Antitrust Policy and Innovation*

Giovanni Morzenti†

June 2023

[Link to latest version]

Abstract

I study whether loosening antitrust policy discourages innovation of merging firms. A natural experiment on a relaxation of pre-merger notification rules allows me to compare mergers notified to the authorities with mergers that are not notified. I develop a novel text analysis methodology to identify horizontal mergers between close competitors. Using the universe of published patents, I can apply this methodology to small and private firms. After the policy change, non-notified horizontal mergers lead to a 30% reduction in patenting activity. Consistently with the deterrence effect of merger policy, the number of non-notified anticompetitive mergers rise after the relaxation of notification rules.

Keywords: Competition, Antitrust Policy, Mergers, Innovation, Patenting
JEL codes: L40 G34 O32

*I thank Axel Gottfries, Basile Grassi, Carlo Ramus Schwarz, Carmine Ornaghi, Cesare Righi, Chiara Fumagalli, Davide Coluccia, Fabiano Schivardi, Francesco Decarolis, Giovanni Rizzi, Julien Sauvagnat, Nicola Pavoni, Marco Ottaviani, Massimo Motta, Michele Polo, Rosa Ferrer, Tacye Hong, Thomas Wollmann and Vincenzo Denicolò for their comments and guidance. I would also like to thank the participants to SMYE 2022, EARIE 2022, 2022 London Workshop on Mergers, Innovation and the Labour Market, EWMES 2022, MACCI 2023 and seminars in DIW Berlin, Bocconi University, UPF Barcelona, Collegio Carlo Alberto, IMT Lucca, Edinburgh University, Universidade NOVA de Lisboa, Analysis Group Paris, Roma Tor Vergata for their contributions and suggestions. All mistakes and omissions are my own.

†Analysis Group. giomorzent@live.it

1 Introduction

Merger activity is on the rise and at an all time high. Figure 1 shows that both the number of mergers and the value of these transactions are increasing. [1] Should we worry about the impact of these mergers on innovation? By fostering competition, merger policy can shape the returns to and incentives for innovation.[2] However, competition affects innovation in an ambiguous way. Less competition allows firms to capture more returns from innovation.[3] At the same time, less competition means that firms have less incentive to innovate to gain a greater market share.[4] I study what is the ultimate effect of competition on innovation. In particular, this paper asks whether loosening merger policy discourages innovation of merging firms.

The first contribution of this work is an identification strategy based on a natural experiment. To exogenously shift antitrust enforcement for a subset of mergers over time, I exploit a year 2000 change of the US pre-merger notification program. This eliminated reporting requirements for thousands of mergers. This change led to a 70% decline in pre-merger notifications. By avoiding pre-merger notifications, firms elude antitrust enforcement almost entirely, resulting in "stealth consolidation."[5] As a result, this treatment abruptly eliminates antitrust scrutiny for a large subset of US mergers. To capture non-reported mergers, I use a database that includes mergers reported on industry journals, news outlets and other publications.[6]

Building on the work of Wollmann (2019), I exploit the difference between horizontal and non-horizontal mergers. Horizontal mergers are defined as those between firms operating in the same product markets. Since such mergers concentrate affected markets, horizontal mergers have the strongest effect on competition. Consequently, they attracted most of the attention of antitrust authorities in the early 2000s.[7] The treated group consists of firms involved in horizontal mergers that were not reported to the authorities. The control group, on the other hand, includes firms involved in either non-horizontal mergers or notified mergers.

[1] 2021 has been a record-setting year in global M&A activity. Total transaction volume topped $5.5 trillion, exceeding prior peaks in 2007 and 2015 that remained below $5 trillion. Consequently, the number of global merger control filings increased. The European Commission alone received 403 merger filings in 2021, which is the second highest figure in European Union merger control. More info on:
https://www.nytimes.com/2021/12/18/business/dealbook/deals-of-the-year.html

[2] Citing the Assistant Attorney General Jonathan Kanter of the DOJ, from the Federal Trade Commission website: "Our country depends on competition to drive progress, innovation, and prosperity...". Commissioner Margrethe Vestager of Directorate General Competition on the 2018 Bayer-Monsanto merger, from the official European Commission ebsite: "...we need competition to push companies to innovate in digital agriculture and to continue to develop new products that meet the high regulatory standards in Europe, to the benefit of all Europeans and the environment."

[3] This gives firms higher incentives to spend on research and development. This argument originates in the literature on creative destruction spurred from Schumpeter (1934).

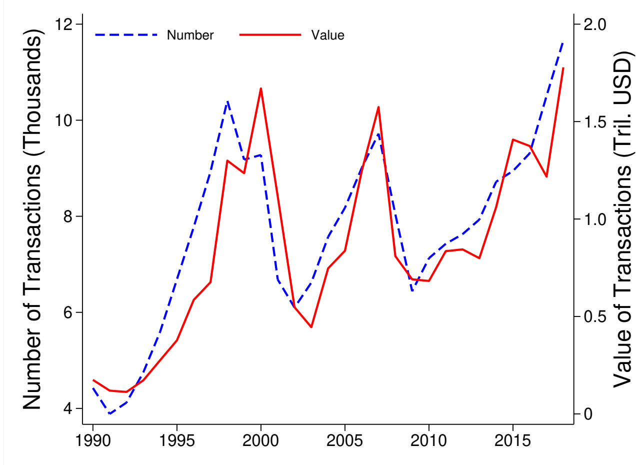
[4] Arrow (1962) explains that a monopolist might innovate less than a competitive firms because it stand to gain less from this innovation.

[5] Stealth consolidation is defined as an abrupt increase in the number of potentially anticompetitive mergers that are not reported to the authorities. See Wollmann (2021) for more details.

[6] I use Thomson Reuters SDC Platinum database: I focus on mergers and acquisitions involving firms operating in the US, which are subject to US antitrust policy.

[7] Just to emphasize how important horizontal mergers were for the authorities, in the 2000s the official US merger guidelines were called *Horizontal Merger Guidelines*.

Figure 1: Rising Merger Activity over Time.



Notes: Data from Thomson Reuters SDC Platinum. The figure reports only mergers involving firms operating in the US. The number of transactions includes only realized mergers, defined as a an acquisition of 100% of assets of the target firm. As a share of US GDP, the value of transactions trebled, from 2.9% in 1990 to 8.7% in 2018.

The second contribution of this paper is a novel methodology to identify horizontal mergers. I train a natural language processing model based on the whole corpus of US patents.[8] This is accomplished with a word embeddings algorithm, which is trained on the title and abstract of each patent. Certain mergers are subsequently classified as horizontal based on how similar the patents that are published by the merging parties. Antitrust authorities have access to internal documents of the merging parties. I approximate their classification using the information contained in the patents. As a validation of the methodology, this definition of horizontal mergers matches the classification of public merger cases of the European Commission and the Federal Trade Commission.

The third contribution of this paper is an approachable theoretical framework to explain the deterrence mechanism underlying the results. Deterrence means that certain firms do not attempt to merge because they know that the authorities will block them. After the policy change, hundreds of mergers were no longer subject to deterrence, as they were not notified. Potentially anticompetitive mergers were thus implemented in great numbers. The effect of a merger on innovation depends on the balance between generated efficiencies and loss of competition. Differentiated products and competition in cost-reducing innovation imply that not all anticompetitive mergers are detrimental to innovation.[9] As the effect of these new mergers on innovation is a priori ambiguous, then so is the effect of looser notification rules. This warrants an empirical study to ascertain the matter.

[8]The corpus of patents published by the United States Patents and Trademark Office and accessed through PatentsView counts approximately 7 million patents.

[9]This is a general result for the class of models described by Motta and Tarantino (2017). The choice of a model of cost reducing innovation à la is justified by results on process innovation, which is the kind of innovation that increases productivity.

I use patenting activity as a firm level measure of innovation.[10] In a descriptive analysis, mergers lead to an average decrease in innovation activity of 28%. However, there is high heterogeneity around this figure, as 36% of these mergers raise innovation.[11] Focusing on mergers that are affected by the policy change, the results of my difference in differences imply that non-notified horizontal mergers lead to a further 30% less innovation. A stronger effect in the short run is consistent with the hurried realization of mergers that are not subject to deterrence. Moreover, the decrease in innovation is characterized by a decline in the quality and originality of published patents.[12] A decrease in process innovation drives the results, while I find less pronounced effects on product innovation. This suggests that affected firms become less productive after these mergers. In addition, I show that the number of horizontal mergers that are not reported to the authorities increases after the policy change, in accordance with deterrence being the mechanism behind the main results.[13]

The main results of this paper are robust to a variety of specification changes. Since a large portion of the sample is comprised of Pharma and Big Tech mergers, I show that results are robust to the exclusion of one or the other from the sample. Given that my definition of horizontal merger based on patent similarity is new to the literature, I propose several variations of this definition, all leading to similar results. I consider also a continuous measure of patent similarity between firms as identification device, which could be interpreted as a measure of "horizontality". Using this, innovation effects are even stronger. Moreover, declining innovation on the side of the acquirer drives the results. Furthermore, I find that innovation declines not only for firms involved in mergers, but also at the industry level. In particular, I show that after the policy change, affected industries become more concentrated, generate more profits and spend less on R&D.

The model yields several predictions for the outcomes of the policy change studied in this paper. First, it implies an increase in the number of horizontal mergers, which is precisely what I observe in the data. Second, since mergers on average decrease innovation, the model predicts less innovation activity after these horizontal mergers. This corresponds with the main results of this work: affected mergers lead to less innovation. Lastly, in the model the effect on the consumer surplus is worse than the effect on innovation. Even mergers that generate enough efficiencies to leave innovation unchanged can be detrimental to consumers. Therefore, there are mergers that increase innovation but harm the consumer surplus, while all mergers that lower innovation also hurt consumers. As a consequence, a negative effect on innovation after the policy change implies a decrease in the consumer surplus, which is indeed what the antitrust authorities seek to prevent.

[10] The main measure I use is called "relative citation average", and Lerner et al. (2011) show that it accounts for differences in popularity between different technology fields. In popular fields, patents may receive many more citations on average. The results hold for several other measures of innovation.

[11] Non-horizontal mergers in sectors such as software programming and computer manufacturing have a high share of innovation improving outcomes.

[12] I find that results are driven by a drop in the number of citations per patent, rather than a contraction in the number of patents. I interpret this as a sign of dwindling innovation quality. Moreover, I show that issued patents tend to cite a less diverse array of other patents, which is considered a drop in patent originality by Lerner et al. (2011).

[13] This result was already discussed extensively by Wollmann (2019) in his first paper on stealth consolidation.

Contribution to the Literature

This paper contributes to the wide literature of competition and innovation by focusing on abrupt changes in competition generated by mergers and acquisitions. Most of the papers in the literature study how the level of competition affects innovation activity of competing firms, among many Aghion et al. (2005), Acemoglu and Akcigit (2012), Gutiérrez and Philippon (2017), De Ridder (2020). My approach to mergers and innovation is empirical. In recent years papers such as Federico et al. (2018) and Denicolò and Polo (2021) present models to understand the effect of mergers on innovation of merging parties and their competitors. Motta and Tarantino (2017) outline a model of competition with cost reducing innovation on top of which I build my model of deterrence. Jullien and Lefouili (2018), on the other hand, propose a model of competition with demand enhancing innovation. My paper proposes an identification strategy based on a relaxation of pre-merger notification rules. Haucap et al. (2019) find empirically that mergers lead to less innovation by comparing merging parties with similar non-merging firms chosen with a matching procedure.

In this paper I find that acquirers tend to decrease innovation effort even more that acquired firms. This relates to the recent literature on reverse killer acquisitions spurred by Caffarra et al. (2020). Cunningham et al. (2019) show that incumbents can find it optimal to stop the development process of acquired start-ups, giving rise to killer acquisitions. My paper reaches similar conclusions for the innovation activity of merging firms, but it focuses on patent creation, a process that comes before product development. Interestingly, the authors find that a significant portion of killer acquisitions happen below notification thresholds, and thus they are not reported to the antitrust authorities.

In general, this paper contributes the wider literature of antitrust policy (Miller (2009), Besley et al. (2021)) with a natural experiment on a relaxation of merger policy. The seminal work of Wollmann (2019) started the literature on stealth consolidation and notification rules. My paper studies the effect of stealth consolidation on innovation. The methodology of this work build upon Wollmann (2021), and I provide a novel way to identify horizontal mergers using patents.[14] My theoretical model shows how antitrust authorities can benefit innovation by deterring harmful mergers. Fumagalli et al. (2020) study optimal antitrust policy when start-ups face financial frictions that can be overcame by an acquiring incumbent. Mermelstein et al. (2018) describe a model of competition with capital accumulation and derive the optimal antitrust policy.[15]

The present paper also contributes to the literature on the deterrence effects of antitrust policy. I focus on merger control, one of the prominent application of antitrust policy.[16] In their seminal work, Besanko and Spulber (1989) build a model of cartel enforcement under asymmetric information. I approach my research question exploiting a natural experiment, identifying a suitable coun-

[14] Wollmann (2021) use single mergers as events in a staggered diff-in-diff, while I use the amendment as a single event.

[15] They find that no antitrust scrutiny is never the optimal, while the optimal policy blocks most of the proposed mergers.

[16] Deterrence is a relevant issue also for the antitrust law literature (see Breit and Elzinga (1973), Baker (1988), Wils (2006), Lande and Davis (2011)).

terfactual for a difference in differences analysis. Miller (2009) studies leniency toward early confessors of cartel behavior. Barrios and Wollmann (2022) incorporate deterrence effects in a model of investor disclosure of merger transactions that may alert antitrust authorities. They find evidence that deterrence is more effective on horizontal mergers, similarly to what I find in my work. [17]

Furthermore, this work contributes to the literature of text analysis in economics (among many recent papers Iaria et al. (2018), Ash et al. (2022), Decarolis and Giorgiantonio (2022)). In this work I use Doc2Vec, a natural language processing tool that exploits word embeddings, and it was first described by Mikolov et al. (2013a) and Mikolov et al. (2013b). I use this algorithm to identify horizontal mergers. I devise a new methodology based on patents that can be applied to small and private firms. A paper with a similar methodology is Hoberg and Phillips (2016), in which the authors use product descriptions for public firms to determine a network of product differentiation. My methodology can be applied to all firms in the economy. Moreover, I use more modern semantic techniques such as word embeddings.[18]

The rest of this paper is organized as follows. Section 2 describes the theoretical framework which provides a mechanism to interpret the results. Section 3 presents data and variables used in the analysis. Section 4 describes the empirical methodology of the natural experiment. Section 5 proposes the main results. Section 6 discusses features and limitations of the main analysis, and provides robustness checks and sensitivity analysis. Section 7 concludes.

[17] Despite the available evidence, deterrence capabilities of antitrust authorities are a contentious issue in the literature. Eckbo (1992) finds no evidence of deterrence comparing US mergers to Canada mergers. Crandall and Winston (2003) find no evidence that antitrust policy deterred firms from engaging in anticompetitive actions, and in some instances they find evidence that it may have lowered consumer welfare.

[18] The authors use term-frequency of each term for a patent, scaled by the inverse document-frequency of each term across the corpus. This methodology is usually called *Tf-idf*, and it is very computationally intensive. Younge and Kuhn (2015) describe a vector space model to compute patent similarity using text analysis methodologies, and they use word counting. In this paper I use more modern techniques based on word embeddings, which are both efficient and more effective at representing semantic meaning.

2 Theoretical Framework

To better understand the implications of changing notification thresholds, I propose an easily approachable theoretical framework. This model features an antitrust authority, firms acquisition, innovation, and product market competition. The model shows that raising notification thresholds has an ambiguous effect on innovation, which can only be determined by an empirical investigation. Appendix A.2 describes a more general model that relaxes several simplifying assumptions of this section, reaching similar conclusions.

There are two symmetric firms i, j that produce differentiated goods and they compete in quantity q and innovation x. Innovation x decreases marginal costs linearly $c(x) = c - x$, and it requires a fixed cost $F_b(x) = x^2/2$. The profit of firm i before the merger is described by equation 1. Demand is linear $p(q_i, q_j) = \alpha - q_i - \gamma q_j$, where a larger $\gamma \in [0, 1]$ means that the goods are more easily substitutes. The FOC with respect to x_i easily implies that $x_i = q_i$. Optimal quantity is $q_i = q_j = \frac{\alpha - c}{1+\gamma}$, and optimal price is $p = c$. Firms still make a profit $\pi_{b,i} = \frac{(\alpha-c)^2}{2(1+\gamma)^2}$, because the price is larger than the actual marginal cost $c(x_i) = c - x_i$.

$$\pi_{b,i} = max_{q_i, x_i}(p(q_i, q_j) - (c - x_i))q_i - x_i^2/2 \tag{1}$$

$$2\pi_M = max_{q_i, x_i, q_j, x_j}(p(q_i, q_j) - (c - x_i))q_i + (p(q_j, q_i) - (c - x_j))q_j - (1-\lambda)x_i^2/2 - (1-\lambda)x_j^2/2 \tag{2}$$

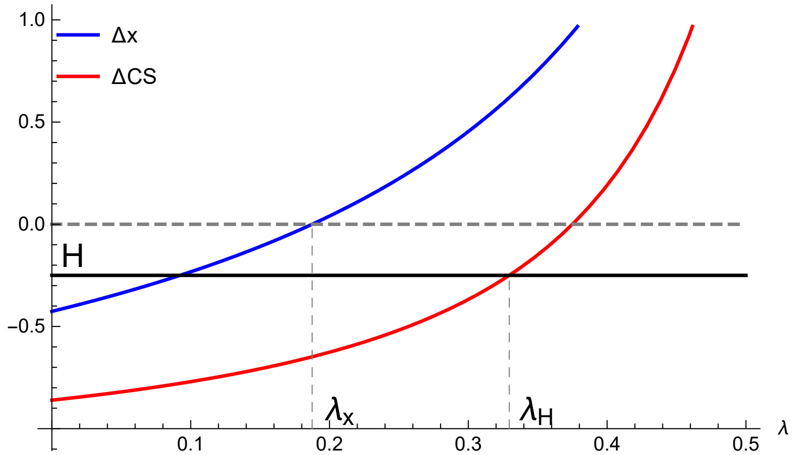
In the event of a merger, the resulting firm maximizes profits generated by the two differentiated products $2\pi_M = \pi_{M,i} + \pi_{M,j}$. Moreover, a merger generates efficiencies in the fixed cost of innovation $F_M(x) = (1-\lambda)x^2/2$, where $\lambda \in [0, 1]$ represents the extent of fixed cost savings. Equation 2 describes the profit of the merged entity. The FOC with respect to innovation implies that $x_i = \frac{q_i}{1-\lambda}$. The FOC with respect to q_i is reported in equation 3, and it equals the FOC before the merger with two additional parts. The efficiency effect is increasing in λ, and it encourages the merged entity to produce and innovate more. On the contrary, the internalization of the negative competition externality that q_i has on sales of product j pushes the merged firm to decrease quantity and innovation. If the efficiency effect prevails, the merger will be beneficial to innovation and to consumers.

$$\partial_{q_i}\pi_M = \underbrace{(\partial_{q_i}p(q_i, q_j))q_i + p(q_i, q_j) - c + q_i}_{\text{FOC before the merger}} + \underbrace{\frac{\lambda}{1-\lambda}q_i}_{\substack{\text{efficiency}\\ \text{effect}}} + \underbrace{(\partial_{q_i}p(q_j, q_i))q_j}_{\substack{\text{internalization of}\\ \text{competition externality}}} = 0 \tag{3}$$

Figure 2 shows that for higher efficiencies λ the effect of a merger on innovation $\Delta x = x_M - x_b$ improves. The same goes for the effect of a merger on consumer surplus ΔCS. Appendix A.1 reports the closed form values of both these effects. The antitrust authority tolerates a level of harm H and blocks any merger such that $\Delta CS < H$. Knowing this, firms with a level of efficiencies $\lambda \in (0, \lambda_H)$ are deterred, and they do not attempt to merge.

After the increase in the notification threshold studied in this paper, the antitrust authority no longer investigates these mergers. This is equivalent to setting an extremely low tolerated level of harm $\tilde{H} << H$, allowing all mergers. After the policy change, all previously deterred mergers with

Figure 2: Effect of merger on innovation and consumer surplus.



Notes: Changes in innovation x and consumer surplus CS depending on efficiencies λ. The horizontal axis reports values of the efficiencies parameter λ. The black line reports the level of harm H that the antitrust authority is willing to tolerate. λ_x is the level of efficiencies for which innovation does not change after the merger ($\Delta x(\lambda_x) = 0$). λ_H is the level of efficiencies under which the antitrust authority blocks a merger ($\Delta CS(\lambda_H) = H$).

$\lambda \in (0, \lambda_H)$ are implemented, generating an increase in the number of realized mergers. The previously deterred mergers with low efficiencies $\lambda \in (0, \lambda_x)$ are detrimental innovation, while the previously deterred mergers with high efficiencies $\lambda \in (\lambda_x, \lambda_H)$ are beneficial to innovation. Therefore, the overall effect of the policy change on innovation is ambiguous, and it depends on the distribution of efficiencies among mergers. The effect of a merger on the consumer surplus ΔCS is worse than its effect on innovation Δx.[19] Therefore, it is possible to infer that the policy change lowered the consumer surplus from the empirical evidence that the policy change decreased innovation.

Although this model relies on several assumptions, it captures the fundamental dynamics of innovation and consumer surplus after the policy change studied in this paper. Appendix A.2.1 extends this model to price competition and multiple asymmetric firms a là Motta and Tarantino (2017). Appendix A.2.2 describes how deterrence arises from profit maximization. Moreover, Appendix A.2.3 introduces an optimal antitrust policy that exploits deterrence with imperfect information, while Appendix A.2.4 describes optimal antitrust policy with imperfect enforcement. Appendix A.2.5 shows that the same conclusions hold in this more complex framework.

3 Data and Variables

There are two main sources of data for this work: mergers and patents. All data refer to the United States. Since the identification strategy of this paper relies on mergers being sufficiently small to go under the radar of the authorities, these data sources must be comprehensive of both public and private held companies.

[19]This is a general property of the class of models described by Motta and Tarantino (2017). This is due to the internalization of negative price externalities on rival sales. This externality is still present, even for mergers that generate enough efficiencies to leave innovation unchanged.

Data on the universe of Merger and Acquisitions come from Thomson Reuters SDC Platinum. This includes mergers of both public and private companies, and it has the advantage of covering even small transactions. For each merger the researcher can access the date of completion and information on the merging parties.[20] For both firms I can access the name, which I use to assign them patents. Moreover, I have access to balance sheet figures and the value of the transaction, which I use to determine if they are required to report the merger to the authorities.[21] Moreover, I gather the state of residence and SIC industry codes, which I use to control for firm heterogeneity. This is the same data source used by Wollmann (2019) to describe stealth consolidation.[22]

I access data on the universe of patents published by the U.S. Patent & Trademark Office (USPTO) through PatentsView. [23] From this database I can access about 7 million patents from 1976 up to the present. For each patent I observe the date of submission and the date of publication, but in my analysis I use the date of submission, as the publication process often takes more than five years. In order to assign a patent to a merging firm I gather the name of the "Assignee", which is the company that owns legal rights related to the patent. For my text analysis exercise I use both titles and abstracts of patents, which I combine in an unique document for each patent. A patent abstract is much similar to a paper abstract in length and content. Furthermore, I can access citations for each patent, which I use to evaluate the quality of innovation. Moreover, I gather each patent's ICP classification, which determines the technology field of each patent.[24]

In order to compute innovation activity of each firm the literature has used several measures derived from patents. Given that patents receive citations similarly to academic works, one can use these citations as a proxy for patent quality. So, rather than the number of patent submitted by a firm each year, one can compute the total number of citations received by these patents. Some technology classes might be more active than others, however, and this might inflate citation numbers for patents in these classes. In order to make patents comparable across technological classes, Lerner et al. (2011) propose a measure of relative citation activity. This is computed as the number of citation received by a patent, divided by the average amount of citations received by patents submitted in the same technological field in the same year. Then, for each firm I compute innovation activity as the average of this relative citation intensity of each patent submitted in a given year.

[20] For the empirical analysis, I consider only completed mergers and I consider the completion date as merger date. Announcement date are available in the dataset, but I do not use them.

[21] For each firm I can access Net and Gross Assets, Income, Turnover for the fiscal year of the merger.

[22] Similarly to Wollmann (2019), I drop some sectors that might show peculiar merger behavior. This includes oil, gas, real estate, and banking. As these sectors tend not to produce patents, this is not affecting my results in any significant way.

[23] PatentsView offers free access to USPTO databases, and it is build specifically for researchers. PatentsView began in 2012 as a team project with the USPTO, American Institutes for Research, University of Massachusetts Amherst, New York University, University of California, Berkeley, Twin Arch Technologies, and Periscopic.

[24] This has a similar purpose to the SIC classification for firms, and it is hierarchical as well.

4 Empirical Methodology

Identifying the effect that Antitrust Policy has on the innovation activity of merging parties requires to identify which mergers are controlled by the authorities and which firms benefit the most from not reporting to the authorities. In this empirical analysis I exploit a change in merger policy that made thousands of mergers exempt from notifying to the authorities. I can compare mergers that become exempt from reporting with mergers that are not affected by this policy change, before and after this policy change, to see if the control of antitrust authorities influences innovation decisions of merging firms.

Moreover, the authorities are much more likely to scrutinize and even block horizontal mergers rather than non-horizontal ones.[25] Horizontal mergers are defined as transactions involving firms operating in the same product markets. Therefore, firms engaged in horizontal mergers are the ones benefiting the most from a possible exemption, since they are the ones carefully controlled by the authorities. Consequently, I can compare firms engaged in horizontal mergers with firm engaged in non-horizontal ones, as a further layer of my identification strategy. How can one identify these horizontal mergers, especially given the limited amount of data that is available for private firms?

4.1 Identify Horizontal Mergers

Identifying horizontal mergers is a challenging task. Most merger cases hang on the definition of relevant markets and actual competitors, and this constitutes a large portion of the work of antitrust authorities and M&A consultants. The authorities have access to a large amount of internal, private documents from both merging firms, and they use them to determine if merging firms are close competitors. In this work I approximate this analysis using the available information I have on patents published by the merging parties. In particular, I compare the abstract of patents owned by merging firms to determine how similar their product lines are. The underlying hypothesis is that firms with similar patents are likely to generate similar products. Patents are both an outcome of firms' innovation process and an essential input of their production process, and thus they contain information on firms' product lines.

In more detail, I use natural language processing to automatically compare the abstract of patents. I train a machine learning model on the universe of patents published in the US. The use of pre-trained models is not warranted, because patents use terms and syntactic structures that are different from standard prose. In particular, I make sure to include also very exotic and infrequent terms in the analysis, since they can represent new products or technologies. This is similar in spirit to what Hoberg and Phillips (2016) do with product descriptions on public firms in COMPUSTAT.[26] The use of patent abstracts allows me to extend this methodology to private firms, which are not covered in

[25] This is such an integral characteristic of antitrust policy that the official 2010 guidelines for merger control in the US are called *Horizontal Merger Guidelines* . See Wollmann (2019) or Wollmann (2021) for more discussion on this matter.

[26] Hoberg and Phillips (2016) use product descriptions to create a new definition of industries. What I do in my work is different in two respects. First, I use the abstract of patents, rather than single product descriptions. Patent abstracts contain more information, but at the same time they are less directly related to products. Second, I am not interested in the definition of industries, rather, I am concerned with similarity between two merging firms.

conventional data-sets and for which information is quite scarce. Patent are carefully collected by the USPTO and there is no discrimination on patent assignee, so that even the smallest private firm has its patents registered.

In order to validate my methodology, then, I predict how the European Commission and the Federal Trade Commission classify mergers in its public decisions, horizontal or non-horizontal. I show that using firm similarity computed with patent abstracts allows me to outperform the Standard Industry Classification (SIC), which has been used in the literature to identify horizontal mergers.[27]

4.1.1 Text Analysis Exercise

To compare two merging firms I need to compare their patents' abstract, and in order to do so I transform texts into comparable objects. Most natural language processing algorithms represent words and texts as vectors of real numbers, so that a notion of distance between texts can be defined as the distance between their representative vectors. In this work I use Doc2Vec, a natural language processing tool that exploits word embeddings, and was published in Mikolov et al. (2013a) and Mikolov et al. (2013b). I describe this algorithm in more details in Appendix B.1. For the purpose of this work it is sufficient to understand that each patent abstract is associated to a vector P_i of 300 real numbers, which is optimized to represent the semantic content of the abstract itself. If two abstract contain exactly the same text, then they are represented by the same vector, and the more similar two abstracts are, the more similar will be their representative vectors.

I compute similarity between patents i and k as cosine similarity of their representative vectors P_i and P_k, as it is standard in the literature. Equation 4 shows that cosine similarity is a generalization of the cosine of the angle between the two vectors, and intuitively it can be considered as their correlation. This similarity measure is lower than one, and it is equal to one if the two vectors are exactly the same.[28]

$$CS_{ik} := \frac{P_i \cdot P_k}{||P_i||\,||P_k||} = \frac{\sum_j P_{ij} P_{kj}}{\sqrt{\sum_j P_{ij}^2}\sqrt{\sum_j P_{kj}^2}} \leq 1 \tag{4}$$

This process is quite standard in natural language processing, but it might appear obscure from the outside. Therefore I propose an example to explain the meaning of semantic similarity. In 2003 Pfizer acquired Pharmacia for $60 billion. This transaction was obviously reported to the FTC, that classified it as an horizontal merger and allowed it.[29] Both these companies have thousands of patents and, among all possible pairs, the two patents with the highest cosine similarity (above 0.95) are the ones shown in Figure 3.

By reading the title and the abstract of these patents it is clear that they refer to chemical compounds with pharmaceutical applications. This is not surprising, given that both Pfizer and Phar-

[27] See Wollmann (2019) for an example.

[28] If two vectors are exactly the same, their angle is 0, and the cosine of 0 is 1. If two vectors are orthogonal the cosine will be 0.

[29] The merger was permitted with some divestitures, for more information one can look on the official FTC website.

Figure 3: Patents with highest similarity between Pfizer and Pharmacia in 2003.

Pfizer

US 6586430 B1: CCR5 modulators
"Compounds.. which are useful as modulators of chemokine activity. The invention also provides pharmaceutical formulations and methods of treatment using these compounds."

[Filed: Dec 1, 1999]

Pharmacia

US 6809111 B2 : Prodrugs of COX-2 inhibitors
"A compound of... or a pharmaceutically-acceptable salt thereof, suitable for use in the treatment of a cyclooxygenase-2 mediated disease is provided... and a method for treatment of a cyclooxygenase-2 mediated disease..."

[Filed: May 15, 2003]

Notes: Two patents with the highest text similarity between all patents filed by Pfixer and all patents filed by Pharmacia at the date of the merger. The identifying codes can be used on lens.org to look for additional information regarding the patent, including the whole abstract. I encourage the reader to do so, just to have a better feeling of these patents.

macia are major pharmaceutical companies. At first, it might seem that the high similarity between the two patents is due to common words, such as "compounds", "pharmaceutical" and "method of treatment". This is how a commonly used term frequency inverse document frequency (TF-IDF) text analysis algorithm would work. If this were the case, however, it would not be sufficient for my exercise. Any patents relating to pharmaceutical products would have high similarity, even if this is too broad to define a product market.

But there is a reason if this particular couple of patents is the one with the highest similarity. Zeidler et al. (2000) mention that COX-2 inhibitors modulate chemokine receptors CCR5, and in doing so they are effective in treating tumor patients. Such a connection between COX-2 inhibitors and CCR5 receptors is surely known to the practitioners, even though it is not apparent from the text alone. This example highlights the advantage of word embeddings methodologies with respect to more traditional tools such as word counting and TF-IDF weighting. Natural language processing tools such as Doc2Vec recover the semantic meaning of a word from the terms that are used close to it. Evidently, in the whole corpus of 7 million patents, "COX-2 inhibitors" and "CCR5 modulators" are often used in a similar context. To be confident that this is not a fluke, in the Appendix B.2 I document that the couple of patents with the second highest similarity follows a similar scheme.

When two innovating firms merge they both have a collection of patents on which I can compute cosine similarity.[30] From the list of all pairwise similarities one can gather a lot of information on the relation between two merging firms. However, most of the information is contained into patents that are most similar, those that show the two firms are operating is the same markets. Therefore, I can measure similarity between two merging firms using the highest values of similarities between their patent portfolios. Some examples are the maximum similarity (hereafter Max), the mean of

[30]In my sample of merging firms the average number of patents owned by a firm at the year of the merger is about 13000, and the median is bout 8000.

the top 20 similarities (hereafter $Max\ 20$), the mean of the top $x\%$ similarities (hereafter $Max\ x\%$). All these measures are meant to represent the same concept, and as such they are very correlated.[31] The main results of this paper are produced using the $Max\ 2\%$ similarity, but they hold for all other measures. These are continuous measures of similarity between two merging firms, and they can be used already as an identification device. Moreover, I identify horizontal mergers as the top quartile of the distribution of similarities across all observed mergers.

4.1.2 Predicting official decisions of the authorities

In order to evaluate patent similarity statistics I use them to predict horizontal mergers as defined by antitrust authorities. This exercise is conducted on both decisions of the Directorate General Competition of the European Commission (hereafter EC) as well as decisions of the Federal Trade Commission (hereafter FTC). Here I report results pertaining to the EC, while the exercise on FTC decisions is described in Appendix B.5. Both exercises lead to similar conclusions. With regard to EC decisions, they are collected from the database developed by Affeldt et al. (2021). Data on almost all merger control decisions by the EC is gathered by hand from legal decision documents.[32] I consider a merger to be horizontal if it is not tagged as vertical and it is not tagged as conglomerate. From the original pool of public decisions I remove mergers between companies that do not have a portfolio of patents. Moreover, I remove transactions that are not considered full mergers.[33] The database is organized by markets, and each merger can influence several markets. As a result there are 111 mergers and 568 markets influenced by EC decisions, in 485 (85%) of these markets the merger are horizontal, while in 83 (15%) they are non-horizontal. This is the sample used in the validation exercise.

For this set of markets controlled by the EC I build a dummy variable that is 1 for horizontal mergers, and 0 otherwise. Then I build a dummy variable that is 1 if the merging parties have the same 4 digits SIC code, and 0 otherwise. This represents the standard in the existing literature, as one can see in Wollmann (2019), and this is the one I compare my measures with. As a first step I compute correlations of these variables in Table 1 in Appendix B.4, to see which one is most similar to FTC definitions. The SIC definition is positively correlated with FTC definition, but with a small value of 0.17. Patent similarity measures have a higher correlation, outperforming the SIC one. It is worth noting that these measures have a high correlation with the SIC dummy. Moreover, all these similarity measures have an even stronger correlation between each other, since they are representing the same concept.

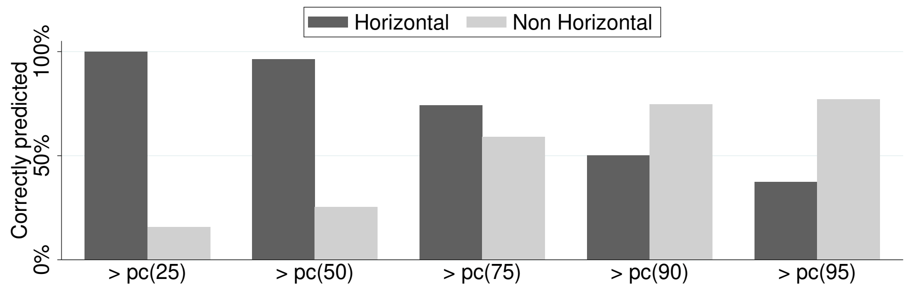
The correlation table compares a dummy variable for EC decisions with continuous measures of patent similarity. These measures are informative by themselves, and they can be used as an identification device. As a robustness exercise I show that using continuous measures of similarity in the identification strategy leads to results similar to the main ones. However, if one wants to gen-

[31] Table 1 on Appendix B.4 shows the correlation table of these variables.

[32] This includes all cases settled in the first phase of an investigation (Art. 6(1)(a), 6(1)(b), 6(1)(c) and 6(2)) and all cases decided in the second phase of an investigation (Art. 8(1), 8(2), and 8(3)). Note that this also includes all cases settled under a 'simplified procedure', provided that a legal decision document exists. More information on the database can be found on the official DIW website.

[33] In the database transactions are considered either full mergers or joint ventures.

Figure 4: Performance of different cutoff rules on EC Decisions.



Notes: Percentage of correctly predicted horizontal (dark grey) and non-horizontal (light grey) mergers, as defined by the EC, with increasing cutoff. The cutoff is in percentiles of the statistic distribution. The exercise is performed on the max 2% statistic, and similar patterns emerge with other statistics. Numbers for the histogram are reported in Table 2 in Appendix B.4

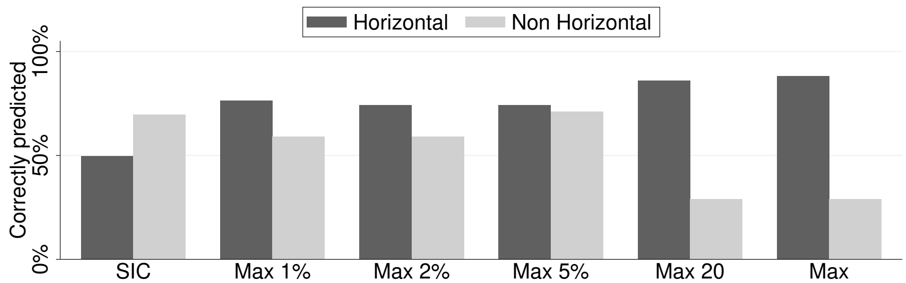
erate a 0-1 dummy variable identifying horizontal mergers using similarity statistics, one needs to determine a threshold above which a merger is considered horizontal. Figure 4 reports variables constructed with various thresholds compared with FTC definitions. Each bar represents the percentage of correct predictions. This figure represent type I and type II errors in predicting horizontal mergers. A lower cutoff, like the 25^{th} percentile is very accurate in predicting horizontal mergers, but does poorly in predicting non-horizontal ones. Conversely, a cutoff like the 95^{th} percentile predicts horizontal mergers poorly. The most reasonable cutoff is the 75^{th} percentile, and this is consistent across various similarity measures. This is also consistent for FTC decisions, as it is shown in Appendix B.5.

Once the cutoff rule is set to the 75^{th} percentile, I compare similarity statistics in Figure 5. Using the SIC industry classification one can predict only 50% of horizontal mergers, while all patent similarity statistics outperform this measure. Similarly to the correlation results in Table 1 in Appendix B.4, the *Max* $x\%$ statistics perform better than the simple maximum value of the similarity matrix. This is the case also for FTC decisions. In the Robustness section I show that all results hold true regardless of the chosen patent similarity statistic. This is to be expected, as all these measures capture the same concept: how close are the products of two merging firms.

4.2 Policy Change

In many jurisdictions, including the US, merging parties are exempted from reporting their transaction to the authorities if they are economically small. The rationale behind this is that mergers between small companies are expected to have little implications on affected markets. Consequently, the legislator prefers to spend resources on larger mergers. In practice, the Hart-Scott-Rodino Act set a threshold under which merging parties are exempt from reporting to the authorities in the US. They can merge, and the Antitrust Authority may never know about their transaction. If the Authority is informed of the transaction by other actors it can investigate the merger, but this is a very rare occurrence, given the already burdensome amount of work given by other mergers. As a matter of fact, one can consider these exempt mergers as transactions that are not controlled by the authorities.

Figure 5: Performance of Similarity Statistics on EC Decisions.



Notes: Percentage of correctly predicted horizontal (dark grey) and non-horizontal (light grey) mergers, as defined by the EC, with different methodologies. SIC stands for horizontal mergers defined using the same 4 digit SIC code, as in Wollmann (2019). The exercise is performed using the $> pc(75)$ cutoff, and similar patterns emerge with other cutoffs. Numbers for the histogram are reported in Table 3 in Appendix B.4.

On December 20, 2000 an Amendment to the Hart-Scott-Rodino Act raised these pre-meger reporting thresholds. Before the Amendment, deals whose target assets were below $\$10$ million were exempt from notifying their transaction.[34] This is commonly referred to as the "size of person test". The Amendment increased the "size of person" test from $\$10$ million to $\$50$ million. Even if it were a sizable increase, this did not affect reporting considerably. What made a significant difference was the introduction of a new "size of transaction" test, which made exempt any merger whose transaction value was below $\$50$ million.[35] This new "size of transaction" threshold, which was effectively raised from $\$0$ to $\$50$ million, explains the bulk of the 70% decrease in merger notifications to the authorities, which fell from more than 4000 per year to about 1000.[36] Figure 6 reports the number of merger notifications in years around the Amendment. Dark grey bars represents mergers that fall below the new "size of person" threshold of $\$50$ million. These mergers were more than half of the transactions reported to the authorities before the Amendment.

As these mergers involve small transactions, it is natural to ask whether they were actually posing any competitive concern for the authorities. Figure 8 in the Appendix C shows that Second Requests issued by the authorities also fell considerably after the policy change, indicating that antitrust enforcement decreased substantially as well.[37] Moreover, there is no evidence of the authorities giving more attention to notified mergers after the policy change. Figure 9 in Appendix C shows that the share of notifications resulting in second requests decreased from about 5% to 3% after the relaxation of notification rules. This happened despite the number of notification decreased drastically.

In order to use the Amendment to identify changes in competition and innovation outcomes the policy change itself must be exogenous to these variables. If the main reason behind the Amend-

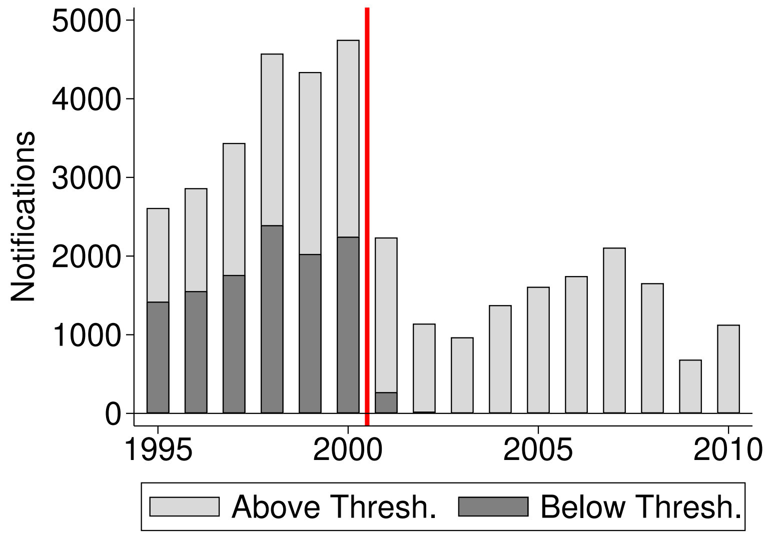
[34] If the target was engaged in manufacturing, also sales were required to be under $\$10$ million

[35] Both the "size of person" and the "size of transaction" thresholds were indexed to the US GNI index, so that now they are much higher. For 2022, that threshold will be $\$101$ million, as explained on the FTC website.

[36] For a more in-depth description of the Amendment to the Hart-Scott-Rodino Act one can refer to Wollmann (2019), who was the first to study this policy change.

[37] Second Requests are issued by the FTC or the DOJ when they want to gather more information after the first 30 days they are given to investigate a merger. Second Requests are a better measure of enforcement, as the authorities typically engage in negotiations with merging companies to meet the specific needs of the investigation.

Figure 6: Number of Notifications received by US Antitrust Authorities.



Notes: The graph reports number of notifications above and below the new threshold of $50 million introduced with the amendment in December 2000. The red vertical line represents the introduction of the amendment.

ment were to focus the attention of the authorities on innovating firms, then this could provide an alternative explanation for my results. This is not the case, however, as the policy change was a response to complaints that the 25-year old threshold was too low, as it was not adjusted to inflation.[38] A second motivation driving the policy change was to make the merger review process more efficient, so as to save taxpayer money and company resources.[39] Moreover, there was no anticipation of the Amendment from merging parties and consultancy firms. The business press largely ignored the new standards and the only news covering the policy change were published after it was voted.[40] A further concern is that firms might be manipulating their numbers to fall below the threshold and avoid antitrust scrutiny. This would introduce a selection bias in my sample that could explain my results. Lowering the amount paid for the transaction is problematic, though, as the acquired company may refuse the deal. Moreover, in the data there is no evidence of bunching below the new threshold of $50 million.[41]

This policy change allows me to identify two categories of mergers, those that are affected by the Amendment, and those that are not. Newly-Exempt mergers are those transactions that were not ex-

[38] Citing directly from the Competition Committee, Directorate for Financial and Enterprise Affairs, 2016, ""In response to complaints that the *25-year old* [...] threshold had become too low, Congress increased it to $50 million and indexed it to GDP." https://one.oecd.org/document/DAF/COMP/WP3/WD(2016)22/en/pdf

[39] Citing directly from the Competition Committee, Directorate for Financial and Enterprise Affairs, 2016, ""The U.S. agencies continually assess how the review process can be made *more efficient* and how the agencies can *reduce the costs and burdens* on parties." https://one.oecd.org/document/DAF/COMP/WP3/WD(2016)22/en/pdf

[40] Moreover, the amendment passed as four pages buried within the 320-page omnibus bill for the Fiscal Year 2001 Commerce-Justice-State Appropriations Bill. I was voted near the end of the Clinton Presidency, on December 21.

[41] Figure 10 in Appendix C shows that there is a spike of mergers just below or at the threshold of $50 million, but this spike is actually smaller than the one of $40 or $35 million. These spikes are due to round numbers, rather than price manipulations.

empt from reporting before the amendment, but they became exempt thanks to the Amendment.[42] These are mergers that are affected by the policy change, and I consider them as my treated group. On the other hand, Never-Exempt mergers are the ones reported to the authorities both before and after the Amendment.[43] By definition, these mergers are not affected by the Amendment due to their larger size. Consequently, I consider them as the control group. Finally, I exclude from the analysis mergers that were exempt from reporting to the authorities both before and after the policy change.[44].

4.3 Natural Experiment

The unit of analysis of my natural experiment is a merging firm, and I have a cross section of them in every year. The outcome of interest is the variation in patenting activity brought on by the merger. This is computed as the average log change $\Delta P_{it} = ln\,(\sum_{i=1}^{n} P_{t+i}/n) - ln\,(\sum_{i=1}^{n} P_{t-i}/n)$, which considers uniformly all years around the merger. As for the measure of patenting activity P_{it} I follow Lerner et al. (2011) and define a measure of relative citations, which is the number of citations received by a patent divided by the average number of citations received by patents in the same technological space in the same year.[45] Patenting activity of merging firms is then the average of this relative citation measure.

Table 1 reports various statistics for this relative citation average in the main sample. The first row reports unconditional moments for the whole sample. On average mergers generate a decrease of innovation of 0.327 log points, which equals a 28% drop. The effect of mergers is quite heterogeneous, though, as the standard deviation is high, and the 90^{th} percentile shows that some mergers can increase innovation substantially. Figure 11 in Appendix C shows that 36% of mergers actually lead to an increase in innovation.[46] A clear difference arises when the sample is split between exempt and non exempt mergers. The ones that are not notified to the authorities have a negative effect that is much stronger than the ones that are notified. The difference is about 0.10 log points, or about 10%. Within the exempt mergers, the horizontal ones have an even more detrimental effect on innovation, and this effect is about 0.10 log points stronger than for non-horizontal ones. Since these are all unconditional moments they have no causal interpretation, and no sound conclusion can be derived from them. However it is worth noting that horizontal mergers that are not reported to the authorities have the most detrimental effect on innovation, and these are going to be the treated group of the natural experiment.

[42] Following the definition of Wollmann (2019), Newly-Exempt mergers are those whose transaction value is below $\$50$ million, or their target asset value is between $\$10$ million and $\$50$ million, or their target sales value is between $\$10$ million and $\$50$ million. In practice, for most of the mergers, the "size of transaction" test is the binding one.

[43] Never-Exempt mergers are defined as those in which transaction value is above $\$50$ million, target assets are above $\$50$ million and target sales are above $\$50$ million.

[44] These Always-Exempt mergers are the ones for which target assets were below $\$10$ million or target sales were below $\$10$ million.

[45] This allows me to compare patents is different technological spaces, even if some spaces are more active than others, meaning that on average patents receive more citations.

[46] Non-horizontal mergers have a higher share transactions with positive effects on innovation. Figure 12 in Appendix C shows that this share of non-horizontal mergers with positive effects is particularly high in Software Programming. This is likely due to high complementarities that can be exploited in this industry.

Table 1: Unconditional Innovation Change generated by Mergers.

	Mean	Std. Dev.	Median	p(10)	p(90)	N
All	-0.327	1.186	-0.295	-1.764	1.044	2601
Exempt	-0.383	1.279	-0.345	-1.974	1.121	1058
Non Exempt	-0.289	1.117	-0.258	-1.639	1.005	1543
Exempt						
Horizontal	-0.456	1.066	-0.374	-1.566	0.737	270
Non-Horizontal	-0.358	1.344	-0.330	-2.087	1.277	788
Non Exempt						
Horizontal	-0.261	0.981	-0.245	-1.398	0.786	471
Non-Horizontal	-0.302	1.172	-0.275	-1.747	1.065	1072

Notes: The table reports summary statistics for innovation change ΔP_{it} generated by mergers in the sample used for the natural experiment. P_{it} is computed as relative citation average. The first row reports statistics compute on all mergers. The subsequent rows reports the same statistics for different groups in the data. Both mergers before and after the Amendment are included in the statistics. In this table mergers are defined as exempt if they were actually exempt at the date of the merger. Therefore, this group does not correspond to the newly exempt mergers, which were exempt only in the years after the amendment, while before the amendment they were notified.

As an identification device I am going to use time variation generated by the Amendment. Consequently, the event is the policy change, it is not a single merger. Each merger is the result of a complex choice of the merging parties, and as such it could not be considered an exogenous event with respect to the innovation activity of the firms. As a matter of fact, my model of deterrence predicts an endogenous increase in non-notified horizontal mergers as consequence of the amendment. This is the main mechanism that drives the results, as these new mergers are also detrimental to innovation. The change in policy, conversely, is an exogenous event, since the main concerns driving the Amendment were unrelated to competitive and innovation outcomes, as discussed in Section 4.2.

As a consequence of this identification design, rather than a staggered difference in differences with several distinct events, I propose a difference in differences with a single event, the Amendment. As a first source of variation I exploit the difference between Newly-Exempt mergers and Never-Exempt mergers. As a second source, I use the variation between horizontal and non-horizontal mergers. The treatment group is composed by horizontal mergers that become exempt from reporting to the authorities. This treated mergers are compared with horizontal mergers that are controlled by the Antitrust Authority, and non-horizontal merges that become exempt from reporting to them.

This results in a triple difference in differences design, which I estimate by OLS following equation 5. I^{Post} is a dummy variable that is equal to 1 if the merger date is after the amendment, meaning years after 2001.[47] I^{Ex} is equal to 1 for Newly-Exempt mergers, and it is equal to 0 for Never-

[47] In the robustness section I show that results hold if I consider years after 2000 as the post period.

Exempt mergers.[48] I^{Hor} is a dummy variable representing horizontal mergers, as defined in Section 4.1. Therefore, the coefficient of interest is β, which represents the difference between Newly-Exempt horizontal mergers and the control group after the Amendment. α_t are year fixed effects, and ξ is the coefficient of additional controls.[49]

$$\begin{aligned}\Delta P_{it} =& \beta I^{Post} I^{Ex} I^{Hor} + \gamma I^{Post} I^{Ex} + \theta I^{Post} I^{Hor} + \delta_1 I^{Ex} I^{Hor} + \delta_2 I^{Ex} + \delta_3 I^{Hor} \\ &+ \alpha_t + \xi X_{it} + \epsilon_{it}\end{aligned} \tag{5}$$

5 Results

Table 2 reports results of the triple difference in difference described by equation 5. Column (2) reports results without clustering standard errors, while column (3) and (4) exclude some fixed effects. Column (1) is the baseline including all fixed effects, and the result for average change in patenting activity is negative and significant. On average innovation activity is 0.357 log points lower for firms involved in horizontal mergers that were not notified to the authorities. This is equivalent to 30% less innovation activity for merging firms that are affected by the Amendment. The size of this effect is comparable to the unconditional innovation change in the whole sample, 0.327 log points as shown in Table 1. Through the lenses of the model we can say that after the amendment some horizontal mergers that were deterred by the authorities now are attempted successfully. These mergers tend to be detrimental to consumers and to innovation, and they lower the average innovation effect of the treated group.

Table 7 in Appendix C shows coefficients from a simple difference in difference between horizontal and non-horizontal mergers. The exercise is conducted for newly exempt and never exempt mergers. The coefficient is not significant for both groups, and the point estimate is negative for newly exempt and positive for never exempt. What is crucial for the identification strategy, though, is that these two coefficients are significantly different. Indeed, the baseline coefficient in Table 2 is the difference between the two.

Rather than computing average effect using all years in the sample, Figure 7 shows triple diff-in-diff coefficients interacted with two year periods around the amendment. This shows that there is no significant difference between treated and untreated firms in the years before the amendment. The coefficients are negative and significant after the amendment. The effect seems to be concentrated in the short run after the amendment. This is confirmed by Figure 13 in Appendix C that shows coefficients in every single year. Figure 8 shows data points used to compute these results, implying that the difference between Newly Exempt mergers and Never Exempt mergers is negative only after the policy change. The figure reports the difference between horizontal and non-horizontal mergers, so as to represent the triple difference-in-differences results. From these exercises it is clear that the parallel trend assumption holds. There is no significant difference in trends between the treated and

[48] As previously explained, Always-Exempt mergers (the smallest ones) are excluded from the analysis.

[49] As additional controls I include SIC 2 digit industry fixed effects, role fixed effect (Acquiror or Target), and State fixed effects.

Table 2: Triple Difference in Differences Results

VARIABLES	(1) Baseline	(2) NoCluster	(3) No FE	(4) No FE
$I^{Post} \cdot I^{Ex} \cdot I^{Hor}$	-0.357**	-0.357**	-0.299**	-0.262*
	(0.152)	(0.175)	(0.148)	(0.148)
Implied Change	-30.0%	-30.0%	-25.8%	-23.0%
Observations	2,601	2,601	2,601	2,601
R-squared	0.080	0.080	0.054	0.020
Year FE	YES	YES	YES	YES
Industry FE	YES	YES	YES	.
State FE	YES	YES	.	.
Cluster SE	SIC4	.	SIC4	SIC4

Notes: This table reports coefficients of equation 5 with various control specifications. The dependent variable of the regression ΔP is computed using the relative citation average described in Section 3. The main specification is reported in column (1). Column (2) shows the main specification without clustering standard errors.

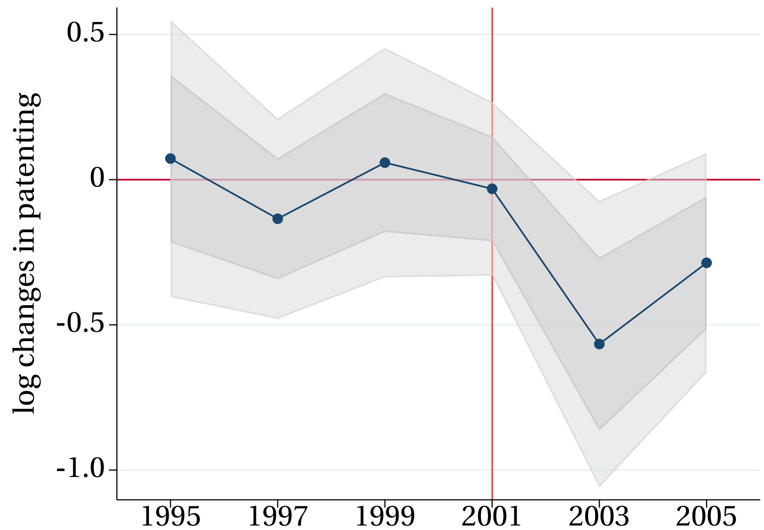
untreated mergers before the amendment, and significant differences appear only after the amendment. The effect is concentrated in the short run after the amendment, though. This is to be expected if deterrence is the main mechanism driving the results. After the Amendment the mergers that were deterred by the authorities are quickly consumed in few years, generating the large results observed in this paper. The effect might still be present afterwards, but the limited sample might not be sufficient to generate significant results.

Rather than inspecting the years around the Amendment, one can study the dynamics in years following every single merger. Regarding single period changes computed as $\Delta P_{t+i} = ln\left(P_{t+i}\right) - ln\left(P_{t-1}\right)$, one can see if the effect is stronger in the long or in the short run after the merger. Figure 16 in Appendix C shows coefficients of triple diff-in-diff for each year after the merger. It shows that the effect is negative and significant in the first 4 years after the merger. The last two coefficients imply that the effect is still negative 6 years after the merger, but it is weaker and less significant. Therefore one can conclude that the effect is more pronounced in the first years after the merger, although it is still negative in the long run.

5.1 Quality and Originality of Innovation

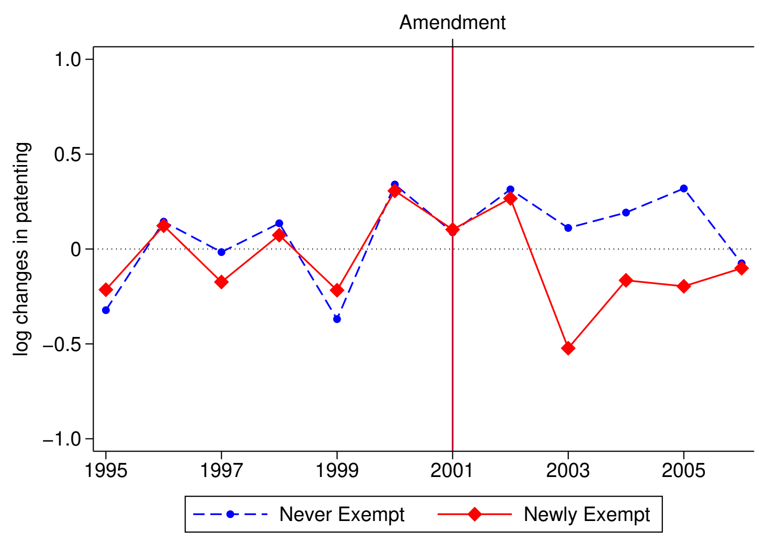
In order to ascertain whether patenting behavior is changing after horizontal mergers that are not notified to the authorities, I apply the triple diff-in-diff analysis to several patenting measures. Figure 9 shows that the main result on relative citations is driven by a decrease in citations rather than a decrease in the number of patents. This means that merging firms are still innovating, but this innovation is of lower quality, since patents receive less citations.

Figure 7: Coefficients of triple Diff in Diff for years around the Amendment.



Notes: Coefficients of triple diff-in-diff term $I^{Ex}I^{Hor}$ interacted with two year periods around the Amendment. The dependent variable of the regression ΔP is computed using the relative citation average described in Section 3. Appendix C reports the same figure for single years and three or four year periods. Results have the same qualitative implications.

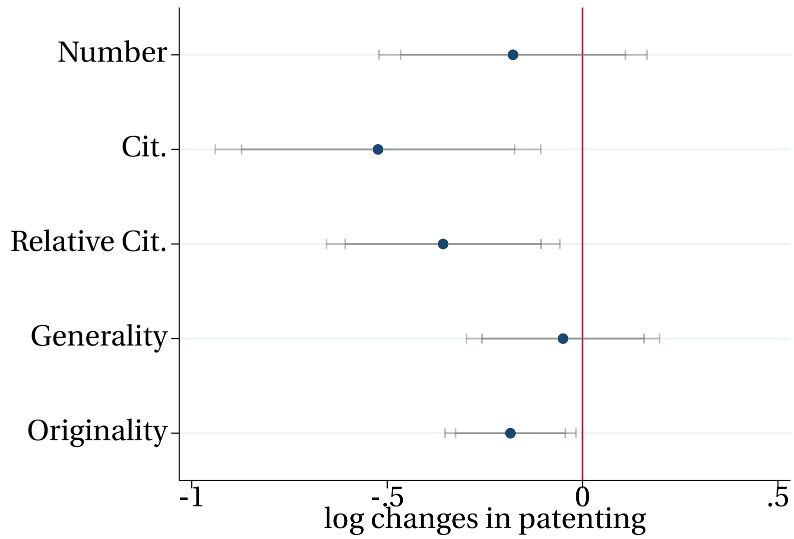
Figure 8: Difference in innovation between Horizontal and Non-Horizontal Mergers



Notes: The data points of the figure are constructed by averaging the innovation effect of mergers in various groups. The variable ΔP is computed using the relative citation average described in Section 3. Data are residualized on fixed effects used in column (1) of Table 2. Each point reports the difference between horizontal and non-horizontal mergers.

Innovation quality can be measured is several ways. Following Lerner et al. (2011) construct two measures using citations and technological fields. First, I compute Generality of a patent as the dispersion of citing patents among technological fields. This is calculated as $(1 - HHI_{citing})$ where

Figure 9: Triple difference in differences results for various innovation activity measures.



Notes: Coefficients of triple diff-in-diff equation 5 with various measures patent activity as dependent variable ΔP. Column (1) reports the total number of patents submitted each year. Column (2) the total number of citations received by patents submitted. Column (3) reports the main results computed with Relative Citation Average, which takes into account varying patenting activity in different technology spaces. Column (4) reports Generality, which increases if patents are cited by a diverse array of patents, as computed by $(1 - HHI)$ of citing patent technology spaces. Column (5) reports Originality, which is higher for patents citing a diverse array of patents, as computed by $(1 - HHI)$ of cited patent technology spaces.

HHI_{citing} is the Herfindahl–Hirschman index of citing patents among technological fields. If all citing patents are concentrated in the same field, then the $HHI_{citing} = 1$ and Generality is equal to 0. This measure captures the fact that the patent is speaking to various fields, and as such it is more general. Figure 9 shows that Generality is decreasing after the policy change, but the decrease is small and non significant. On the other hand, the decrease in Originality is larger and significant. I compute Originality as the dispersion of cited patents. Similarly to Generality, this is calculated as $(1 - HHI_{cited})$. Therefore, a patent with high Originality is citing other patents from a divers array of technological fields. My results show that horizontal mergers that are not notified to the authorities lead to a decrease in Originality of patents, and have little effect on Generality of patents.

5.2 Product or Process Innovation

The literature has identified two categories of innovation. Process innovations comprises new methods of production that increase firm's productivity. This can be modeled as cost reducing innovation, and it can be added to models of competition as in Motta and Tarantino (2017).[50] Product innovation, on the other hand, means updated products that respond to consumer preferences. This can be modeled as demand enhancing innovation, and Bourreau et al. (2021) include it in a model of competition. In order to identify which patents represent process or product innovation I follow the methodology of Ganglmair et al. (2022). Using text analysis they classify the individual claims of each patent as process or product claim. Claims describe the possible applications of a patent,

[50]The authors also show that for a large category of models cost-reducing and quality-enhancing innovation are equivalent.

Table 3: Triple Difference in Differences Results for Product and Process Innovation

VARIABLES	(1) Both	(2) Process	(3) Product
Relative Citation Average			
$I^{Post} \cdot I^{Ex} \cdot I^{Hor}$	-0.357**	-0.406*	-0.280*
	(0.152)	(0.214)	(0.163)
Implied Change	-30.0%	-33.4%	-24.4%
Citations			
$I^{Post} \cdot I^{Ex} \cdot I^{Hor}$	-0.523**	-0.499*	-0.358
	(0.212)	(0.301)	(0.229)
Implied Change	-40.7%	-39.3%	-30.1%
Observations	2,601	1,599	2,242
All FE	YES	YES	YES
Cluster SE	SIC4	SIC4	SIC4

Notes: This table reports coefficients of equation 5 applied on process or product innovation. Column (1) shows baseline results. Column (2) shows results computed considering only process patents. Column (3) reports results computed considering only product patents.

and as such they are the natural choice for this exercise. Then I classify each patent as a process or product patent based on the classification of its first claim. All other claims tend to refer to the first one, as it is usually the broadest one.[51]

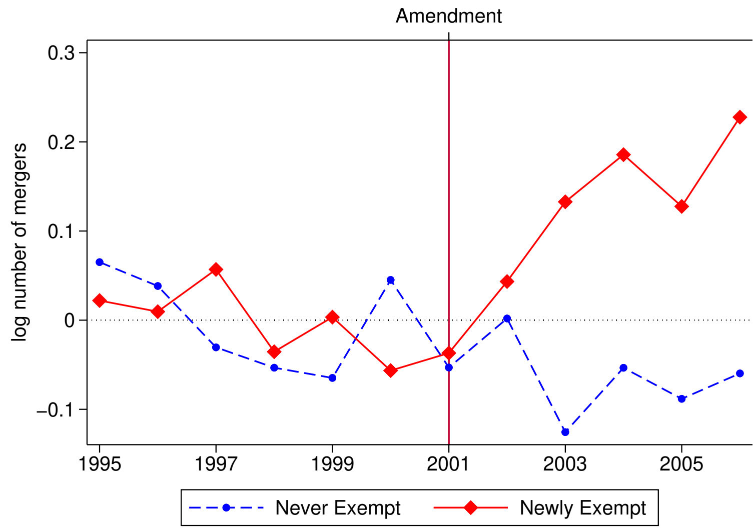
In order to assess which kind of innovation is most affected by the policy change, I repeat the main analysis described by equation 5 but considering only patents classified as one kind of innovation. Table 3 reports results of this analysis. Column (1) shows results on the whole sample, which is the baseline of this paper. Column (2) reports coefficients computed considering only process patents, while column (3) considers only product patents. For both relative citation average and total citations the effect is stronger and more significant for process innovation. Actually, the effect on process innovation is stronger than the overall effect on innovation, a decrease of 0.406 log points, compared to the 0.357 log points of the main result. This means that affected firms are becoming less productive with respect to the control group. This result justifies the choice of a model of competition with cost reducing innovation, where firms invest resources in innovation to increase their productivity.

5.3 Number of Mergers and Deterrence

In order to test the predictions of the deterrence model, I inspect the number of mergers before and after the amendment. Wollmann (2019) shows extensive evidence that after the amendment there is an increase in the number of horizontal mergers that are not reported to the authorities. The author

[51] In this patent classification based on the first claim I follow the methodology used by Ganglmair et al. (2022).

Figure 10: Time series of the number of mergers



Notes: The data points of the figure are constructed as the log number of mergers in various groups. Data are residualized on year fixed effects. Each point reports the difference between the number of horizontal and non-horizontal mergers.

defines this merger wave as stealth consolidation, a series of potentially anticompetitive mergers that escape antitrust scrutiny. In Figure 10 I report the number of newly exempt and never exempt mergers before and after the amendment. Before the policy change there is no particular difference between the two groups, as both are notified to the authorities. After the amendment newly exempt mergers are not notified anymore, and Figure 10 shows that they increase by about 0.2 log points or 20%. This agrees with the interpretation of deterrence as a plausible mechanism driving the results. Before the amendment several potentially anticompetitive mergers were deterred by the authorities, as they had a significant probability of being blocked. After the policy change the antitrust authority does not control them anymore, and so these mergers are successfully attempted by the merging parties. I report also the number of never exempt mergers, the ones that are large enough to be notified to the authorities both before and after the amendment, to show that the increase in the number of unreported mergers is not due to a general trend in merger dynamics.

Several other countries experienced similar policy changes to the one analyzed in this paper. In Table 8 in Appendix C I show that these policy changes resulted in significant decreases in the number of notifications received by each antitrust authority. In the case of Italy, for example, the number of notifications fell by as much as 90%. In another paper, Morzenti (2022), I analyze the effect of these policy changes. I find evidence of stealth consolidation in all countries included in the study, meaning that the number of horizontal mergers that are not notified to the authorities increases. This is further evidence of the deterrence effect of antitrust authorities even in countries outside the US. Moreover, I find that these policy changes generate an increase in concentration in affected industries, a decrease of labor share by 2% and a decrease in investment by 4%, on average. This shows that stealth consolidation can have far reaching effects on the whole economy, and not only on innovation.

6 Discussion and Robustness

A feature of the analysis that is worth emphasizing is the size of the sample. Table 10 in Appendix C shows the number of treated and untreated mergers, both before and after the amendment. The overall sample size is of 2601 merging firms, which is already a small number. This is due to the nature of the analysis. I consider only merging firms whose transaction involves companies that are actively patenting before the merger, so as to compute patent similarity between the merging parties. This exclude a great deal of mergers from the sample, as not so many firms were actively patenting and chose to merge in the time span between 1995 and 2006. Considering the small size of the sample it is remarkable that identified effects are statistically significant. This is a further sign of the magnitude and the economic importance of the results identified in this paper. This might explain also why results appear to be significant only in the short run after the policy change. In few years after the amendment many of the potentially anticompetitive mergers that were deterred by the authorities were attempted successfully, resulting in an immediate effect strong enough to be seen in the data. The effect might still be there in the long run, but the power of such a small sample is not sufficient to identify it.

Given the definition of horizontal mergers in Section 4.1, the number of non-horizontal mergers is higher than the number of non-horizontal ones. Moreover, the number of never exempt horizontal mergers is much lower than the number of newly exempt ones. This is due to two reasons. First, most of the mergers are below the new threshold introduced by the amendment.[52] Second, it is harder to identify never exempt mergers, as they need to satisfy more conditions simultaneously. To be exempt, on the other hand, mergers have to satisfy just one of the conditions provided by the Hart-Secott-Rodino Act.[53] In panel B and C of Table 10 one can see the group sizes before and after the amendment. The smallest group is never exempt horizontal mergers after the amendment, and it counts 124 mergers. Even if small, this number is well above the standard boundaries for statistical meaningful results, and indeed results of this paper are significant. However, the size of the sample limits the kind of exercise I can perform with it, as slicing it further leads to sub-samples that are too small to yield any meaningful insight.

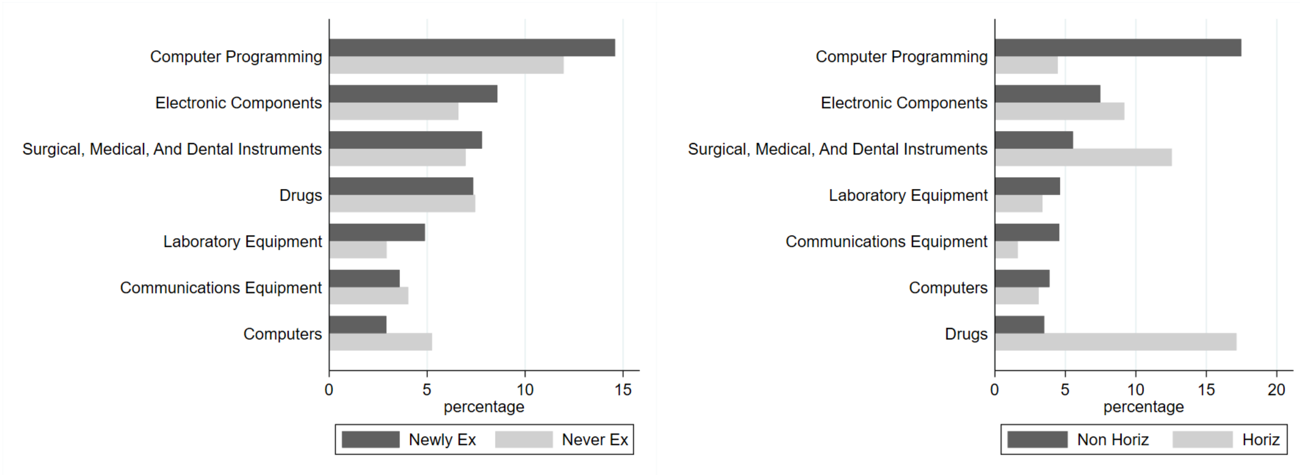
6.1 Sectors and Industries

Given the nature of my analysis I include in the sample all sectors available in the data. Consequently the composition of sectors in treated and untreated groups might be affecting my results. Figure 11 shows the main sectors that are represented in the analysis. Two are the main industries that compose the sample. The first is Big Tech, which includes "computer programming", "electronic

[52] One can see this also in the Figure 10, where it is clear that there are many more mergers below \$50 million than above this threshold.

[53] The conditions are summarized in the "size of person test" and the "size of transaction test" described in Section 4.2

Figure 11: Sector composition differences between merger categories



(a) Newly vs. Never Exempt

(b) Horizontal vs. Non-Horizontal

Notes: The graphs report the distribution of mergers in various industries. Panel (a) on the left shows the difference between the distribution of Newly and Never Exempt mergers. Panel (b) on the right shows the difference between the distribution of horizontal and non-horizontal mergers. Only industries with highest share of mergers are reported in the graphs.

components" and "communication equipment". The second is the Pharma industry, which comprises "drugs", "Surgical, medical and dental equipment" and "laboratory equipment". Figure 11a shows that the sector distribution for Newly and Never-Exempt mergers is quite similar. Conversely, Figure 11b implies that horizontal mergers are more common in the Pharma industry, whereas non-horizontal ones are more common in the Big Tech industry. This means that merging firms in the Pharma industry tend to have more similar patent portfolios.[54]

As a consequence of this sector differences one might be concerned that the main results of this paper are driven by a different sector composition of treated and untreated mergers. This is accounted for by the triple difference in difference nature of the identification strategy. The treated horizontal mergers that are not notified by the authorities are actually compared with untreated horizontal mergers that are not notified to the authorities. These two groups have no meaningful composition difference. As a proof of this I conduct a leave-one-out exercise in which I exclude one sector at a time from the sample and test whether the results are still significant. A sector by sector analysis could not be implemented given the limited sample size. Table 4 shows results of the leave-one-out exercise, and it is clear that results are not affected by removing sectors. Column (4) and (5) show that If one removes sector in the Pharma industry such as "Drugs" or "Medical, Surgical and Dental Equipment" results are even stronger, a decrease in innovation of bout 0.4 log points. Overall, one can interpret these results as evidence that sector composition is not driving the effects found in this paper.

Even if the number of mergers in particular sectors is limited, Table 11 in Appendix C shows re-

[54]This might be due to more homogeneous patents in the Pharma industry, compared to other industries.

Table 4: Triple difference in differences leaving out particular sectors

VARIABLES	(1) Baseline	(2) Computers	(3) Software	(4) Drugs	(5) Med. Eq.
$I^{Post} \cdot I^{Ex} \cdot I^{Hor}$	-0.357**	-0.358**	-0.352**	-0.420**	-0.397**
	(0.152)	(0.158)	(0.177)	(0.165)	(0.158)
Implied Change	-30.0%	-30.1%	-29.7%	-34.3%	-32.8%
Observations	2,601	2,506	2,243	2,409	2,405
R-squared	0.080	0.080	0.092	0.086	0.083
Year FE	YES	YES	YES	YES	YES
Industry FE	YES	YES	YES	YES	YES
State FE	YES	YES	YES	YES	YES
Cluster SE	SIC4	SIC4	SIC4	SIC4	SIC4

Notes: This table reports coefficients of equation 5 with various sample specifications. Column (1) reports the baseline result computed on the whole sample. Column (2) reports results computed on the sample without "Computers", column (3) without "Computer Programming", column (4) without "Drugs", column (5) without "Medical, Surgical and Dental Equipment".

sults of the triple diff-in-diff methodology applied to single sectors. The only two sectors that display significant coefficients are "software programming" and "drugs". It is worth mentioning that these are also among the most numerous sectors in the sample, and this contributes to the statistical significance of their coefficients. The very high coefficient displayed by the "software programming" can be due to the fact that non-horizontal mergers in this sector tend to have more positive effects on innovation, as shown in Figure 12 in Appendix C. This could be explained by the high level of complementarities that can be attained in this particular sector.

6.2 Macro Effect

The main results of this paper pertain only to merging firms. These firms are small and private, and thus it is natural to ask whether the overall level of innovation was affected. Indeed Motta and Tarantino (2017) study the innovation effect of a merger on the rivals of merging firms. They show that if the merging firms decrease innovation, then the rivals increase innovation, but not enough to compensate the main effect on the insiders. As a consequence, the total level of innovation decreases. Haucap et al. (2019) show a similar result empirically for mergers in the EU.

To study the effect of a relaxation of pre-merger notification rules on the overall level of innovation, I propose an industry level analysis. Equation (6) describes the difference in differences identification strategy. The unit of analysis is an industry in a year. The variable I^{Post} is a dummy equal to 1 for year after 2001. The variable $Treat$ is computed as the number of newly exempt horizontal mergers divided by the total number of mergers. This can be considered as a continuous measure of the intensity of stealth consolidation in a given industry. δ_i and α_t represent industry and year fixed effects. I inspect several outcome variables Y_{it}, so as to study also the channels that bring to the effect

Table 5: Results of Industry Level analysis

VARIABLES	COMPUSTAT HHI (10,000)	COMPUSTAT Profit (100%)	EU Klems Investment (Bn $)	EU Klems R&D (Bn $)
$I^{Post} \cdot Treat$	605.5**	5.215***	-887.7***	-21.41***
	(258.4)	(1.947)	(213.3)	(8.231)
Observations	483	483	357	357
N. industries	23	23	17	17
R-squared	0.871	0.758	0.985	0.961
Avg	784.1	10.1	1055.7	17.85
Year FE	YES	YES	YES	YES
Industry FE	YES	YES	YES	YES

Notes: Coefficients of diff-in-diff equation 6 with different outcome variables Y_{it}. In COMPUSTAT the concentration HHI is computed using turnover, while profit margins are computed as EBIT+FINREV+FINX divided by turnover. In EU Klems I use the stock of real investment (Kq_OCon) and real R&D flow (Iq_RD). Industries are defined in COMPUSTAT using the SIC 2 digit level, while in EU Klems industries are defined using the Klems classification (See more info on the EU Kelms website).

on innovation.

$$Y_{it} = \beta I^{Post} \cdot Treat + \gamma Treat + \delta_i + \alpha_t + \xi X_{it} + \epsilon_{it} \tag{6}$$

Table 5 reports results of the identification described in equation (6). Using data from COMPUSTAT I compute industry level HHI and profit margin. The first two columns show that both of them are increasing. The coefficient in the second column implies that profit margins would increase by 5.2% after the policy change if all mergers in the data became part of stealth consolidation. This implies that industries are becoming more concentrated, and thanks to more market power, firms are able to increase their profits. Then, using EU Klems data one can compute investment stock and RD expenditure. The third and fourth column of Table 5 show that both investment and R&D decrease after the policy change. This implies that the overall effect on innovation of a relaxation of pre-merger notification rules is negative.

6.3 Definition of Horizontal Mergers

The definition of horizontal mergers presented in Section 4.1 is a key part of the identification strategy. One might be concerned that the particular choice of patent similarity measure between merging firms can drive the final results. Table 12 to 15 in Appendix C show coefficient of the main equation 5 where the dummy I^{Hor} is defined using different similarity statistics. Overall the qualitative results do no change, and in some cases the coefficients are even larger than the main results shown in Table 2.

Table 6: Triple difference in differences results computed using continuous patent similarity.

VARIABLES	(1) Max	(2) Max 20	(3) Max 1%	(4) Max 2%	(5) Max 5%
Relative Citation Average					
$I^{Post} \cdot I^{Ex} \cdot Sim$	-1.054***	-1.099***	-0.848**	-0.946**	-0.935*
	(0.397)	(0.405)	(0.406)	(0.460)	(0.489)
$I^{Post} \cdot I^{Ex}$	0.818***	0.689***	0.582**	0.640**	0.629**
	(0.283)	(0.243)	(0.239)	(0.265)	(0.271)
$I^{Post} \cdot Sim$	0.566*	0.401	0.330	0.452	0.476
	(0.326)	(0.326)	(0.332)	(0.389)	(0.414)
R-squared	0.081	0.081	0.081	0.080	0.080
Citation Count					
$I^{Post} \cdot I^{Ex} \cdot Sim$	-1.512***	-1.410***	-1.294***	-1.157**	-1.400**
	(0.462)	(0.510)	(0.486)	(0.554)	(0.644)
$I^{Post} \cdot I^{Ex}$	1.170***	0.912***	0.882***	0.843**	0.945**
	(0.353)	(0.321)	(0.305)	(0.334)	(0.368)
$I^{Post} \cdot Sim$	0.508	0.210	0.331	0.319	0.436
	(0.440)	(0.452)	(0.406)	(0.473)	(0.554)
R-squared	0.158	0.158	0.157	0.156	0.156
Observations	2,610	2,610	2,610	2,610	2,610
Year FE	YES	YES	YES	YES	YES
Industry FE	YES	YES	YES	YES	YES
State FE	YES	YES	YES	YES	YES
Cluster SE	SIC4	SIC4	SIC4	SIC4	SIC4

Notes: Coefficients of triple diff-in-diff equation 7 with different continuous similarity statistics used to compute the variable Sim. The first panel reports results computed using Relative Citation Average as patenting measure P_{it}. The second panel shows results computed using citation count as patenting measure P_{it}.

$$\begin{aligned}\Delta P_{it} =& \beta I^{Post} I^{Ex} Sim + \gamma I^{Post} I^{Ex} + \theta I^{Post} Sim + \delta_1 I^{Ex} Sim + \delta_2 I^{Ex} + \delta_3 Sim \\ &+ \alpha_t + \xi X_{it} + \epsilon_{it}\end{aligned} \tag{7}$$

The main exercise of this paper relies on the identification of horizontal mergers using a 0-1 discrete rule. This reflect the inner workings of antitrust authorities that spend time and resources in classifying notified mergers. However, it is possible to use patent similarity as a continuous measure to identify the effect that the amendment has on firms with similar patent portfolios. This can be considered a continuous measure of "horizontality" of the merger. Equation 7 shows that similarity measures Sim can be put in place of the horizontal dummy I^{Hor} for an alternative identification strategy. Table 6 shows results of this identification strategy. The panel of relative citation average shows an average decrease in innovation of about 0.9 log points, which means a decrease of about 60%. The magnitude of coefficients of this table has a different interpretation with respect to the

Table 7: Triple diff-in-diff result computed for Acquirers and Targets

	Relative Cit. Average			Citations		
VARIABLES	Baseline	Target	Acquirer	Baseline	Target	Acquirer
$I^{Post} \cdot I^{Ex} \cdot I^{Hor}$	-0.357**	-0.246	-0.401**	-0.523**	-0.204	-0.654**
	(0.152)	(0.293)	(0.175)	(0.212)	(0.418)	(0.325)
Implied Change	-30.0%	-21.8%	-33.0%	-40.7%	-18.4%	-48.0%
Observations	2,601	1,015	1,586	2,610	1,021	1,589
R-squared	0.080	0.145	0.144	0.156	0.196	0.209
All FE	YES	YES	YES	YES	YES	YES
Cluster SE	SIC4	SIC4	SIC4	SIC4	SIC4	SIC4

Notes: Coefficients of triple diff-in-diff equation 7 focusing only on firms making the acquisition (Acquirer) or firms being acquired (Target). The first three columns report results computed using relative citation average, the measure used in the baseline result in Table 2. The other columns report results computed using total number of citations.

main results: an increase in similarity of a merger from 0 to 1 (from completely orthogonal to identical patent portfolios) implies a 60% lower innovation outcome for unreported mergers. Thus, one can conclude that the more similar are two merging companies, the more they are affected by the amendment.

6.4 Acquirer and Target

Firms that acquire a competitor and firms that are acquired might enter the merger for different reasons. This may shape their incentives and their behavior. Therefore, Table 7 proposes the main identification strategy applied only on acquirers or on targets of the acquisition. The point estimate is negative for both, but it is not statistically significant for target firms. The coefficient or acquirers, on the other hand, is high and significant. This means that acquirers involved in horizontal mergers that are not notified to the authorities decrease their innovation effort after the merger. The effect is even stronger than the baseline.

This result relates to the growing literature on killer acquisitions spurred by Cunningham et al. (2019). The authors show that in some instances the target firm gets shut down after the merger. My data does not allow to see whether a firm cease operations or defaults. However, the results of Table 7 imply that the strongest decline in innovation happens on the side of the acquirer. This is in line with the recent literature on reverse killer acquisitions. Caffarra et al. (2020) propose that the acquirer might terminate its own innovation projects as a consequence of the merger. My results show that this is the prevailing channel explaining the drop in innovation after the relaxation of pre-merger notification rules.

6.5 Patents as Measure of Innovation

Patents are a in imperfect measure of innovation. A strategic choice determines whether a firm will patent or not a new discovery.[55] Patenting ensures protection only for a limited amount of time, and there are large economies in which intellectual property rights are not safeguarded as in the US. As a consequence, published patents provide only a partial measure of the innovation activity of a company. The triple difference in differences identification strategy explained in Section 4 can account for this, though. Strategic patenting can affect the main results only if firms involved in horizontal mergers that are not notified to the authorities have different strategic motives than firms involved in non-horizontal ones or notified ones. Moreover, Kuhn et al. (2020) show that the nature of patent citations has changed dramatically in recent years, with few patents receiving most of citations. The findings of Kuhn et al. (2020) apply mostly to the years after 2005, and so they do not represent a concern for the analysis of this paper, which spans from 1995 to 2005. Lastly, Table 5 shows that also R&D expenditure and the stock of investment decline after the policy change.

6.6 Remedies

An alternative mechanism that might explain the results on innovation found in this paper is that remedies imposed by the authorities improve the innovation outcome after notified mergers. If the antitrust authorities were able to negate the anticompetitive effects of mergers through remedies, in a model of competition with cost reducing innovation then mergers controlled by the authorities would have a positive effect on innovation.[56] After the amendment, only mergers that are notified to the authorities benefit from remedies, and as a consequence non notified mergers appear to lead to less innovation.

Remedies, however, are known to have limited effects on competition outcomes such as prices. In his recent work Kwoka (2014) reports merger retrospective studies done on 119 product prices. The author reports that remedies have become more and more important as the authorities are looking for alternatives to binary decisions such as blocking. This despite strong evidence that mergers resulted in higher prices, regardless of whether the agencies imposed remedies or not, and of the type of remedies chosen. Kwoka (2014) finds that mergers on average lead to a 7.22% increase in prices, whereas mergers in which the authorities impose remedies implied a price increase of 7.71%. Mergers with conduct remedies resulted in an increase of 16.03%, which is particularly troubling because conduct remedies are becoming more common, whereas structural remedies and divestitures are pursued less.

6.7 Amendment

The year 2000 was marked by the dot-com bubble, a financial crisis that influenced the entire US economy. The triple difference-in-differences identifications strategy should control for this, as both

[55] Among others, Righi and Simcoe (2020) explore strategic motives behind patenting decisions related to future patent applications.

[56] This would be due to merger efficiencies. If a merger has zero or positive effect on consumer surplus, then this merger has positive effect on innovation, as shown in Figure 1.

treated and control mergers were affected by the dot-com bubble. As a further test, Figure 17 in Appendix C shows a placebo exercise using the 2008 financial crisis as a placebo event. The Figure shows that there is no discernible effect of the financial crisis on newly or never exempt mergers.

Moreover, I test whether the choice of 2001 as year of the amendment affects significantly the results. As the policy change was voted in December 2000 and it became effective the next year, one could consider also 2000 as the year of amendment. Table 16 in Appendix C shows that the effect is still negative, significant, and of similar size. A further concern might be that never exempt mergers are too different from newly exempt ones because they are much larger. In Table 18 in Appendix C I shows results of the main analysis conducted on mergers that have a transaction size below $\$500$ million, so as to exclude the largest mergers. The effect is still present and negative, however it is smaller in magnitude and it is not significant.

Lastly, a higher "size of transaction" threshold of $\$200$ million is considered when merging parties do not satisfy the "size of person test", which is likely to be the case only for firms with a lot of intangible assets, as the ones that are considered in this paper.[57] Following the existing literature on the amendment I use the $\$50$ million threshold for the main analysis. However, my results hold even if I consider the $\$200$ million threshold to define newly and never exempt mergers. Table 17 in Appendix C shows that results are unaffected by this change, the coefficient is still negative and significant and of similar size to the main results.

7 Conclusion

This paper studies the effect of competition on innovation. In particular, I explore whether loosening antitrust policy discourages innovation of merging firms. To examine this issue, I take advantage of a natural experiment arising from a unique relaxation of pre-merger notification rules that affected hundreds of M&A. This amendment to existing regulation was so dramatic that the number of pre-merger notifications fell by as much as 70%. This event has allowed me to exogenously shift antitrust policy control for a subset of mergers. Given the nature of these transactions, I use a data-set containing mergers reported in news outlets and industry journals. Moreover, to measure the innovation activity of these firms I combine this data with the universe of patents published in the US. I focus on horizontal mergers, transactions that are the most likely to be anticompetitive and to attract the attention of the antitrust authorities. To identify horizontal mergers in this unconventional data-set, I use a natural language processing algorithm. I train a word embedding machine learning model on the whole corpus of US patents. I show that my algorithm performs better than standard industry classification at predicting EC and FTC classification of horizontal mergers.

In this sample, mergers lead to an average innovation reduction of about 28%. After the relaxation of notification rules, the difference in differences exercise indicates that non reported horizontal mergers imply a further 30% drop in innovation. Moreover, merging firms become less productive

[57] See Wollmann (2019) for more details on this second threshold. Moreover, also Cunningham et al. (2019) use this $\$200$ million threshold in their analysis in lieu of the $\$50$ million one.

after the policy change. Lastly, I find that the number of unreported horizontal mergers increases, in accordance with deterrence being the mechanism behind my results. To explore this avenue, I build an approachable theoretical framework. In this model, mergers that are more detrimental to consumers are also mergers that lead to decreasing innovation. At the same time, they are most susceptible to deterrence. This model predicts that after the policy change there should be an increasing number of horizontal mergers, and that these mergers lead to less innovation. Both these testable implications are verified by the main results of this work. Through the model, I infer that these mergers are harmful to the consumer surplus.

The policy implications of these results are that relaxing antitrust policy could have far reaching and unintended consequences. In particular, a surge of unreported mergers can stifle innovation. The policymaker should not dismiss small mergers as negligible for competition and innovation either. Quite to the contrary, a large number of these transactions may have a profound impact in several product markets. The FTC already cited stealth consolidation when it issued special orders compelling big tech to disclose previously non-reportable deals.[58] Moreover, the New York Senate passed a bill creating a first-of-its-kind $9.2 million state-specific pre-merger notification threshold, specifically aimed at the big tech sector.[59] The current paper, however, is focused only on one side of the coin. A general equilibrium analysis would consider also the potential benefits of lowering antitrust scrutiny. After a relaxation of notification rules, enforcement cost are expected to decrease, while profits should rise. Thus, this paper cannot be considered as evidence to support a strengthening of notification rules. This would require an analysis of both costs and benefits.

This work can be the foundation for several future research avenues. Using the measure of firm similarity based on text analysis of patents, it is possible to create a network of competitive relations and demand elasticity such as the one provided by Hoberg and Phillips (2016) and used by Pellegrino (2021) to study the implications of rising concentration and market power. The advantage of a patent-based measure would be to encompass also small and private firms, and not only large and public ones. This would make it possible to extend the model of deterrence to become a fully fledged structural model. Gathering more data on the actual pricing decisions of merging firms affected by the amendment would provide empirical evidence on consumer surplus and overall welfare.

[58] "We support the Commission's decision to issue a 6(b) study designed to assess the sufficiency of the Hart-Scott-Rodino Antitrust Improvement Act of 1976 ("HSR Act") thresholds with respect to *technology mergers* and acquisitions of competitive significance." As cited from the Joint Statement by the FTC Commissioners, 2020.

[59] "The Bill applies to all industries. But... concerns about purported anticompetitive behavior in the "Big Tech" sector were the spark." As cited from the White & Case summary, 2021

References

Acemoglu, Daron and Ufuk Akcigit, 2012, "Intellectual Property Rights, Competition and Innovation," *Journal of the European Economic Association*, *10* (1).

Affeldt, Pauline, Tomaso Duso, and Florian Szücs, 2021, "25 years of European merger control," *International Journal of Industrial Organization*, *76.*

Aghion, Philippe, Nick Bloom, Richard Blundell, Rachel Griffith, and Peter Howitt, 2005, "Competition and Innovation: an Inverted-U Relationship," *Quarterly Journal of Economics.*

Arrow, Kenneth, 1962, "Economic Welfare and the Allocation of Resources for Invention," in "The Rate and Direction of Inventive Activity: Economic and Social Factors," Princeton University Press, pp. 609–626.

Ash, Elliott, Daniel L. Chen, and Suresh Naidu, 2022, "Ideas Have Consequences: The Impact of Law and Economics on American Justice."

Baker, Jonathan B., 1988, "Private Information and the Deterrent Effect of Antitrust Damage Remedies," *Journal of Law, Economics & Organization*, *4* (2).

Barrios, John M and Thomas G Wollmann, 2022, "A New Era of Midnight Mergers: Antitrust Risk and Investor Disclosures."

Besanko, David and Daniel F. Spulber, 1989, "Antitrust Enforcement Under Asymmetric Information," *The Economic Journal*, *99* (396).

Besley, Timothy, Nicola Fontana, and Nicola Limodio, 2021, "Antitrust Policies and Profitability in Nontradable Sectors," *American Economic Review: Insights*, *3* (2).

Bourreau, Marc, Bruno Jullien, and Yassine Lefouili, 2021, "Mergers and Demand-Enhancing Innovation."

Breit, William and Kenneth G. Elzinga, 1973, "Antitrust Penalties and Attitudes toward Risk: An Economic Analysis," *Harvard Law Review*, *86* (4). Publisher: The Harvard Law Review Association.

Caffarra, Cristina, Greg Crawford, and Tommaso Valletti, ""How Tech Rolls": Potential Competition and "Reverse" Killer Acquisitions," 2020.

Crandall, Robert W. and Clifford Winston, 2003, "Does Antitrust Policy Improve Consumer Welfare? Assessing the Evidence," *Journal of Economic Perspectives*, *17* (4).

Cunningham, Colleen, Florian Ederer, and Song Ma, 2019, "Killer Acquisitions."

Decarolis, Francesco and Cristina Giorgiantonio, 2022, "Corruption red flags in public procurement: new evidence from Italian calls for tenders," *EPJ Data Science*, *11* (1).

Denicolò, Vincenzo and Michele Polo, 2021, "Acquisitions, innovation and the entrenchment of monopoly."

Dickens, David S., Rafal Kozielski, Javed Khan, Anne Forus, and Timothy P. Cripe, 2002, "Cyclooxygenase-2 Expression in Pediatric Sarcomas," *Pediatric and Developmental Pathology*, *5* (4).

Eckbo, B. Espen, 1992, "Mergers and the Value of Antitrust Deterrence," *The Journal of Finance*, *47* (3).

Federico, Giulio, Gregor Langus, and Tommaso Valletti, 2018, "Horizontal mergers and product innovation," *International Journal of Industrial Organization*, *61.*

Fumagalli, Chiara, Massimo Motta, and Emanuele Tarantino, 2020, "Shelving or developing? The acquisition of potential competitors under financial constraints."

Ganglmair, Bernhard, W. Keith Robinson, and Michael Seeligson, 2022, "The Rise of Process Claims: Evidence from a Century of U.S. Patents."

Gutiérrez, Germán and Thomas Philippon, 2017, "Declining Competition and Investment in the U.S.," Technical Report w23583, National Bureau of Economic Research, Cambridge, MA.

Haucap, Justus, Alexander Rasch, and Joel Stiebale, 2019, "How mergers affect innovation: Theory and evidence," *International Journal of Industrial Organization*, *63.*

Hoberg, Gerard and Gordon Phillips, 2016, "Text-Based Network Industries and Endogenous Product Differentiation," *Journal of Political Economy*, *124* (5).

Iaria, Alessandro, Carlo Schwarz, and Fabian Waldinger, 2018, "Frontier Knowledge and Scientific Production: Evidence from the Collapse of International Science*," *The Quarterly Journal of Economics*, *133* (2).

Jullien, Bruno and Yassine Lefouili, 2018, "Horizontal Mergers and Innovation."

Kuhn, Jeffrey, Kenneth Younge, and Alan Marco, 2020, "Patent citations reexamined," *The RAND Journal of Economics*, *51* (1).

Kwoka, John, 2014, *Mergers, Merger Control, and Remedies*, The MIT Press.

Lande, Robert H. and Joshua P. Davis, 2011, "Comparative Deterrence from Private Enforcement and Criminal Enforcement of the U.S. Antitrust Laws," *Brigham Young University Law Review*, *2011* (2).

Lerner, Josh, Morten Sorensen, and Per Strömberg, 2011, "Private Equity and Long-Run Investment: The Case of Innovation," *The Journal of Finance*, *66* (2).

Mermelstein, Ben, Bates White, Volker Nocke, Mark A Satterthwaite, and Michael D Whinston, 2018, "Internal versus External Growth in Industries with Scale Economies: A Computational Model of Optimal Merger Policy."

Mikolov, Tomas, Ilya Sutskever, Kai Chen, Greg Corrado, and Jeffrey Dean, 2013, "Distributed Representations of Words and Phrases and their Compositionality," *arXiv:1310.4546 [cs, stat]*. arXiv: 1310.4546.

_, Kai Chen, Greg Corrado, and Jeffrey Dean, 2013, "Efficient Estimation of Word Representations in Vector Space."

Miller, Nathan H., 2009, "Strategic Leniency and Cartel Enforcement," *The American Economic Review*, *99* (3). Publisher: American Economic Association.

Morzenti, Giovanni, 2022, "A Cross Country Analysis of Stealth Consolidation and its effects on Inequality."

Motta, Massimo and Emanuele Tarantino, 2017, "The Effect of Horizontal Mergers, When Firms Compete in Prices and Investments."

Pellegrino, Bruno, 2021, "Product Differentiation and Oligopoly: a Network Approach."

Ridder, Maarten De, 2020, "Market Power and Innovation in the Intangible Economy."

Righi, Cesare and Timothy Simcoe, 2020, "Patenting Inventions or Inventing Patents? Continuation Practice at the USPTO," Technical Report w27686, National Bureau of Economic Research, Cambridge, MA.

Schumpeter, Joseph A., 1934, *The Theory of Economic Development: An Inquiry into Profits, Capital, Credit, Interest, and the Business Cycle.*

Wils, Wouter P. J., 2006, "Optimal Antitrust Fines; Theory and Practice," *World Competition*, *29* (2).

Wollmann, Thomas G., 2019, "Stealth Consolidation: Evidence from an Amendment to the Hart-Scott-Rodino Act," *American Economic Review: Insights*, *1* (1).

—, 2021, "How to Get Away with Merger: Stealth Consolidation and Its Effects on US Healthcare," Technical Report w27274, National Bureau of Economic Research, Cambridge, MA.

Younge, Kenneth A. and Jeffrey M. Kuhn, 2015, "Patent-to-Patent Similarity: A Vector Space Model," *SSRN Electronic Journal.*

Zeidler, Reinhard, Miklos Csanady, Olivier Gires, Stephan Lang, Bärbel Schmitt, and Barbara Wollenberg, 2000, "Tumor cell-derived prostaglandin E2 inhibits monocyte function by interfering with CCR5 and Mac-1," *The FASEB Journal*, *14* (5).

APPENDIX

A Model

A.1 Closed form solutions of the simple model

Figure 2 reports the changes in innovation and consumer surplus depending on the level of efficiencies λ generated by a merger. All formulas are reported without firm identifier because firms are symmetric. Therefore, the valuse are the same for firm i and firm j. The following equations report the closed form solutions for these changes due to a merger:

$$\Delta x = x_M - x_b = \frac{\alpha - c}{1 + 2\gamma(1-\lambda) - 2\lambda} - \frac{\alpha - c}{1+\gamma}$$

$$\Delta CS = CS_M - CS_b = \frac{(\alpha - c)^3(1+2\gamma)(1-\lambda)^3}{6(1+2\gamma(1-\lambda)-2\lambda)^3} - \frac{(\alpha-c)^3(1+2\gamma)}{6(1+\gamma)^3}$$

Moreover, the following equations show changes in quantity q, price p and profit π due to the merger:

$$\Delta q = q_M - q_b = \frac{(\alpha - c)(1-\lambda)}{1+2\gamma(1-\lambda)-2\lambda} - \frac{\alpha-c}{1+\gamma} = x_M(1-\lambda) - x_b$$

$$\Delta p = p_M - p_b = \frac{1}{2}\left(\alpha + c + \frac{\alpha - c}{1+2\gamma(1-\lambda)-2\lambda}\right) - c$$

$$\Delta \pi = \pi_M - \pi_b = \frac{(\alpha-c)^2(1-\lambda)}{1+2\gamma(1-\lambda)-2\lambda} - \frac{(\alpha-c)^2}{2(1+\gamma)^2} > 0$$

Note that the change in profit due to a merger is always positive in this model. The change in other variables depends on the level of efficiencies λ.

A.2 More General Framework

To explain the deterrence effect of Antitrust Authorities I build a model with endogenous merger choice and an active antitrust policy. Mergers that are more detrimental to consumers have a lower chance of being proposed to the authorities, since they have a lower chance of being accepted. This generates deterrence of the most anticompetitive mergers. Merger decisions affect not only competitive outcomes such as prices, but also the innovation incentives of the merging parties. Therefore, I describe firm behavior with a model of competition in prices and cost reducing innovation à la Motta and Tarantino (2017). Mergers that lead to lower consumer surplus, and thus are more susceptible to deterrence, are also mergers that generate less efficiencies and less innovation. This is the mechanism underlying the main results.

There are two kinds of agents in the model. Two firms have a merging opportunity and they maximize expected profit π. They have an imperfect ability to arrange the merger, as managers exert an effort to convince their shareholders and to negotiate the merger conditions. This results in a probability of proposing the merger $\varphi \in [0, 1]$. The Antitrust Authority maximizes expected consumer surplus CS. The authority has imperfect information on the consumer surplus effect of the merger, and it blocks any merger that exceed a determined level of harm. This results in a probability of allowing the merger $\alpha \in [0, 1]$. This is effectively the antitrust policy, which is known to the merging parties.[60]

At time $t = 0$ the Antitrust Authority chooses its policy rule, which relates each possible merger to the probability α^* that it is allowed. Knowing this rule, at time $t = 1$ firms decide their probability to propose the merger $\varphi^*(\alpha^*)$. Lastly, at time $t = 2$ firm merge with probability $\alpha^*\varphi^*(\alpha^*)$ and they compete obtaining profits π and generating consumer surplus CS. I start to solve the model from period $t = 2$ and then move backwards toward period $t = 0$.

A.2.1 Competition (t=2)

Firms compete in prices p_i and cost reducing innovation x_i. In order to innovate firms have to pay a fixed cost $F(x_i)$. This kind of innovation is most akin to process innovation, which makes firms more efficient, in contrast with product innovation, which creates new products and enhances demand.[61]. Before the merger, or in case the merger does not realize, each merging firm gains profit π_b as in equation **??**.

Following Motta and Tarantino (2017), firms compete in prices p_i and cost reducing innovation x_i. Before the merger thy maximize their own profits:

$$\pi_b = \max_{p_i, x_i} (p_i - c(x_i))q_i(p_i, \bar{p}_{-i}) - F(x_i)$$

Where marginal cost $c(x_i)$ satisfies $c(0) > 0,\ c' < 0,\ c'' \geq 0,\ c''' \geq 0$. Research fixed costs $F(x_i)$ satisfy $F(0) = 0,\ F' \geq 0,\ F'' \geq 0,\ F''' \geq 0$. The associated FOC are:

$$\begin{aligned}
\partial_{p_i}\pi_b &= q_i(p_i, \bar{p}_{-i}) + \partial_{p_i} q_i(p_i, \bar{p}_{-i})(p_i - c(x_i)) = 0 \\
\partial_{x_i}\pi_b &= -\, c'(x_i)q_i(p_i, \bar{p}_{-i}) - F'(x_i) = 0
\end{aligned}$$

This implies that for each value of $(p_i, \bar{p}_{-i})$ there is a unique value of x_i, pinned down by the following condition:

$$-\frac{F'(x_i)}{c'(x_i)} = q_i(p_i, \bar{p}_{-i})$$

Two companies merging generate efficiencies $\lambda G(x_i, x_k)$ satisfying $F(x_i) + F(x_k) - \lambda G(x_i, x_k) \geq 0$. After the merger they maximize the profits of the merged entity:

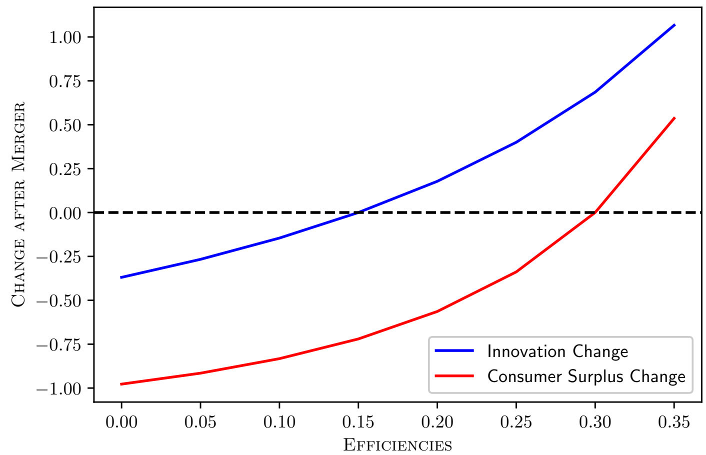
$$\begin{aligned}
2\pi_M = \max_{p_i, x_i, p_k, x_k} & (p_i - c(x_i))q_i(p_i, \bar{p}_{-i}) + (p_k - c(x_k))q_k(p_k, \bar{p}_{-k}) \\
& - F(x_i) - F(x_k) + \lambda G(x_i, x_k)
\end{aligned}$$

A general result of these models is that higher efficiencies λ imply better outcomes, both in terms

[60] It is reasonable to assume that merging parties have this knowledge. They always employ consultancy firms to organize the merging process, and consultants have good knowledge of the merger review process. In practice, a lot of policy work is done by consultants who discourage merging firms from proposing their transactions when they know that it would not stand in court.

[61] For a model of competition with demand enhancing innovation the reader can refer to Jullien and Lefouili (2018)

Figure 1: Changes in Consumer Surplus and Innovation after a merger



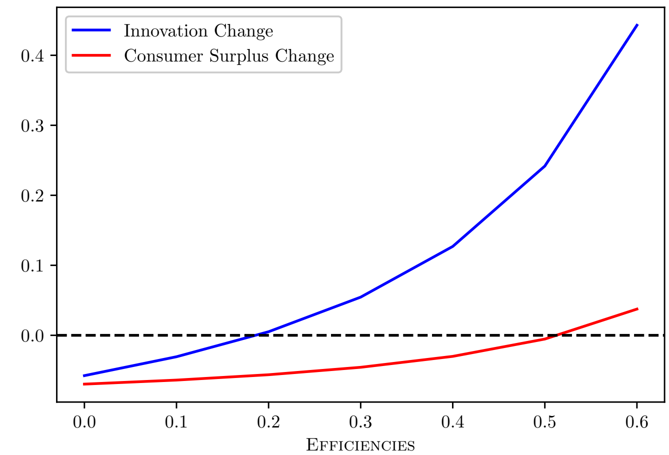
Notes: Closed form solution of a model with 3 firms (2 merging and 1 outsider). The demand system is assumed to be linear $q_i(p_i, \bar{p}_{-i}) = 2 - p_i + 0.3\sum_{j\neq i} p_j$, the same figure arises with CES and Logit demand as shown in Appendix **??**. Marginal cost is linear in innovation $c(x_i) = 1 - x_i$. Fixed costs and efficiencies are quadratic $F(x_i) = \frac{1}{2}x_i^2,\; G(x_i, x_k) = \frac{1}{2}x_i x_k = \frac{1}{2}x_i^2$. The blue line represents changes in innovation x_i, while the red line shows changes in consumer surplus CS. Negative numbers mean a decrease, and vice-versa for positive numbers.

of innovation and in terms of competition. Since consumer surplus is determined by prices, then also consumer surplus increases with efficiency gains λ.

$$\partial_\lambda x^M(\lambda) > 0, \begin{matrix} \partial_\lambda q^M(\lambda) > 0 \\ \partial_\lambda p^M(\lambda) < 0 \end{matrix}$$

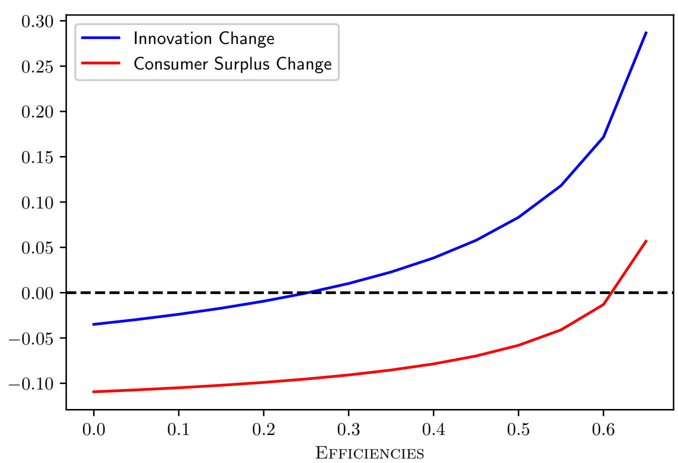
What then is the effect of a merger for varying levels of efficiencies λ? Figure 1 shows changes (after the merger minus before the merger) of consumer surplus and innovation depending on the level of efficiencies. Figure 2 and Figure 3 show the same graph for logit and CES demand systems. As a general feature of these models, the figure shows that for no efficiencies $\lambda = 0$ the merger results in a decrease of both innovation and consumer surplus. As efficiencies grow there is a point where innovation does not change, but consumer surplus is still decreasing. Above this point the effect on innovation is positive, while the competitive outcome on consumer surplus is still negative. Deterrence will be most effective for mergers that have low efficiencies λ, and these tend to be merger that are most detrimental to innovation. However, some anticompetitive mergers have positive effect on innovation. Therefore, the deterrence effect of antitrust authorities on innovation is a priori ambiguous. If most of the deterred mergers have positive effects on innovation, then forcing firms to report their transaction to the authorities has a negative impact on overall innovation. This explains the need for an empirical analysis of the issue.

Figure 2: Changes in Consumer Surplus and Innovation after a merger with logit demand



Notes: Data coming from closed form solution of a model with 3 firms (2 merging and 1 outsider). The demand system is assumed to be logit $q_i(p_i, \bar{p}_{-i}) = \exp((1-p_i)/0.4)/\sum_j \exp((1-p_j)/0.4)$. Marginal cost is linear in innovation $c(x_i) = 1 - x_i$. Fixed costs and efficiencies are quadratic $F(x_i) = \frac{1}{2}x_i^2,\ G(x_i, x_k) = \frac{1}{2}x_i x_k = \frac{1}{2}x_i^2$. The blue line represents changes in innovation x_i, while the red line shows changes in consumer surplus CS. Negative numbers mean a decrease, and vice-versa for positive numbers.

Figure 3: Changes in Consumer Surplus and Innovation after a merger with CES demand



Notes: Data coming from closed form solution of a model with 3 firms (2 merging and 1 outsider). The demand system is assumed to be CES $q_i(p_i, \bar{p}_{-i}) = p_i^{-2}/\sum_j p_j^{-1}$. Marginal cost is linear in innovation $c(x_i) = 1 - x_i$. Fixed costs and efficiencies are quadratic $F(x_i) = \frac{1}{2}x_i^2,\ G(x_i, x_k) = \frac{1}{2}x_i x_k = \frac{1}{2}x_i^2$. The blue line represents changes in innovation x_i, while the red line shows changes in consumer surplus CS. Negative numbers mean a decrease, and vice-versa for positive numbers.

A.2.2 Merger Decision (t=1)

Whenever two firms have the possibility to merge the respective managers have to convince their shareholders and they have to negotiate conditions that satisfy both parties. This process yields an uncertain outcome, and managers can exert effort to increase the chance that an agreement is reached and the merger is proposed to the authorities. Managers care about expected profits, they obtain π_b if the merger does not realize with probability $(1-\alpha^*\varphi(\alpha^*))$, while they obtain π_M if the merger realizes with probability $\alpha^*\varphi(\alpha^*)$. This means that their expected profits is $\hat{\pi}\alpha^*\varphi(\alpha^*))$ where $\hat{\pi} = \pi_M - \pi_b$ is the change in profits after merger. On the other hand, the manager pays a cost $\Gamma(\varphi)$ to raise the probability of proposing the merger φ that satisfies $\Gamma(0) = 0, \Gamma' > 0, \Gamma'' > 0, \Gamma''' \leq 0$. Equation (1) describes the resulting problem of the manager, who chooses optimally φ.

$$\varphi^*(\alpha^*) = \underset{\varphi}{\text{argmax}}\ \hat{\pi}\alpha^*\varphi - \Gamma(\varphi) \tag{1}$$

From the First Order Condition of the problem $\Gamma'(\varphi) = \hat{\pi}\alpha^*$ and the fact that Γ is increasing, one can immediately derive that the optimal $\varphi^*(\alpha^*)$ is increasing in α^*. This is a key step of the deterrence mechanism: the lower is the probability of being approved α, the lower is the effort made by the manager φ.

A.2.3 Antitrust Policy with imperfect information (t=0)

In the European Union, the main problem of an antitrust authority is to discern anticompetitive mergers from procompetitive ones. The European Commission has the power to block mergers, and it spends resources to investigate proposed mergers. Therefore, one can assume that the authority observes the change in consumer surplus $\Delta CS = CS_M - CS_b$ with an error:

$$\widehat{\Delta CS} = \Delta CS + \varepsilon, \ \ \varepsilon \sim \phi(0,\sigma)\ symmetric$$

Then the authority will allow the merger if the consumer surplus outcome $\widehat{\Delta CS}$ is above a certain threshold H, which can be considered the level of harm that the authority is willing to tolerate. Therefore, equation (2) shows the probability that a merger is allowed.

$$\begin{aligned} \alpha^*(\Delta CS,\sigma) &= P(\widehat{\Delta CS} > H) = P(\Delta CS + \varepsilon > H) = P(\varepsilon > -\Delta CS + H) \\ &\overset{sym}{=} P(\varepsilon < \Delta CS - H) = \Phi_\sigma(\Delta CS - H) \end{aligned} \tag{2}$$

Given the properties of CDF Φ_σ then the optimal α^* is increasing in the true consumer surplus outcome ΔCS. Lemma 1 follows from the endogenous merger choice, which implies that also the probability that firms propose a merger φ^* is increasing in ΔCS.

Lemma 1 *Mergers with lower ΔCS have a lower chance of being approved α, and through deterrence a lower chance of being proposed φ.*

A.2.4 Alternative Antitrust Policy with Imperfect Enforcement

In the US system, the antitrust authority has to challenge mergers in court and it has to convince a judge to rule against the transaction. This is an uncertain process, and thus the authority has an imperfect ability to block mergers. For simplicity of exposition I assume the the authority knows the effect that the merger has on consumer surplus CS. Similarly to the firms, authorities care about expected consumer surplus. Therefore, their expected payoff is $\hat{C}S\alpha^*\varphi(\alpha^*)$, where $\hat{C}S = CS_M - CS_b$

is the change in consumer surplus caused by the merger. In order to decrease the probability that the merger is allowed by the judge, the authority exerts a costly effort that results in a cost $\Phi(\alpha)$ that satisfies $\Phi(1) = 0, \Phi' < 0, \Phi'' \geq 0$. This cost represents resources, employees time and effort that the authority needs to spend to convince the judge. Equation (3) shows the problem of the antitrust authority.

$$\alpha^* = \underset{\alpha}{\text{argmax}} \ \hat{CS}\alpha\varphi^*(\alpha) - \Phi(\alpha) \tag{3}$$

If the merger is procompetitive, meaning that it is beneficial to consumers because $\hat{CS} \geq 0$, then it is clear that the authority has no incentive to challenge the merger. In this case the merger is allowed with probability $\alpha^* = 1$. On the other hand, if the merger is anticompetitive, if $\hat{CS} < 0$, then the problem is well defined and the SOC holds, given the properties of Φ. Then, from the Implicit Function Theorem, one can derive that the optimal α^* is increasing with consumer surplus changes $\hat{CS}$. Lemma 1 follows from firms deterrence, which implies that also the probability that firms propose a merger φ^* is increasing in $\hat{CS}$.

Under some assumptions on the functional form to Γ and Φ, it is possible to derive a closed form solution for the optimal antitrust policy α^*. The firm decides optimally the probability of merger φ paying a cost $\Gamma(\varphi) = \gamma\varphi^2/2$. Therefore, the manager problem is:

$$\varphi^*(\alpha) = \underset{\varphi}{\text{argmax}} \ \hat{\pi}\alpha\varphi - \gamma\varphi^2/2$$

From FOC of the manager: $\gamma\varphi = \hat{\pi}\alpha \ \Rightarrow \ \varphi^*(\alpha) = \hat{\pi}\alpha/\gamma$ increasing.

Antitrust Authority chooses a probability of allowing the merger α by paying a cost $\Phi(\alpha) = \phi(1-\alpha)$ Therefore the authority problem is:

$$\alpha^* = \underset{\alpha}{\text{argmax}} \ \frac{\hat{CS}\hat{\pi}}{\gamma}\alpha^2 - \phi(1-\alpha)$$

From FOC of the authority: $2\frac{\hat{CS}\hat{\pi}}{\gamma}\alpha + \phi = 0$. Then one can derive the closed form solution of the optimal antitrust policy:

$$\alpha^* = \frac{\phi\gamma}{2(-\hat{CS})\hat{\pi}} \ \Rightarrow \ \begin{array}{ll} d_\gamma\alpha^* > 0 & d_\phi\alpha^* > 0 \\ d_{\hat{CS}}\alpha^* > 0 & d_{\hat{\pi}}\alpha^* < 0 \end{array}$$

A.2.5 Predictions

After the policy change described in this paper all mergers below a certain threshold become non-notifiable to the antitrust authorities. This is equivalent to allowing every merger, meaning that $\alpha = 1$. The immediate consequence of this is that the number of mergers increases, as it is shown by Lemma 2. This comes not only from the fact that mergers are not blocked, but also from the fact that firms are more likely to propose their transactions now that there is no chance they will be blocked. This prediction will be verified in the data, as there is an increase in the number of horizontal mergers that are not notified to the authorities.

Lemma 2 *The total number of mergers increases when mergers become non-notifiable.*

Proof: Call the probability of merger realization $\alpha\varphi(\lambda) = \alpha^*(\hat{CS}(\lambda))\varphi^*(\alpha^*(\hat{CS}(\lambda)))$. Normalize the total number of possible mergers to 1, so that it is equal to $\int_\Lambda dF(\lambda) = 1$, where $\lambda \sim F(\lambda) \in \mathbb{R}^\Lambda$. When mergers are non-reportable, $\alpha = 1$ and the probability of merger realization is $\varphi(1)$ for any $\lambda \in \Lambda$. Since $\alpha(\lambda) \leq 1$ for any λ such that $\hat{CS}(\lambda) \leq 0$, then

$$\varphi(1) \geq \alpha\varphi(\lambda) \quad \forall \lambda \in \Lambda$$

Which implies that:

$$\underbrace{\int_\Lambda \varphi(1)dF(\lambda)}_{\text{\# of non-reportable mergers}} \geq \underbrace{\int_\Lambda \alpha\varphi(\lambda)dF(\lambda)}_{\text{\# of reportable mergers}}$$

Q.E.D.

The effect on innovation is more ambiguous. Some of the mergers that realize due to the policy change might have a positive effect on innovation, as one can see from Figure 1. Therefore, as Proposition 1 shows, the overall effect on innovation depends on the distribution of mergers. If enough anticompetitive mergers have a negative effect on innovation, then the policy change will result in less innovation. Consequently this is an empirical question, and indeed the results of this paper shows that after the amendment non-notified horizontal mergers lead to less innovation.

Proposition 1 *If the average innovation change generated by all possible mergers is negative, then the average innovation change generated by realized mergers is lower when mergers become non-notifiable.*

Proof: Call the probability of merger realization $\alpha\varphi(\lambda) = \alpha^*(\hat{CS}(\lambda))\varphi^*(\alpha^*(\hat{CS}(\lambda)))$. Call innovation change implied by mergers $\hat{x}(\lambda) = x_M(\lambda) - x_B$. From the assumption that average innovation change is negative:

$$\int_\Lambda \hat{x}(\lambda)dF(\lambda) \leq 0 \Rightarrow \int_\Lambda \hat{x}(\lambda)\overbrace{[\varphi(1) - \alpha\varphi(0)]}^{\geq 0}dF(\lambda) \leq 0$$

The implication comes from the fact that $[\varphi(1) - \alpha\varphi(0)]$ is a positive scalar and does not depend from λ. Since $[\varphi(1) - \alpha\varphi(\lambda)]$ is decreasing in λ and $\min_\Lambda \lambda = 0$, then $[\varphi(1) - \alpha\varphi(\lambda)] \leq [\varphi(1) - \alpha\varphi(0)] \ \forall \lambda \in \Lambda$. Therefore:

$$\Rightarrow \int_\Lambda \hat{x}(\lambda)[\varphi(1) - \alpha\varphi(\lambda)]dF(\lambda) \leq \int_\Lambda \hat{x}(\lambda)[\varphi(1) - \alpha\varphi(0)]dF(\lambda) \leq 0$$

$$\Rightarrow \underbrace{\int_\Lambda \hat{x}(\lambda)\varphi(1)dF(\lambda)}_{\substack{\text{innov. change} \\ \text{non-reportable mergers}}} \leq \underbrace{\int_\Lambda \hat{x}(\lambda)\alpha\varphi(\lambda)dF(\lambda)}_{\substack{\text{innov. change} \\ \text{reportable mergers}}}$$

Q.E.D.

One last result of this model is that a negative effect on innovation implies less consumer surplus, as per Corollary 1. The effect on consumer surplus is worse than the effect on innovation, as shown by

Figure 1. Therefore if mergers have negative effects on innovation, a fortiori even more mergers will decrease consumer surplus. This prediction I cannot verify with the available data, but it warrants future empirical analysis of the price effects of these mergers.

Corollary 1 *If the average innovation change generated by realized mergers is lower when mergers become non-notifiable, then so is the average consumer surplus change, when consumer surplus is well behaved.*

Proof: Call the probability of merger realization $\alpha\varphi(\lambda) = \alpha^*(\hat{CS}(\lambda))\varphi^*(\alpha^*(\hat{CS}(\lambda)))$. Call CS change implied by mergers $\hat{CS}(\lambda) = CS_M(\lambda) - CS_B$. Consider consumer surplus well behaved if $\hat{CS}(\lambda) \leq \hat{x}(\lambda)$ for all $\lambda \in \Lambda$

$$\underbrace{\int_\Lambda \hat{x}(\lambda)\varphi(1)dF(\lambda)}_{\substack{\text{innov. change}\\\text{non-reportable mergers}}} \leq \underbrace{\int_\Lambda \hat{x}(\lambda)\alpha\varphi(\lambda)dF(\lambda)}_{\substack{\text{innov. change}\\\text{reportable mergers}}} \Rightarrow \int_\Lambda \hat{x}(\lambda)[\varphi(1) - \alpha\varphi(\lambda)]dF(\lambda) \leq 0$$

Since consumer surplus is well behaved, then $\hat{CS}(\lambda) \leq \hat{x}(\lambda)$, which implies:

$$\int_\Lambda \hat{CS}(\lambda)[\varphi(1) - \alpha\varphi(\lambda)]dF(\lambda) \leq 0 \Rightarrow \underbrace{\int_\Lambda \hat{CS}(\lambda)\varphi(1)dF(\lambda)}_{\substack{\text{CS change}\\\text{non-reportable mergers}}} \leq \underbrace{\int_\Lambda \hat{CS}(\lambda)\alpha\varphi(\lambda)dF(\lambda)}_{\substack{\text{CS change}\\\text{reportable mergers}}}$$

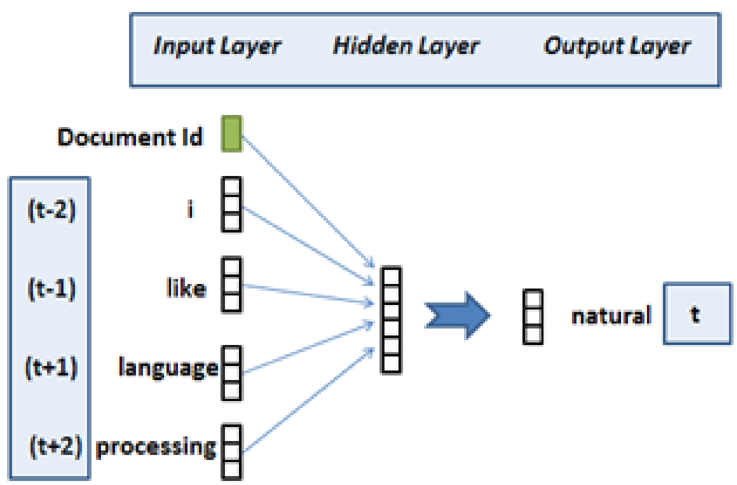
Q.E.D.

B Natural Language Processing

B.1 Doc2Vec and word embeddings

Word Embeddings are vectors representing the semantic meaning of words and text. To each word it is assigned a vector W_j of 300 real numbers, and the document itself is assigned a vector P_i of 30o real numbers. Then, the algorithm cycles through all words in all documents, trying to predict them using the surrounding words, the context. To understand the training process, consider as an example the document "I like natural language processing", as if it were one of the patent abstract of the corpus. Figure 4 shows the step the algorithm takes to predict the word "natural". In the Input Layer, the algorithm uses the vector representing words surrounding "natural" and the vector representing the document. In the Hidden Layer, a shallow neural network combines all the vectors in the Input Layer to predict the Output Layer.[62] This is the vector representing the word "natural". Then the algorithm repeats this process for every word, and for every patent, several times, changing the vectors W_j and P_i to maximize the predicting power. The performance of the prediction exercise is not the goal, however. After this training process the representing vectors incorporate the semantic meaning as represented by the context in which the word is used more often.[63]

Figure 4: Doc2Vec algorithm on the sentence "I like natural language processing"



B.2 Second Example

Here I report a further example of patents with high similarity. In particular, this is the couple of patents with the second highest similarity in the Pfizer-Pharmacia merger.

[62] A shallow neural network is the simplest neural network. It is composed by a single layer of neurons, which transforms the input in the output.

[63] See also a gentle introduction to Doc2Vec

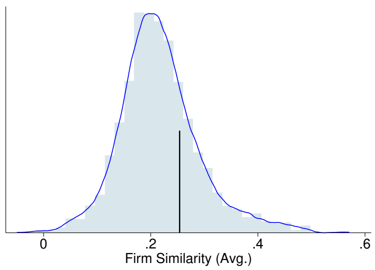
US 6090852 A: Substituted... acids as therapeutic agents
"Compounds... and salts thereof, are matrix metalloprotease inhibitors." [Filed: Jan 20, 1999]

US 6809111 B2 : Prodrugs of COX-2 inhibitors
"A compound of... or a pharmaceutically-acceptable salt thereof, suitable for use in the treatment of a cyclooxygenase-2 mediated disease is provided... and a method for treatment of a cyclooxygenase-2 mediated disease..." [Filed: May 15, 2003]

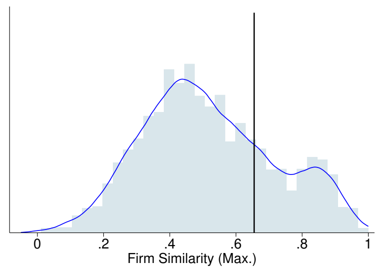
The Pharmacia patent is the same as the first example, as it seems that this is a field in which there is quite the overlap between the two companies. Also in this case there is a reason if this particular couple of patents has such a high similarity. Dickens et al. (2002) reports that COX-2 inhibitors and matrix metalloprotease inhibitors are effective against various cancer types. Then the authors propose that the combined use of both compounds could prove even more beneficial. This is an example of possible merger efficiencies, although it remains to be proven that they are merger specific. Again, such a connection between COX-2 inhibitors and matrix metalloprotease inhibitors is likely due to both terms appearing in similar contexts in other patents.

B.3 Similarity Statistics

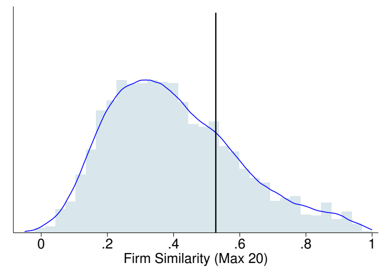
Figure 5: Distribution of Patent Similarity Statistics



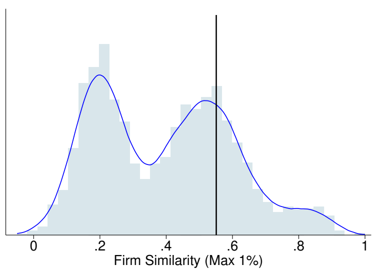
(a) Average Similarity



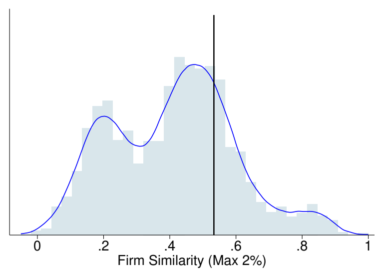
(b) Max Similarity



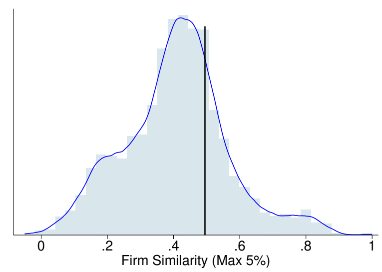
(c) Max 20 Similarity



(d) Max 1% Similarity



(e) Max 2% Similarity



(f) Max 5% Similarity

Notes: Distribution of similarity statistics computed on all mergers available in the dataset.

B.4 Predict EC decisions

Table 1: Correlation Table of Similarity Measures and EC Decisions

Variables	FTC	SIC	Max	Max 20	Max 1%	Max 2%	Max 5%
EC Decisions	1.00						
SIC 4 Digits	0.17	1.00					
Max Patent Sim	0.33	0.57	1.00				
Max 20 Patent Sim	0.34	0.61	0.99	1.00			
Max 1% Patent Sim	0.35	0.50	0.81	0.85	1.00		
Max 2% Patent Sim	0.33	0.52	0.78	0.83	0.99	1.00	
Max 5% Patent Sim	0.26	0.54	0.73	0.77	0.96	0.97	1.00

Notes: Correlation table of various definitions of horizontal mergers in the sample used for the validation exercise. The variable EC Decisions is a dummy equal to 1 for mergers defined as horizontal and 0 otherwise. The variable SIC 4 is another dummy equal to 1 if the merging firms have the same 4 digit SIC code. The remaining variables are continuous measures of similarity.

B.5 Predict FTC decisions

In order to evaluate patent similarity statistics I use them to predict horizontal mergers as defined by antitrust authorities. I collected by hand all available official decisions of the Federal Trade Commis-

Table 2: Performance of different cutoff rules on EC Decisions.

EC	$> pc(25)$	$> pc(50)$	$> pc(75)$	$> pc(90)$	$> pc(95)$
Horizontal	100	96	74	50	37
Non-Horizontal	16	25	59	75	77

Notes: Percentage of correctly predicted horizontal (dark grey) and non-horizontal (light grey) mergers, as defined by the EC, with increasing cutoff. The cutoff is in percentiles of the statistic distribution. The exercise is performed on the max 2% statistic, and similar patterns emerge with other statistics.

Table 3: Performance of Similarity Statistics with threshold rule $[> pc(75)]$ on EC Decisions.

EC	SIC	Max	Max 20	Max 1%	Max 2%	Max 5%
Horizontal	50	86	88	76	74	74
Non-Horizontal	70	29	29	59	59	71

Notes: Percentage of correctly predicted horizontal (dark grey) and non-horizontal (light grey) mergers, as defined by the EC, with different methodologies. SIC stands for horizontal mergers defined using the same 4 digit SIC code, as in Wollmann (2019). The exercise is performed using the $> pc(75)$ cutoff, and similar patterns emerge with other cutoffs.

sion (hereafter FTC).[64] The FTC tags each decision as *Horizontal* or *Vertical*. A few cases are tagged as both. I consider a decision to be horizontal if it is tagged as *Horizontal* and it is not tagged as *Vertical*. From the original pool of public decisions I remove mergers between companies that do not have a portfolio of patents. These are mainly transactions between hospitals and clinics, or exchanges of pipelines and extraction rigs between oil companies. As a result I have 20 FTC decisions, 17 of which horizontal. Given this number, the rest of this exercise should be considered as narrative evidence, rather than significant statistical evidence. Regardless, this limit is given by decisions published by the FTC, since no other decisions on innovating firms have been issued by the antitrust authority.

For this set of mergers controlled by the FTC I build a dummy variable that is 1 for horizontal mergers, and 0 otherwise. Then I build a dummy variable that is 1 if the merging parties have the same 4 digits SIC code, and 0 otherwise. This represent the standard in the existing literature, as one can see in Wollmann (2019), and the one I compare my measures with. As a first step I compute correlations of these variables in Table 4, to see which one is most similar to FTC definitions. The SIC definition is positively correlated with FTC definition, but with a small value of 0.15. Patent similarity measures have a higher correlation, outperforming the SIC one. It is worth noting that these measures have a positive, although small, correlation with the SIC dummy. Moreover, all these similarity measures have a strong correlation between each other, since they are representing the same

[64] I access the full set of public FTC decisions from their website, using their Advanced Search option (https://www.ftc.gov/enforcement/cases-proceedings/advanced-search).

concept.

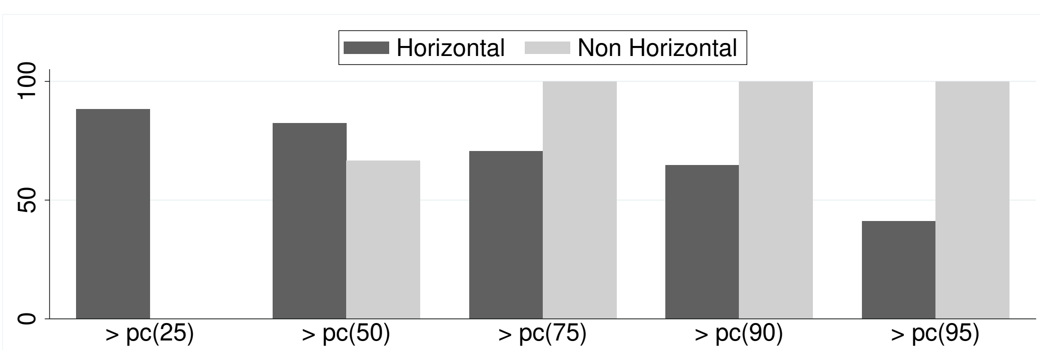
Table 4: Correlation Table of Similarity Measures and FTC Decisions

Variables	FTC	SIC	Max	Max 20	Max 1%	Max 2%	Max 5%
FTC Decisions	1.00						
SIC 4 Digits	0.15	1.00					
Max Patent Sim	0.29	0.05	1.00				
Max 20 Patent Sim	0.24	0.03	0.92	1.00			
Max 1% Patent Sim	0.35	0.03	0.86	0.92	1.00		
Max 2% Patent Sim	0.36	0.03	0.87	0.91	0.90	1.00	
Max 5% Patent Sim	0.40	0.03	0.87	0.85	0.79	0.88	1.00

Notes: Correlation table of various definitions of horizontal mergers in the sample used for the validation exercise. The variable FTC Decisions is a dummy equal to 1 for mergers defined as horizontal and 0 otherwise. The variable SIC 4 is another dummy equal to 1 if the merging firms have the same 4 digit SIC code. The remaining variables are continuous measures of similarity.

The correlation table compares a dummy variable for FTC with continuous measures of patent similarity. These measures are informative by themselves, and they can be used as an identification device. As a robustness exercise I show that using continuous measures of similarity in the identification strategy leads to results similar to the main ones. However, if one wants to generate a 0-1 dummy variable identifying horizontal mergers using similarity statistics, one needs to determine a threshold above which a merger is considered horizontal. Figure 6 reports variables constructed with various thresholds compared with FTC definitions. Each bar represents the percentage of correct predictions. This figure represent type I and type II errors in predicting horizontal mergers. A lower cutoff, like the 25^{th} percentile is very accurate in predicting horizontal mergers, but does poorly in predicting non-horizontal ones. Conversely, a cutoff like the 95^{th} percentile predicts horizontal mergers poorly. The most reasonable cutoff is the 75^{th} percentile, and this is consistent across various similarity measures.

Figure 6: Performance of different cutoff rules on FTC Decisions.



Notes: Percentage of correctly predicted horizontal (dark grey) and non-horizontal (light grey) mergers, as defined by the FTC, with increasing cutoff. The cutoff is in percentiles of the statistic distribution. The exercise is performed on the max 2% statistic, and similar patterns emerge with other statistics. Numbers for the histogram are reported in Table 5

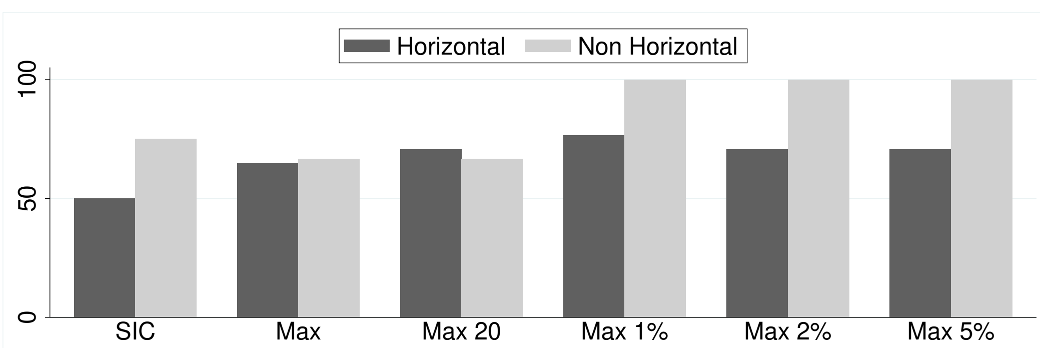
Table 5: Performance of different cutoff rules on FTC Decisions.

FTC	$> pc(25)$	$> pc(50)$	$> pc(75)$	$> pc(90)$	$> pc(95)$
Horizontal	88	82	71	65	41
Non-Horizontal	0	67	100	100	100

Notes: Percentage of correctly predicted horizontal (dark grey) and non-horizontal (light grey) mergers, as defined by the FTC, with increasing cutoff. The cutoff is in percentiles of the statistic distribution. The exercise is performed on the max 2% statistic, and similar patterns emerge with other statistics.

Once the cutoff rule is set to the 75^{th} percentile, I compare similarity statistics in Figure 7. Using the SIC industry classification one can predict only 50% of horizontal mergers, while all patent similarity statistics outperform this measure. Similarly to the correlation results in Table 4, the *Max* $x\%$ statistics perform better than the simple maximum value of the similarity matrix. In the Robustness section I show that all results hold true regardless of the chosen patent similarity statistic. This is to be expected, as all these measures capture the same concept: how close are the products of two merging firms.

Figure 7: Performance of Similarity Statistics on FTC Decisions.



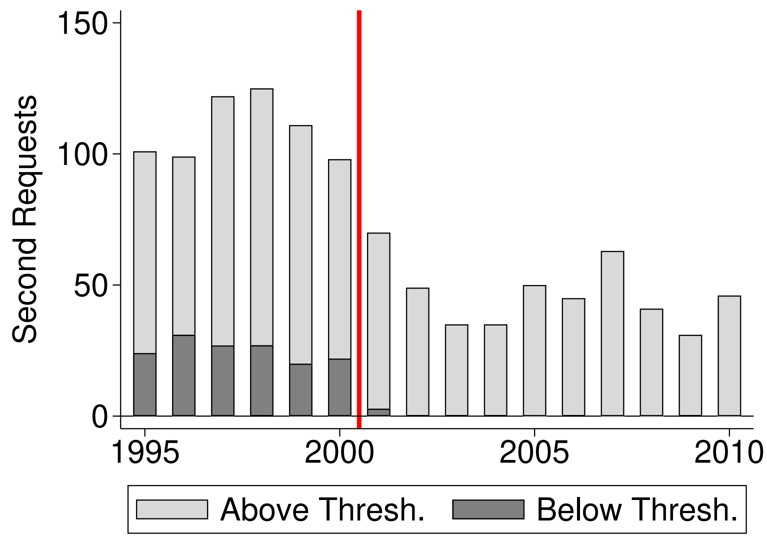
Notes: Percentage of correctly predicted horizontal (dark grey) and non-horizontal (light grey) mergers, as defined by the FTC, with different methodologies. SIC stands for horizontal mergers defined using the same 4 digit SIC code, as in Wollmann (2019). The exercise is performed using the $> pc(75)$ cutoff, and similar patterns emerge with other cutoffs. Numbers for the histogram are reported in Table 6.

Table 6: Performance of Similarity Statistics with threshold rule $[> pc(75)]$ on FTC Decisions.

FTC	SIC	Max	Max 20	Max 1%	Max 2%	Max 5%
Horizontal	50	65	71	76	71	71
Non-Horizontal	75	67	67	100	100	100

Notes: Percentage of correctly predicted horizontal (dark grey) and non-horizontal (light grey) mergers, as defined by the FTC, with different methodologies. SIC stands for horizontal mergers defined using the same 4 digit SIC code, as in Wollmann (2019). The exercise is performed using the $> pc(75)$ cutoff, and similar patterns emerge with other cutoffs.

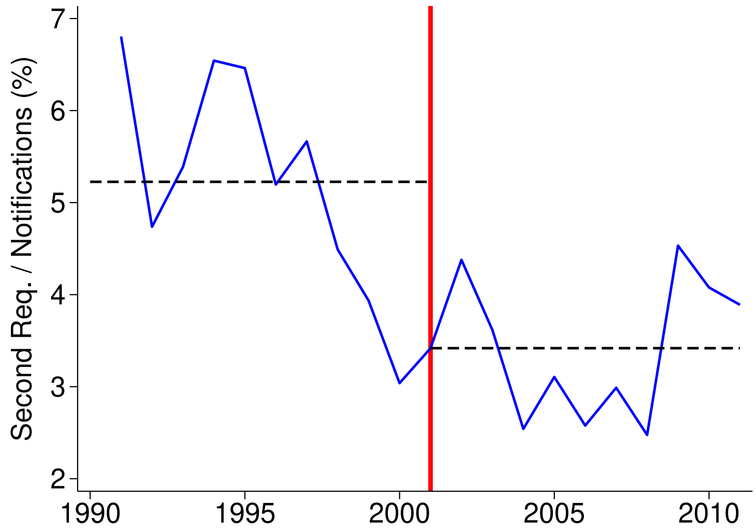
Figure 8: Number of Notifications received by US Antitrust Authorities.



Notes: The graph reports number of second requests above and below the new threshold of $\$50$ mln introduced with the amendment in December 2000. The red vertical line represents the introduction of the Amendment.

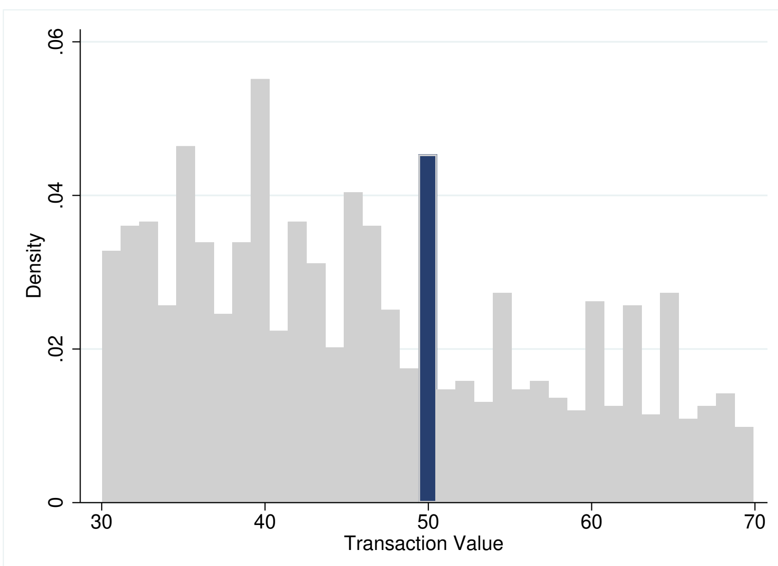
C Additional Results

Figure 9: Share of Notifications resulting in Second requests above new threshold.



Notes: The graph reports number of second requests above new threshold of $50 mln divided by the number of notifications above the same threshold. The red vertical line represents the introduction of the Amendment.

Figure 10: Distribution of Mergers by Transaction Value.



Notes: The graph reports the distribution of mergers by their transaction value in the year 2001-2004 following the Amendment. Transaction value is defined as the sum of money which was paid to the acquired firm by the acquirer. In these years the threshold of $50 million was not adjusted to inflation. The blue bar comprises all transaction that are below bt close to the $50 million threshold.

Table 7: Coefficients of triple Diff-in-Diff and simple Diff-in-Diff.

Variables	Relative Cit. Average All Mergers	Relative Cit. Average Newly Exempt	Relative Cit. Average Never Exempt	Citations All Mergers	Citations Newly Exempt	Citations Never Exempt
$I^{Post} \cdot I^{Ex} \cdot I^{Hor}$	-0.357**			-0.523**		
	(0.152)			(0.212)		
$I^{Post} \cdot I^{Hor}$		-0.136	0.118		-0.191	0.345
		(0.118)	(0.149)		(0.145)	(0.229)
Implied Change	-30.0%	-12.7%	12.5%	-40.0%	-17.4%	41.2%
Observations	2,601	1,782	819	2,610	1,789	821
R-squared	0.080	0.087	0.215	0.156	0.145	0.301
Year FE	YES	YES	YES	YES	YES	YES
Industry FE	YES	YES	YES	YES	YES	YES
State FE	YES	YES	YES	YES	YES	YES
Cluster SE	SIC4	SIC4	SIC4	SIC4	SIC4	SIC4

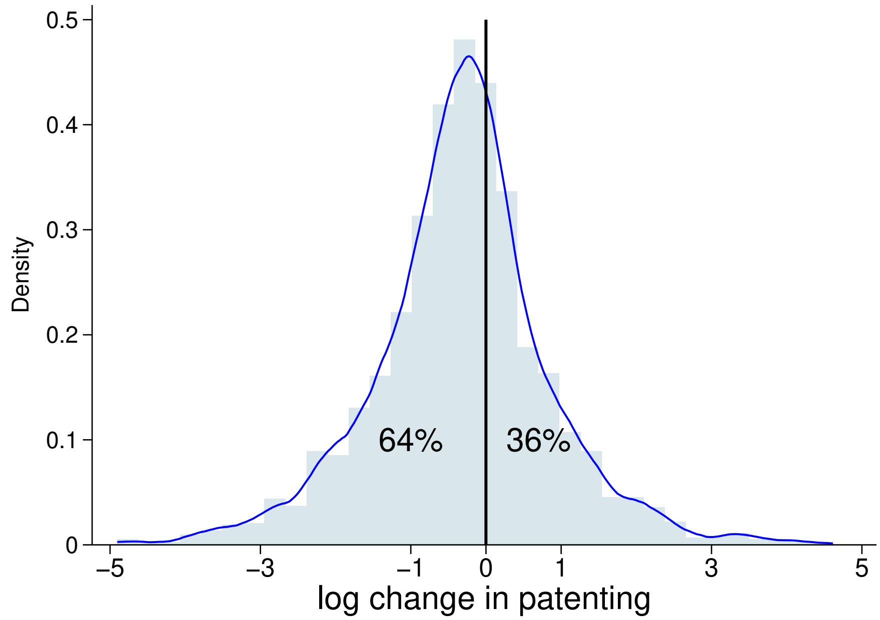
Notes: Coefficients of triple diff-in-diff equation 5 with *Single Period Changes* as dependent variable ΔP are shown in the first and the third column. The other columns represent coefficients of a simple diff-in-diff between horizontal and non-horizontal mergers. This exercise is conducted for newly exempt mergers and for never exempt mergers. The coefficient of the triple diff-in-diff represents the difference between the two coefficients of the simple diff-in-diff.

Table 8: Policy changes that relaxed notifications requirements

COUNTRY	Year of Amendment	Change in Merger Notifications	Actual Numbers
United States	2000	-70%	From 3500 in 2000 to 1000 in 2001
Italy	2012	-90%	From 459 in 2012 to 59 in 2013
Germany	1999	-37%	From 1888 in 1998 to 1182 in 1999
Spain	2007	-55%	From 132 in 2006 to 58 in 2013
Belgium	2006	-70%	From 60 in 1997 to 17 in 2007
Sweden	2000	-50%	From 168 in 1999 to 84 in 2001
Hungary	2005	-40%	From an average of 70 in 2000-2005 to 42 in 2006-2010
Canada	2009	-9%	From 236 in '08-'09 to 216 in '09-'10
Japan	2010	-70%	From 1000 in 2009 to 300 in the following years
Russia	2005	-48%	From 12000 in 2004 to 6265 in 2005

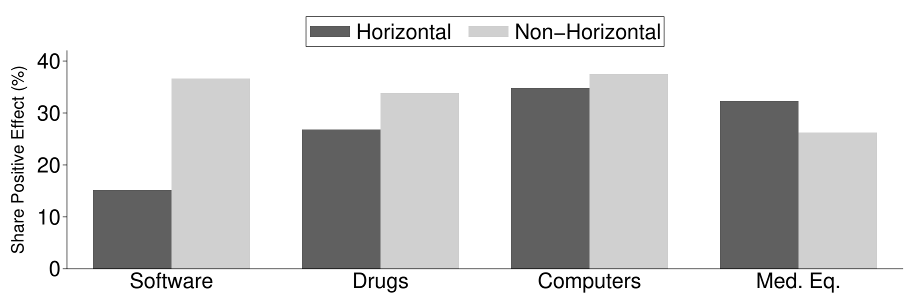
Notes: The table reports information on antitrust policy changes that affected various countries. The third column reports the change in number of merger notifications in percentages. The last column reports the actual notification numbers recovered from official documents of the antitrust authorities of these countries.

Figure 11: Distribution of average change in patenting activity after merger.



Notes: Unconditional distribution of the variable ΔP in the overall sample. The vertical black line divides negative changes from positive ones. The figure shows that 36% of mergers lead to an increase in innovation.

Figure 12: Share of mergers with positive effect on innovation, by sector.



Notes: Share of mergers for which $\Delta P > 0$. Non-horizontal mergers tend to have a higher share of positive effect. The sector with the highest shares are Software and Computers. These are sectors with high complementarities between non-competitors.

Figure 13: Coefficients of triple Diff in Diff for single years around the Amendment.

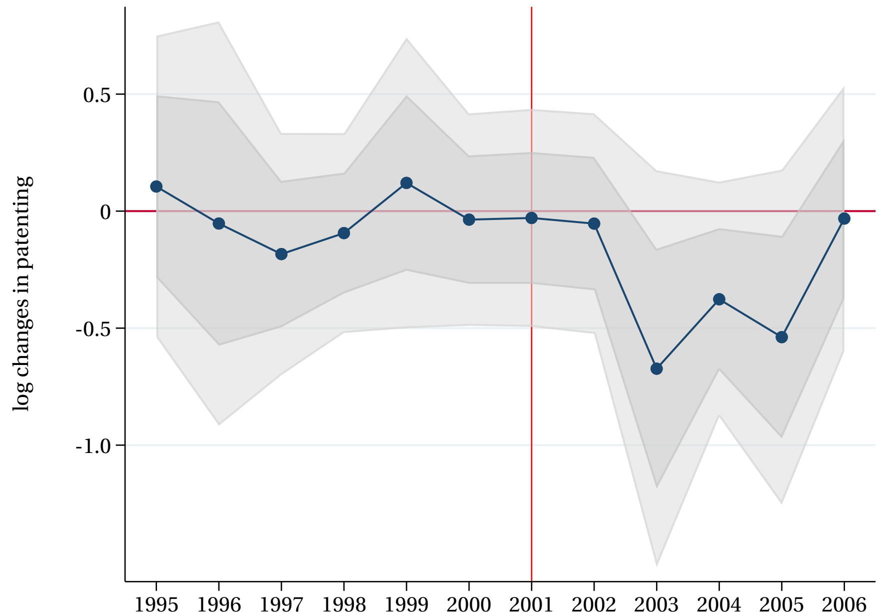


Figure 14: Coefficients of triple Diff in Diff for years in groups of 3 around the Amendment.

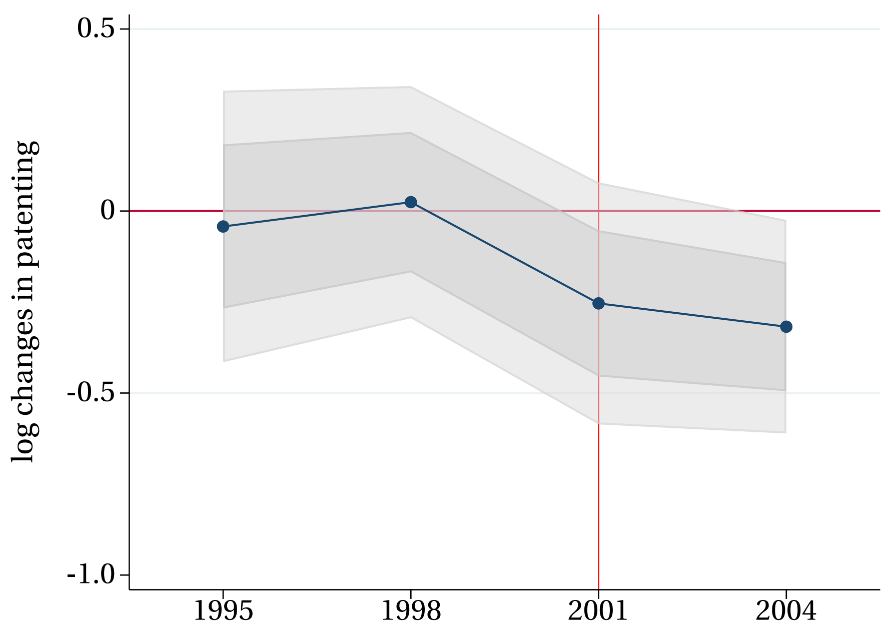


Figure 15: Coefficients of triple Diff in Diff for years in groups of 4 around the Amendment.

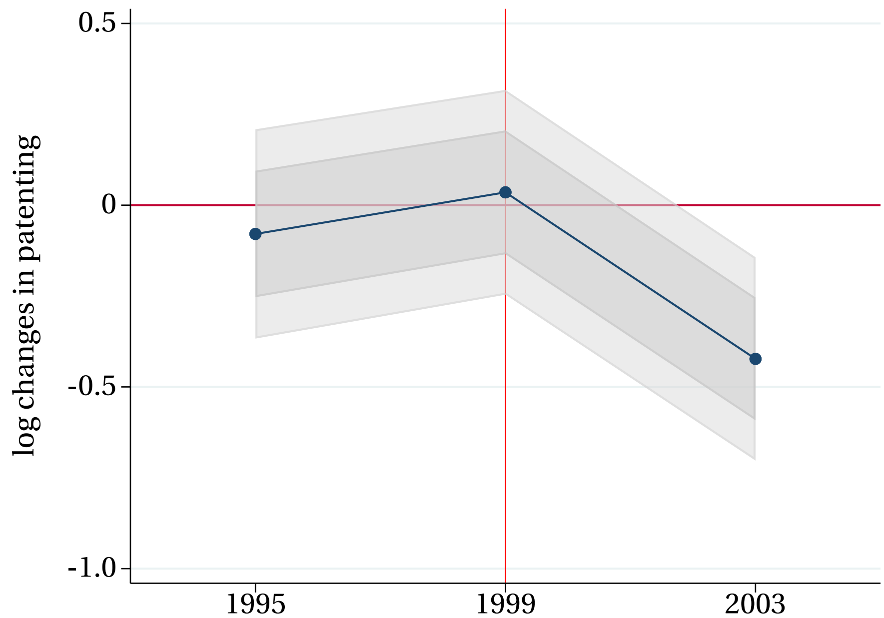
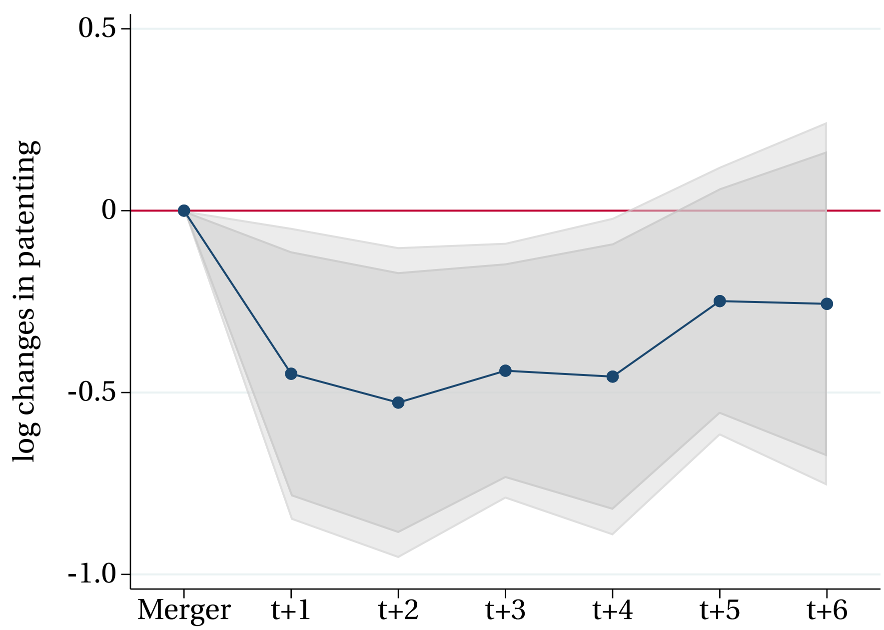


Figure 16: Coefficients for different time span after the merger.



Notes: Coefficients of triple diff-in-diff equation 5 with *Single Period Changes* as dependent variable ΔP. This shows how the effect evolves in the years after a merger. The coefficient in the year of the merger is artificially put to 0, with 0 standard error.

Table 9: Triple difference in differences results for various innovation activity measures.

VARIABLES	(1) Number	(2) Cit.	(3) Relative Cit.	(4) Generality	(5) Originality
$I^{Post} \cdot I^{Ex} \cdot I^{Hor}$	-0.178	-0.523**	-0.357**	-0.0499	-0.184**
	(0.175)	(0.212)	(0.152)	(0.126)	(0.0854)
Observations	2,677	2,610	2,601	2,393	2,480
R-squared	0.105	0.156	0.080	0.106	0.062
Year FE	YES	YES	YES	YES	YES
Industry FE	YES	YES	YES	YES	YES
State FE	YES	YES	YES	YES	YES
Cluster SE	SIC4	SIC4	SIC4	SIC4	SIC4

Notes: Coefficients of triple diff-in-diff equation 5 with various measures patent activity as dependent variable ΔP. Column (1) reports the total number of patents submitted each year. Column (2) the total number of citations received by patents submitted. Column (3) reports the main results computed with Relative Citation Average, which takes into account varying patenting activity in different technology spaces. Column (4) reports Generality, which increases if patents are cited by a diverse array of patents, as computed by $(1 - HHI)$ of citing patent technology spaces. Column (5) reports Originality, which is higher for patents citing a diverse array of patents, as computed by $(1 - HHI)$ of cited patent technology spaces.

Table 10: Sample size by categories before and after the Amendment

	Horizontal		
Exempt	No	Yes	Total
A: All			
Never	494	325	819
Newly	1,366	416	1,782
Total	1,860	741	2,601
B: Before Amendment			
Never	313	201	514
Newly	817	250	1,067
Total	1,130	451	1,581
C: After Amendment			
Never	181	124	305
Newly	549	166	715
Total	730	290	1,020

Notes: This table reports the size of various groups of merging firms composing the sample. Panel A reports the whole sample, comprising both transactions before and after the amendment. Panel B includes only transactions before the amendment, Panel C includes only the ones after the policy change. The Total row in each panel is computed as the sum of Never Exempt ans Newly Exempt rows. The last column is the sum of the first two columns.

Table 11: Triple difference in differences considering only particular sectors

VARIABLES	(1) Baseline	(2) Drugs	(3) Software	(4) Computers	(5) Med. Eq.
$I^{Post} \cdot I^{Ex} \cdot I^{Hor}$	-0.357**	-0.333***	-0.648***	-0.0951	0.175
	(0.152)	(0.0546)	(0.145)	(0.780)	(0.812)
Implied Change	-30.0%	-28.3%	-47.7%	-9.1%	19.2%
Observations	2,601	192	358	95	196
R-squared	0.080	0.192	0.190	0.649	0.117
All FE	YES	YES	YES	YES	YES
Cluster SE	SIC4	SIC4	SIC4	SIC4	SIC4

Notes: This table reports coefficients of equation 5 with various sample specifications. Column (1) reports the baseline result computed on the whole sample. Column (2)-(5) reports results computed on the sample including only mergers from a particular sector, as defined by SIC 3 digit code.

Table 12: Triple difference in differences results computed using "Max" patent similarity.

VARIABLES	(1) Avg	(2) Time t+1	(3) Time t+2	(4) Time t+3	(5) Time t+4	(6) Time t+5	(7) Time t+6
$I^{Post} \cdot I^{Ex} \cdot I^{Hor}$	-0.406**	-0.585***	-0.626***	-0.335	-0.297	-0.369	-0.506*
	(0.162)	(0.186)	(0.211)	(0.222)	(0.211)	(0.226)	(0.269)
$I^{Post} \cdot I^{Ex}$	0.382***	0.196	0.261	-0.00734	0.0856	0.236*	0.0919
	(0.129)	(0.170)	(0.172)	(0.170)	(0.168)	(0.142)	(0.215)
$I^{Post} \cdot I^{Hor}$	0.242*	0.396**	0.356*	0.207	0.235	0.330*	0.609**
	(0.137)	(0.166)	(0.181)	(0.186)	(0.214)	(0.180)	(0.251)
Observations	2,601	1,534	1,408	1,289	1,217	1,105	1,047
R-squared	0.080	0.096	0.111	0.119	0.120	0.143	0.186
Year FE	YES	YES	YES	YES	YES	YES	YES
Industry FE	YES	YES	YES	YES	YES	YES	YES
State FE	YES	YES	YES	YES	YES	YES	YES
Cluster SE	SIC4	SIC4	SIC4	SIC4	SIC4	SIC4	SIC4

Notes: Coefficients of triple diff-in-diff equation 5 with various measures patent activity as dependent variable ΔP. Horizontal mergers are defined using the maximum of patent similarities between the merging firms patent portfolios. Column (1) reports results for average change considering all years around the merger, computed as $\Delta P = ln\left(\sum_{i=1}^{n} P_{t+i}/n\right) - ln\left(\sum_{i=1}^{n} P_{t-i}/n\right)$. Column (2) to (7) report results for single years change after the merger, computed as $\Delta P^{t+i} = ln\left(P_{t+i}\right) - ln\left(P_{t-1}\right)$.

Table 13: Triple difference in differences results computed using "Max 20" patent similarity.

VARIABLES	(1) Avg	(2) Time t+1	(3) Time t+2	(4) Time t+3	(5) Time t+4	(6) Time t+5	(7) Time t+6
$I^{Post} \cdot I^{Ex} \cdot I^{Hor}$	-0.273*	-0.110	-0.628***	-0.334	-0.0526	-0.168	-0.206
	(0.146)	(0.185)	(0.228)	(0.208)	(0.245)	(0.212)	(0.252)
$I^{Post} \cdot I^{Ex}$	0.300**	-0.0761	0.204	-0.0252	-0.0710	0.0805	-0.0827
	(0.120)	(0.157)	(0.184)	(0.178)	(0.187)	(0.147)	(0.210)
$I^{Post} \cdot I^{Hor}$	0.0651	-0.0442	0.196	0.171	-0.0650	-0.0239	0.336
	(0.126)	(0.150)	(0.198)	(0.199)	(0.217)	(0.182)	(0.254)
Observations	2,601	1,534	1,408	1,289	1,217	1,105	1,047
R-squared	0.080	0.091	0.113	0.119	0.119	0.141	0.182
Year FE	YES	YES	YES	YES	YES	YES	YES
Industry FE	YES	YES	YES	YES	YES	YES	YES
State FE	YES	YES	YES	YES	YES	YES	YES
Cluster SE	SIC4	SIC4	SIC4	SIC4	SIC4	SIC4	SIC4

Notes: Coefficients of triple diff-in-diff equation 5 with various measures patent activity as dependent variable ΔP. Horizontal mergers are defined using the mean of the top 20 patent similarities between the merging firms patent portfolios. Column (1) reports results for average change considering all years around the merger, computed as $\Delta P = ln\left(\sum_{i=1}^{n} P_{t+i}/n\right) - ln\left(\sum_{i=1}^{n} P_{t-i}/n\right)$. Column (2) to (7) report results for single years change after the merger, computed as $\Delta P^{t+i} = ln\left(P_{t+i}\right) - ln\left(P_{t-1}\right)$.

Table 14: Triple difference in differences results computed using "Max 1%" patent similarity.

VARIABLES	(1) Avg	(2) Time t+1	(3) Time t+2	(4) Time t+3	(5) Time t+4	(6) Time t+5	(7) Time t+6
$I^{Post} \cdot I^{Ex} \cdot I^{Hor}$	-0.424***	-0.469***	-0.506**	-0.345*	-0.354	-0.236	-0.243
	(0.157)	(0.176)	(0.224)	(0.183)	(0.243)	(0.183)	(0.243)
$I^{Post} \cdot I^{Ex}$	0.365***	0.0909	0.150	-0.0310	0.0815	0.113	-0.110
	(0.123)	(0.151)	(0.156)	(0.148)	(0.154)	(0.132)	(0.196)
$I^{Post} \cdot I^{Hor}$	0.207	0.187	0.158	0.149	0.191	-0.0366	0.315
	(0.130)	(0.151)	(0.162)	(0.168)	(0.205)	(0.166)	(0.237)
Observations	2,601	1,534	1,408	1,289	1,217	1,105	1,047
R-squared	0.081	0.095	0.110	0.119	0.120	0.143	0.181
Year FE	YES	YES	YES	YES	YES	YES	YES
Industry FE	YES	YES	YES	YES	YES	YES	YES
State FE	YES	YES	YES	YES	YES	YES	YES
Cluster SE	SIC4	SIC4	SIC4	SIC4	SIC4	SIC4	SIC4

Notes: Coefficients of triple diff-in-diff equation 5 with various measures patent activity as dependent variable ΔP. Horizontal mergers are defined using the mean of the top 1% patent similarities between the merging firms patent portfolios. Column (1) reports results for average change considering all years around the merger, computed as $\Delta P = ln\left(\sum_{i=1}^{n} P_{t+i}/n\right) - ln\left(\sum_{i=1}^{n} P_{t-i}/n\right)$. Column (2) to (7) report results for single years change after the merger, computed as $\Delta P^{t+i} = ln\left(P_{t+i}\right) - ln\left(P_{t-1}\right)$.

Table 15: Triple difference in differences results computed using "Max 5%" patent similarity.

VARIABLES	(1) Avg	(2) Time t+1	(3) Time t+2	(4) Time t+3	(5) Time t+4	(6) Time t+5	(7) Time t+6
$I^{Post} \cdot I^{Ex} \cdot I^{Hor}$	-0.219	-0.493**	-0.450**	-0.532**	-0.846***	-0.489*	-0.412
	(0.135)	(0.215)	(0.206)	(0.246)	(0.237)	(0.261)	(0.292)
$I^{Post} \cdot I^{Ex}$	0.288***	0.0748	0.129	0.0268	0.240*	0.204*	-0.0611
	(0.108)	(0.136)	(0.135)	(0.130)	(0.145)	(0.121)	(0.165)
$I^{Post} \cdot I^{Hor}$	0.0990	0.319	0.472***	0.396*	0.666***	0.411**	0.589**
	(0.105)	(0.201)	(0.155)	(0.230)	(0.217)	(0.189)	(0.237)
Observations	2,601	1,534	1,408	1,289	1,217	1,105	1,047
R-squared	0.079	0.094	0.110	0.122	0.129	0.144	0.187
Year FE	YES	YES	YES	YES	YES	YES	YES
Industry FE	YES	YES	YES	YES	YES	YES	YES
State FE	YES	YES	YES	YES	YES	YES	YES
Cluster SE	SIC4	SIC4	SIC4	SIC4	SIC4	SIC4	SIC4

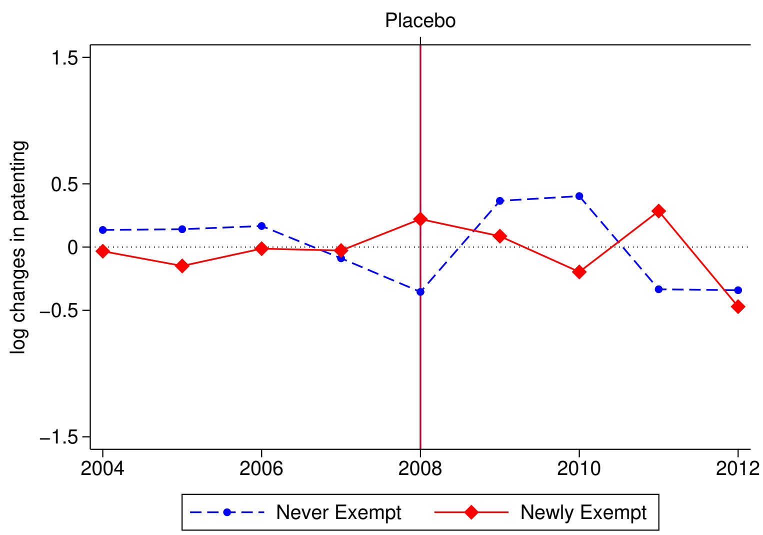
Notes: Coefficients of triple diff-in-diff equation 5 with various measures patent activity as dependent variable ΔP. Horizontal mergers are defined using the mean of the top 5% patent similarities between the merging firms patent portfolios. Column (1) reports results for average change considering all years around the merger, computed as $\Delta P = ln\left(\sum_{i=1}^{n} P_{t+i}/n\right) - ln\left(\sum_{i=1}^{n} P_{t-i}/n\right)$. Column (2) to (7) report results for single years change after the merger, computed as $\Delta P^{t+i} = ln\left(P_{t+i}\right) - ln\left(P_{t-1}\right)$.

Table 16: Triple difference in differences results computed using 2000 as Amendment year.

VARIABLES	(1) Avg	(2) Time t+1	(3) Time t+2	(4) Time t+3	(5) Time t+4	(6) Time t+5	(7) Time t+6
$I^{Post} \cdot I^{Ex} \cdot I^{Hor}$	-0.285**	-0.428**	-0.487***	-0.361**	-0.390*	-0.306*	-0.340
	(0.138)	(0.189)	(0.179)	(0.181)	(0.220)	(0.165)	(0.243)
$I^{Post} \cdot I^{Ex}$	0.228**	0.0644	0.199*	-0.0451	0.0799	0.136	0.0348
	(0.0992)	(0.129)	(0.109)	(0.120)	(0.136)	(0.116)	(0.186)
$I^{Post} \cdot I^{Hor}$	0.166	0.236	0.335***	0.255*	0.241	0.133	0.438**
	(0.110)	(0.151)	(0.129)	(0.138)	(0.181)	(0.139)	(0.183)
Observations	2,601	1,534	1,408	1,289	1,217	1,105	1,047
R-squared	0.078	0.094	0.109	0.121	0.121	0.142	0.182
Year FE	YES	YES	YES	YES	YES	YES	YES
Industry FE	YES	YES	YES	YES	YES	YES	YES
State FE	YES	YES	YES	YES	YES	YES	YES
Cluster SE	SIC4	SIC4	SIC4	SIC4	SIC4	SIC4	SIC4

Notes: Coefficients of triple diff-in-diff equation 5 where the variable I^{Post} is computed considering 2000 as Amendment year. Column (1) reports results for average change considering all years around the merger, computed as $\Delta P = ln\left(\sum_{i=1}^{n} P_{t+i}/n\right) - ln\left(\sum_{i=1}^{n} P_{t-i}/n\right)$. Column (2) to (7) report results for single years change after the merger, computed as $\Delta P^{t+i} = ln\left(P_{t+i}\right) - ln\left(P_{t-1}\right)$.

Figure 17: Placebo exercise with 2008 financial crisis



Notes: The data points of the figure are constructed by averaging the innovation effect of mergers in various groups. The variable ΔP is computed using the relative citation average described in Section 3. Data are residualized on fixed effects used in column (1) of Table 2. Each point reports the difference between horizontal and non-horizontal mergers.

Table 17: Triple difference in differences results using $200 million as HSR threshold.

VARIABLES	(1) Avg	(2) Time t+1	(3) Time t+2	(4) Time t+3	(5) Time t+4	(6) Time t+5	(7) Time t+6
$I^{Post} \cdot I^{Ex} \cdot I^{Hor}$	-0.334**	-0.295*	-0.460***	-0.534***	-0.428**	-0.257	-0.256
	(0.131)	(0.169)	(0.139)	(0.189)	(0.196)	(0.159)	(0.189)
$I^{Post} \cdot I^{Ex}$	0.249***	0.0249	0.158	0.176	0.247**	0.143	-0.0480
	(0.0946)	(0.127)	(0.109)	(0.112)	(0.110)	(0.112)	(0.145)
$I^{Post} \cdot I^{Hor}$	0.160	0.136	0.337***	0.416**	0.335**	0.146	0.361**
	(0.106)	(0.131)	(0.114)	(0.170)	(0.167)	(0.125)	(0.160)
Observations	2,608	1,861	1,743	1,598	1,509	1,393	1,326
R-squared	0.082	0.068	0.078	0.089	0.122	0.128	0.138
Year FE	YES	YES	YES	YES	YES	YES	YES
Industry FE	YES	YES	YES	YES	YES	YES	YES
State FE	YES	YES	YES	YES	YES	YES	YES
Cluster SE	SIC4	SIC4	SIC4	SIC4	SIC4	SIC4	SIC4

Notes: Coefficients of triple diff-in-diff equation 5 where the variable I^{Ex} is computed considering $200 million as HSR threshold. Column (1) reports results for average change considering all years around the merger, computed as $\Delta P = ln\left(\sum_{i=1}^{n} P_{t+i}/n\right) - ln\left(\sum_{i=1}^{n} P_{t-i}/n\right)$. Column (2) to (7) report results for single years change after the merger, computed as $\Delta P^{t+i} = ln\left(P_{t+i}\right) - ln\left(P_{t-1}\right)$.

Table 18: Triple difference in differences results considering transactions close to the threshold.

VARIABLES	(1) Avg	(2) Time t+1	(3) Time t+2	(4) Time t+3	(5) Time t+4	(6) Time t+5	(7) Time t+6
$I^{Post} \cdot I^{Ex} \cdot I^{Hor}$	-0.109	-0.563**	-0.547*	-0.520**	-0.270	0.0770	-0.195
	(0.143)	(0.267)	(0.291)	(0.208)	(0.222)	(0.222)	(0.261)
$I^{Post} \cdot I^{Ex}$	0.273***	-0.00929	0.160	-0.0120	-0.0664	0.00670	-0.204
	(0.0828)	(0.136)	(0.155)	(0.127)	(0.134)	(0.139)	(0.150)
$I^{Post} \cdot I^{Hor}$	-0.0272	0.287	0.334	0.326*	0.111	-0.295	0.365
	(0.119)	(0.206)	(0.219)	(0.190)	(0.199)	(0.183)	(0.229)
Observations	3,534	1,542	1,409	1,321	1,248	1,118	1,070
R-squared	0.063	0.098	0.117	0.144	0.142	0.143	0.188
Year FE	YES	YES	YES	YES	YES	YES	YES
Industry FE	YES	YES	YES	YES	YES	YES	YES
State FE	YES	YES	YES	YES	YES	YES	YES
Cluster SE	SIC4	SIC4	SIC4	SIC4	SIC4	SIC4	SIC4

Notes: Coefficients of triple diff-in-diff equation 5 where only mergers with transaction size below $500 million are considered in the analysis as never exempt. Column (1) reports results for average change considering all years around the merger, computed as $\Delta P = ln\left(\sum_{i=1}^{n} P_{t+i}/n\right) - ln\left(\sum_{i=1}^{n} P_{t-i}/n\right)$. Column (2) to (7) report results for single years change after the merger, computed as $\Delta P^{t+i} = ln\left(P_{t+i}\right) - ln\left(P_{t-1}\right)$.