.044)

Table 7 presents the instrumental variable estimation results for the impact of early and late stage R&D cooperation on treated firms' level of activity in the technological and product market, one and two periods after forming a R&D cooperation. Number of previous early and late stage cooperations are used as an instrument. Standard errors are shown in the parenthesis. ***, ** and * denote 99%, 95% and 90% level of confidence.

Table 8: Average treatment effect on nontreated: impact of early/late stage R&D cooperations on technology and product markets

Technological and product market level of activity (ATENT(x))				
Early stage cooperations	TMA_{t+1}	0.0348	PMA_{t+1}	0.0596*
		(0.141)		(0.022)
	TMA_{t+2}	0.0652	PMA_{t+2}	0.0369
		(0.032)		(.039)
Late stage cooperations	TMA_{t+1}	0.0181*	PMA_{t+1}	0.0112**
		(0.007)		(0.005)
	TMA_{t+2}	0.0162	PMA_{t+2}	-0.1163
		(0.013)		(0.047)

Table 8 presents the instrumental variable estimation results for the impact of early and late stage R&D cooperation on non-treated firms' level of activity in the technological and product market, one and two periods after forming a R&D cooperation. Number of previous early and late stage cooperations are used as an instrument. Standard errors are shown in the parenthesis. ***, ** and * denote 99%, 95% and 90% level of confidence.

Table 9: Average treatment effect: impact of early/late stage R&D cooperations on technology and product markets

Technological and product market level of activity (ATE(x))				
Early stage cooperations	TMA_{t+1}	0.0362*	PMA_{t+1}	0.0524
		(0.114)		(0.001)
	TMA_{t+2}	0.0622***	PMA_{t+2}	0.0485
		(0.000)		(0.049)
Late stage cooperations	TMA_{t+1}	0.0176	PMA_{t+1}	0.0102**
		(0.022)		(0.009)
	TMA_{t+2}	0.0160*	PMA_{t+2}	-0.0904
		(0.008)		(0.009)

Table 9 presents the instrumental variable estimation results for the impact of early and late stage R&D cooperation on firms' level of activity in the technological and product market, one and two periods after forming a R&D cooperation. Table 9 presents the average treatment effect. Number of previous early and late stage cooperations are used as an instrument. Standard errors are shown in the parenthesis. ***, ** and * denote 99%, 95% and 90% level of confidence.