Università Commerciale "Luigi Bocconi"
PhD School

PhD program in: Economics and Finance

Cycle: 33^{rd}

Disciplinary Field (code): SECS-P/01

Essays on Competition Policy

Advisor: Basile Grassi

Co-Advisor : Luca Sala

PhD Thesis by

Giovanni Morzenti

ID number: 1638790

Year: 2023

Abstract

The aim of competition policy is to balance market power so as to protect and improve consumer welfare. In this thesis I study how antitrust policy affects firms market power, and in turn how this power influences other important economic outcomes, ranging from innovation to redistribution and inequality.

I study how antitrust authorities encourage innovation of merging firms by fostering competition. A change in merger notification rules allows me to build an event study comparing mergers notified to the authorities with non-notified ones. I develop a new text analysis methodology to identify mergers between close competitors, even for small private firms. As a result of the event study, non-notified horizontal mergers lead to 30% less innovation effort, measured as patenting activity. To understand the mechanism driving these findings, I build a model with endogenous merger choice where optimal antitrust policy deters anticompetitive mergers, which are also most detrimental to innovation. An increase in the number of non-notified anticompetitive mergers is consistent with the deterrence effect of antitrust authorities.

Changes in antitrust policy provide unique instances of variation in market structures, which allow to study how market power affects surplus distribution. In the second chapter I analyze the effect of amendments on merger regulation in several countries, showing that such changes in antitrust policy resulted in stealth consolidation, increasing the number of potentially anticompetitive transactions. Furthermore, I show that such policy changes increased industry level concentration, decreased investment, decreased labor shares in

affected industries, and consequently increased income inequality in affected countries.

In the third chapter I answer a similar research question, but from a macroeconomic perspective. I use stealth consolidation in a dynamic factor model to identify exogenous variations in market power and their effect on the economy, a novel methodology that allows to overcome limitations in the data. Results show that the identified market power shock lowers output, but it increases the share of output that goes into profits. Moreover, it increases income and labor earnings inequality on impact, and this is mainly due to an earnings loss for the poor. The identified shock accounted for an increase in income Gini index by 0.4 between 2001 and 2006, and it can account for 20% of the variation in inequality. Therefore, this chapter provides evidence of a causal link between market power and income inequality.

Contents

Acknowledgements

thank you!

Introduction

Chapter 1: Antitrust Policy and Innovation

I study how antitrust authorities encourage innovation of merging firms by fostering competition. A change in merger notification rules allows me to build an event study comparing mergers notified to the authorities with non-notified ones. I develop a new text analysis methodology to identify mergers between close competitors, even for small private firms. For this exercise, a natural language processing model is trained on the corpus of US published patents. As a result of the event study, non-notified horizontal mergers lead to 30% less innovation effort, measured as patenting activity. To understand the mechanism driving these findings, I build a model with endogenous merger choice where optimal antitrust policy deters anticompetitive mergers, which are also most detrimental to innovation. This implies that authorities deter mergers that would be harmful to innovation. An increase in the number of non-notified anticompetitive mergers is consistent with the deterrence effect of antitrust authorities.

Chapter 2: A Cross Country Analysis of Stealth Consolidation and its effects on Inequality

Antitrust policy can influence the balance of power between producers and consumers, and thus the redistribution of resources between them. Consequently, changes in antitrust policy provide unique instances of variation in market structures, which allow to study how market power affects surplus distribution. This work analyzes the effect of amendments on merger regulation in several countries, showing that such changes in antitrust policy resulted in stealth consolidation, increasing the number of potentially anticompetitive transactions. Furthermore, this work shows that such policy changes increased industry level concentration, decreased investment, decreased labor shares in affected industries, and consequently increased income inequality in affected countries.

Chapter 3: Stealth Consolidation, Market Power and Income Inequality

Market power allows firms to capture a larger share of society surplus and to concentrate it in the hands of few. However, there is scant evidence on the relationship between market power and income inequality. This paper uses stealth consolidation in a dynamic factor model to identify exogenous variations in market power and their effect on the economy, a novel methodology that allows to overcome limitations in the data. Results show that the identified market power shock lowers output, but it increases the share of output that goes into profits. Moreover, it increases income and labor earnings inequality on impact, and this is mainly due to an earnings loss for the poor. The identified shock accounted for an increase in income Gini index by 0.4 between 2001 and 2006, and it can account for 20% of the variation in inequality. Therefore, this paper provides evidence of a causal link between market power and income inequality.

Chapter 1

JMP: Antitrust Policy and Innovation

1.1 Introduction

Competition for costumers and market shares can be a strong incentive for investments and innovation. At the same time, profits earned through market power can be essential for funding new investments and developing new products. As it stands, the relationship between competition and innovation is hotly debated in economics. This is not only academically significant, but it is also policy relevant for competition authorities, who have directed their attention to innovation in the recent years.[1][2] While there exists a wide literature on the effects of the level of competition on innovation, the literature on

[1]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... We need to understand why so many industries have too few competitors, and to think carefully about how to ensure our merger enforcement tools are fit for purpose in the modern economy."

[2]Commissioner Margrethe Vestager of Directorate General Compeptition on the 2018 Bayer-Monsanto merger, from the official European Commission Website: "...Our decision ensures that there will be effective competition and innovation in seeds, pesticides and digital agriculture markets also after this merger. ...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."

mergers and innovation has only developed in recent years.[3] This despite mergers providing a sudden change in competition and innovation incentives. Moreover, mergers and acquisition (hereafter M&A) activity is at an all time high, while large scale acquisition of innovative startups is concerning regulators.

This paper asks whether antitrust policy can encourage innovation by fostering competition. In particular, I study the effect of antitrust policy on the innovation activity of merging firms. This question poses three major challenges. First, antitrust authorities have jurisdiction over every merger. Thus, it appears to be impossible to build a reliable counterfactual of M&As that are not scrutinized by the authorities. In reality, though, only mergers that satisfy certain requirements are notified to the authorities, who have no time or resources to inspect most of the non-notified transactions. Therefore, I construct an event study exploiting a change in these notification rules to identify a suitable counterfactual. Second, non-notified mergers involve small and private firms for which it is particularly challenging to gather reliable information on products and production processes. As a solution, patents are used to measure their innovation efforts. Moreover, I compare patent abstracts using a natural language processing algorithm to identify close competitors of the merging firms. Third, the mechanism driving my results has to account for antitrust authorities effectively blocking very few mergers. Consequently, I build a model of merger choice describing the deterrence effect of antitrust authorities. Predictions of this model are in line with the data.

The first contribution of this paper is to provide a reliable identification strategy to study how antitrust policy influences innovation through its effect on competition. In the US, mergers have to be notified to the authorities if merging parties are large enough.[4]

[3]One can look at Vives (2008) for a review of the literature on competition and innovation. Recent papers in the literature on mergers and innovation include Motta and Tarantino (2017), Mermelstein et al. (2018), Denicolò and Polo (2021) and Haucap et al. (2019).

[4]Notification rules in US are based on the size of the transaction, which correspond to the amount of

As the source of time variation I exploit a 2000 change in notification rules that made hundreds of mergers exempt from notification. Wollmann (2019) shows that this policy change lead to a 70% decline in pre-merger notifications. As a first layer of cross-sectional variation I use the difference between mergers that are affected by the policy change and mergers that are not affected. Since mergers that became exempt from notification involve small and private firms, this methodology requires data on this kind of companies. Wollmann (2019) defines these M&A as Stealth Consolidation, to emphasize their nature of stealthy and hidden transactions.[5] Therefore, I use a database in which mergers collected from conventional channels are combined with mergers announced on industry journals, news outlets and other publications.[6]

As a second layer of cross-sectional variation, I exploit the difference between horizontal and non horizontal mergers. Horizontal transactions involve close competitors that operate in the same product markets, and they attracted most of the attention of antitrust authorities in the early 2000s. Just to emphasize how important horizontal mergers were for the authorities, in the 2000s the official US merger guidelines were called *Horizontal Merger Guidelines*. Therefore, my treated group consists of firms involved in horizontal mergers that were made exempt from notification after the policy change. The control group, on the other hand, includes firms involved in non horizontal mergers or mergers that were notified to the authorities.

Antitrust authorities spend a large portion of their time and resources to discern which mergers are horizontal. This is particularly challenging for the small and private firms that are the focus of this study, since there is very little information available on their products and production process. To overcome this challenge, my second contribution is

money that is paid, and the size of the target company that is acquired, in terms of assets and sales.

[5]In a further paper Wollmann (2020) shows that this policy change generated concentration in the dialysis industry and it caused a decline in survival rate of treated patients.

[6]I use Thomson Reuters SDC Platinum database on mergers and acquisitions for the US.

to use the information contained in patents published by the merging parties. In particular, I train a natural language processing model based on word embeddings on the whole corpus of US patents.[7][8] Then, horizontal mergers are identified based on how similar are the patents of the merging parties. Antitrust authorities have access to internal documents to ascertain if two firms operate in the same product market. I approximate their decision using information on patents used in firms' production process. As a validation of my methodology, this definition of horizontal mergers based on patent similarity can match the European Commission and the Federal Trade Commission classification of public merger cases reasonably well.

I measure innovation activity using patent citations weighted by average citations in the same technology field.[9] In my sample mergers lead to an average decrease in innovation activity of about 30%. The results of my difference in differences identification strategy imply that horizontal mergers that are not reported to the authorities lead to a further 30% less innovation in the years following the policy change. The effect is stronger in the short run after the event. This result is consistent with the hurried realization of several mergers that were deterred by the authorities. Moreover, the decrease in innovation is characterized by a decline of quality and originality of published patents.[10] Furthermore, a decrease in process innovation drives the results. This implies that affected firms are becoming less productive after the policy change. In addition, I show that the number of horizontal mergers that are not reported to the authorities increases after the policy

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

[8] I train a Doc2Vec model, which is an extension of the popular Word2Vec word embeddings model, on the title and abstract of each patent.

[9] The 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.

[10] 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 then to cite a less diverse array of other patents, which is considered a drop in patent originality by Lerner et al. (2011). I find no significant effect on patent generality, which is measured as diversity of citing patents.

change, in accordance with deterrence being the mechanism behind the main results.[11]

The main results of this paper are robust to a variety of specification changes. Since a large portion of the sample is comprised by 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 similarity between firms as identification device, and this implies results that are even stronger. On top of that, remedies imposed by the authorities as condition for merger approval might explain my results. I discuss evidence on the meager effectiveness of remedies in preventing anticompetitive effects on prices, implying that they cannot provide a convincing explanation for the large effects found in this paper.

The third contribution of this paper is a model of deterrence to explain the mechanism behind my results. Very few blocked mergers are not sufficient to explain large effects on innovation. Therefore, I build a model with endogenous merger choice determined by firm profit maximization and an optimal antitrust policy derived from consumer harm minimization. Through deterrence, mergers that are more detrimental to consumers have a lower probability of being proposed because they have a lower chance of being accepted. To relate this deterrence model to my innovation results, I build it on top of a state of the art model of competition in prices and cost reducing innovation a là Motta and Tarantino (2017).[12] In this class of models mergers that lead to lower consumer surplus are also merger that generate less efficiencies and less innovation, and these are the most susceptible to deterrence.

[11]This result was already discussed extensively by Wollmann (2019) in his first paper on Stealth Consolidation.

[12]The choice of a model of cost reducing innovation is justified by the results on process innovation, which is the kind of innovation that increases productivity.

The model has several predictions for changes in the level of scrutiny that mergers are subject to. First, it implies an increase in the number of horizontal mergers, which is precisely what I observe in the data. Second, it predicts a decrease in innovation activity if antitrust policy is more lenient. This correspond to the main results of this work: affected mergers lead to less innovation. The final implication of the model is that these mergers should decrease consumer surplus, which is indeed what the antitrust authorities seek to prevent.

Related Literature

This paper contributes to the wide literature of competition and innovation. There are several works that 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). The present paper instead focuses on abrupt changes in competition generated by mergers and acquisitions. 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. Among the work that study this issue empirically, Haucap et al. (2019) find that mergers lead to less innovation by comparing merging parties with similar non-merging firms chosen with a matching procedure. My paper proposes a different identification strategy based on a policy change that affected merger incentives of small and innovating firms. 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. Incidentally, they 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)). Wollmann (2019) started the literature on Stealth Consolidation by studying merger notification rules. The theoretical literature on antitrust policy has recently explored also the innovation activity of merging parties. For instance, 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. They find that no antitrust scrutiny is never the optimal, while the optimal policy blocks most of the proposed mergers. How the authorities affect innovation activity of merging firms is of interest even for the management literature, as Thatchenkery and Katila (2021) show.

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. (2013b) and Mikolov et al. (2013a). I apply this methodology to patents and I use it to identify horizontal mergers. 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. Their methodology can be applied only to large public firms available in COMPUSTAT, while I devise a new methodology based on patents that can be applied to small and private firms. Younge and Kuhn (2015) describe a vector space model to compute patent similarity using text analysis methodologies. They use word counting and weighting, while I use more modern semantic techniques such as word embeddings.[13]

[13]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

In my work I use patents as a measure of innovation. Lerner et al. (2011) use patents to study the effect of private equity leveraged buyouts on the innovation activity of acquired firms, and I use several of their innovation measures. Ganglmair et al. (2022) describe an innovative algorithm to classify claims and patents as process or product innovation. Arts et al. (2021) develop natural language processing to identify patents that generate radically new technologies with a major impact on technological progress. Nevertheless, there is still an open debate on the use of patents and patent citations as measure of innovation. Kuhn et al. (2020) and Kuhn and Younge (2019) show that citation patterns have changed significantly in recent years, possibly affecting existing results. What the authors find mostly applies to recent years, but not to my sample of years around the policy change in 2000. Moreover, patents can be filed for strategical reasons, as a large literature has shown (Levin et al. (1987), Hall and Ziedonis (2001), Rysman and Simcoe (2008), Kang and Motohashi (2015), Lerner and Tirole (2015), Righi and Simcoe (2020))

The present paper also contributes to the literature on deterrence effects of antitrust policy. Besanko and Spulber (1989) build a model of antitrust enforcement under asymmetric information. Similarly to my framework, they model antitrust policy as a probability of enforcement which depends on observed market outcomes.[14] More recently, Miller (2009) studies leniency toward early confessors of cartel behavior and he finds significant evidence of authorities deterrence. The author uses time series techniques to identify a causal effect, while in my work I identify a suitable counterfactual for a difference in differences analysis. 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. Despite the available evidence, deterrence capabilities of antitrust

intensive. In my work I use more modern techniques based on word embeddings, which are both efficient and more effective at representing semantic meaning.

[14]The authors also find that it is optimal to tolerate some collusion from the most efficient firms.

authorities are a contentious issue in the literature. Eckbo (1992) finds no evidence of deterrence comparing US mergers to Canada mergers, exploiting more lenient antitrust scrutiny in Canada. 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. Deterrence is a relevant issue also for the antitrust law literature and it is covered in several works such as Breit and Elzinga (1973), Baker (1988), Wils (2006), Lande and Davis (2011).

The rest of this paper is organized as follows. Section 1.2 presents data and variables used in the analysis. Section 1.3 describes the empirical methodology of the event study. Section 1.4 proposes the main results. Section 1.5 discusses features and limitations of the main analysis, and provides robustness checks and sensitivity analysis. Section 1.6 describes the model of deterrence which provides a mechanism explaining the results. Section 1.7 concludes.

1.2 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.

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.[15] For both firms I can

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

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.[16] 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.

I accessed data on the universe of patents published by the U.S. Patent & Trademark Office (USPTO) through PatentsView. [17] 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.[18]

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

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

[17]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.

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

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.

1.3 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.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* .[19] Horizontal mergers are defined as transactions involving firms operating in the same product markets, firms that are close competitors. 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

[19] See Wollmann (2019) or Wollmann (2020) for more discussion on this matter.

mergers, especially given the limited amount of data that is available for private firms?

1.3.1 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. If two firms have very similar patents, what they produce based on these patents is likely to be similar as well. 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.[20] The use of patent abstracts allows me to extend this methodology to private firms, which are not covered in conventional data-sets and for which information is quite scarce. Patent are carefully

[20] 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.

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 try to 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.[21]

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. (2013b) and Mikolov et al. (2013a). I describe this algorithm in more details in Appendix A. 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 1.1 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

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

one, and it is equal to one if the two vectors are exactly the same.[22]

$$CS_{ik} := \frac{P_i \cdot P_k}{||P_i||\ ||P_k||} = \frac{\sum_j P_{ij} P_{kj}}{\sqrt{\sum_j P_{ij}^2}\ \sqrt{\sum_j P_{kj}^2}} \leq 1 \tag{1.1}$$

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.[23] 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 following.[24]

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]

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]

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 Pharmacia are major pharmaceutical companies. At first, it might seem

[22] I f 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.

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

[24] The identifying codes can be used on lens.org to look for the patent. There one can find the whole abstract, as well as more information on the patents themselves. I encourage the reader to do so, just to have a better feeling of these patents.

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 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.[25] 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.

[25] 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.

Some examples are the maximum similarity (hereafter Max), the mean of 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.[26] 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.

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. 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.[27] I consider a decision 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.[28] 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

[26]Table 6 on Appendix A shows the correlation table of these variables.

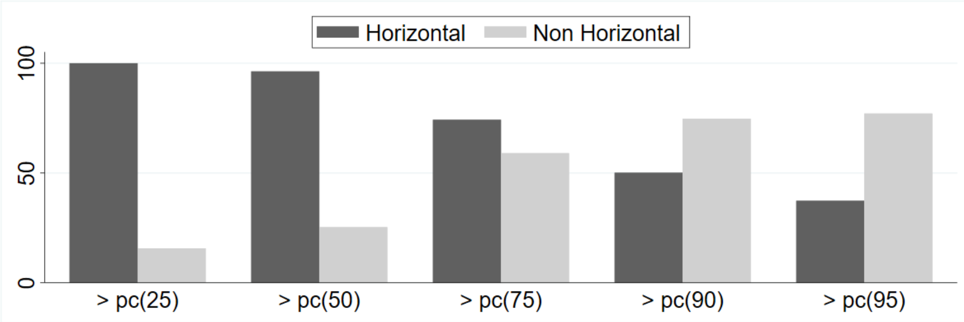
[27]This includes all cases settled in the first phase of an investigation (Art. 6(1)(a), 6(1)(b), 6(1)(c) and 6(2)) and all cases decided in the second phase of an investigation (Art. 8(1), 8(2), and 8(3)). Note that this also includes all cases settled under a 'simplified procedure', provided that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

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

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 6 in Appendix A, 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 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 1.1 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.



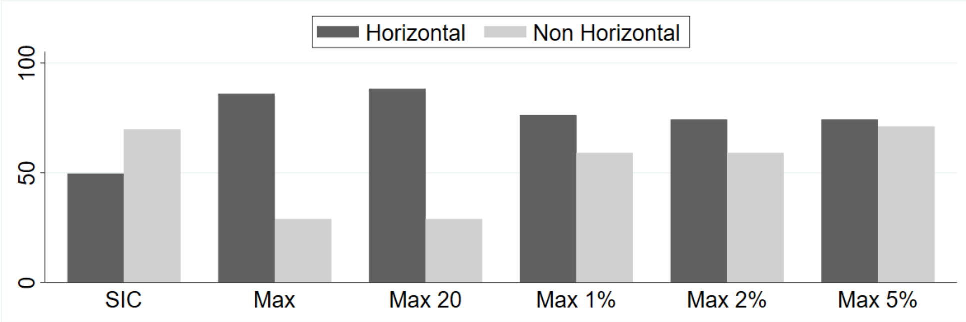
Notes: Percentage of correctly predicted horizontal (dark grey) and non horizontal (light grey) mergers, as defined by the EC, with increasing cutoff. The cutoff is in percentiles of the statistic distribution. The exercise is performed on the max 2% statistic, and similar patterns emerge with other statistics. Numbers for the histogram are reported in Table 7 in Appendix A

Figure 1.1: Performance of different cutoff rules on EC Decisions.

Once the cutoff rule is set to the 75^{th} percentile, I compare similarity statistics in Figure 1.2. 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 6 in Appendix A, 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.

1.3.2 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, merger guidelines set a threshold under which merging parties are exempt from reporting to the authorities. They can merge, and the Antitrust Authority



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 8 in Appendix A.

Figure 1.2: Performance of Similarity Statistics on EC Decisions.

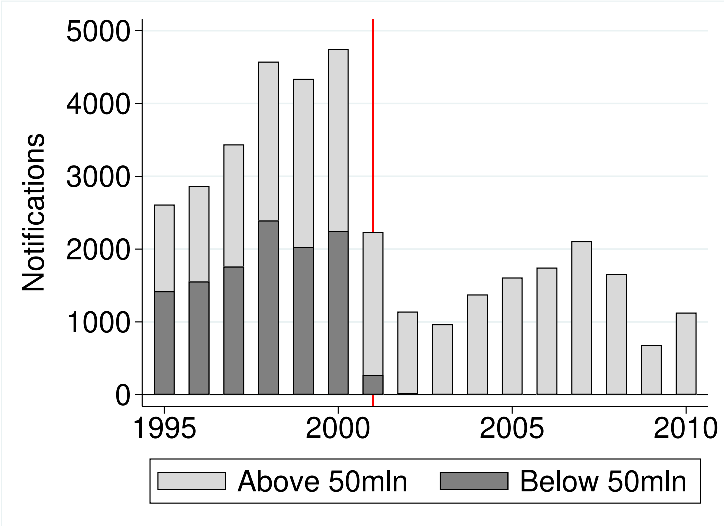
will never know about their transaction. If the Authority is informed of the transaction by other actors it can investigate the merger, but this is a very rare occurrence, given the already burdensome amount of work given by other mergers. As a matter of fact, one can consider these exempt mergers as transactions that are not controlled by the authorities.

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.[29] 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.[30] 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.[31] Figure 1.3

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

[30]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.

[31]For a more in-depth description of the Amendment to the Hart-Scott-Rodino Act one can refer to



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.

Figure 1.3: Number of Notifications received by US Antitrust Authorities.

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. Figure 16 in the Appendix .4 shows that Second Requests issued by the authorities also fell considerably after the policy change, indicating that antitrust enforcement decreased substantially as well.[32]

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 Amendment 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

Wollmann (2019), who was the first to study this policy change.

[32]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.

was too low, as it was not adjusted to inflation.[33] 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.[34] 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. 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 quite difficult, though, as the acquired company is likely to refuse the deal. Moreover, in the data there is no evidence of bunching below the new threshold of $50 million.[35]

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 exempt from reporting before the amendment, but they became exempt thanks to the Amendment.[36] 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.[37] By definition, these mergers are not affected by the Amendment due to their larger size.

[33]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`

[34]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`

[35]Figure 17 in Appendix .4 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.

[36]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.

[37]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.

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.[38].

1.3.3 Event Study

The unit of analysis of my event study 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 $\Delta P_{it} = ln\left(\sum_{i=1}^{n} P_{t+i}/n\right) - ln\left(\sum_{i=1}^{n} P_{t-i}/n\right)$, 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.[39] Patenting activity of merging firms is then the average of this relative citation measure.

Table 1.1 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. 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 innova-

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

[39] 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.

	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 event study. 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.

Table 1.1: Unconditional Innovation Change generated by Mergers.

tion, 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 event study.

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 1.3.2.

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 1.2. I^{Post} is a dummy variable that is equal to 1 if the merger date is after the amendment, meaning years after 2001.[40] I^{Ex} is equal to 1 for Newly-Exempt mergers, and it is equal to 0 for Never-Exempt mergers.[41] I^{Hor} is a dummy variable representing horizontal mergers, as defined in Section 1.3.1. 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.[42]

$$\begin{aligned}\Delta P_{it} = &\beta I^{Post} I^{Ex} I^{Hor} + \gamma I^{Post} I^{Ex} + \theta I^{Post} I^{Hor} + \delta_1 I^{Ex} I^{Hor} + \delta_2 I^{Ex} + \delta_3 I^{Hor} \\ &+ \alpha_t + \xi X_{it} + \epsilon_{it}\end{aligned} \tag{1.2}$$

[40] In the robustness section I show that results hold if I consider years after 2000 as the post period.
[41] As previously explained, Always-Exempt mergers (the smallest ones) are excluded from the analysis.
[42] As additional controls I include SIC 2 digit industry fixed effects, role fixed effect (Acquiror or Target), and State fixed effects.

1.4 Results

Table 1.2 reports results of the triple difference in difference described by equation 1.2. From column (1) to column (3) I add Industry and State fixed effects. Column (3) 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 1.1. 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.

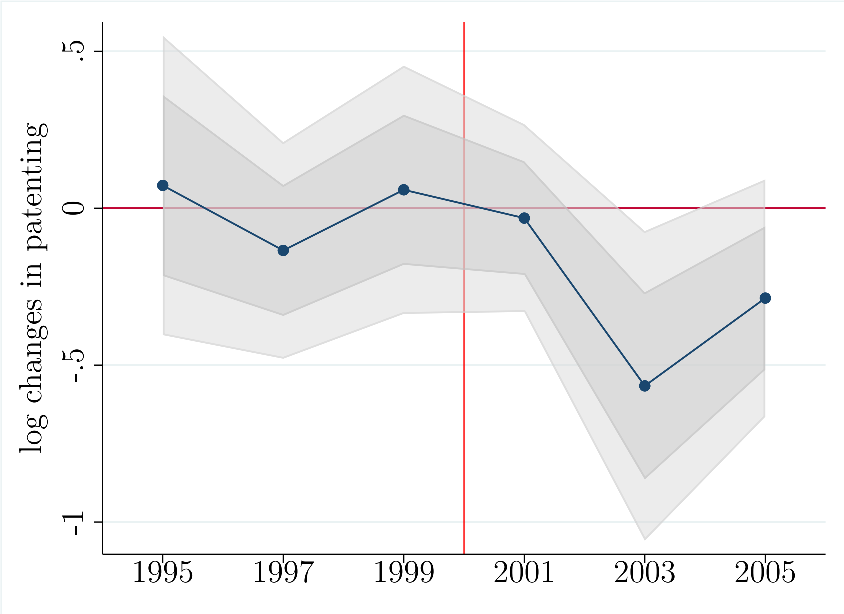
VARIABLES	(1) Avg	(2) Avg	(3) Avg	(4) Avg
$I^{Post} \cdot I^{Ex} \cdot I^{Hor}$	-0.262*	-0.299**	-0.357**	-0.357**
	(0.148)	(0.148)	(0.152)	(0.175)
Observations	2,601	2,601	2,601	2,601
R-squared	0.020	0.054	0.080	0.080
Year FE	YES	YES	YES	YES
Industry FE	.	YES	YES	YES
State FE	.	.	YES	YES
Cluster SE	SIC4	SIC4	SIC4	.

Notes: This table reports coefficients of equation 1.2 with various control specifications. The main specification is reported in column (3). Column (4) shows the main specification without clustering standard errors.

Table 1.2: Triple Difference in Differences Results

Rather than computing average effect using all years in the sample, Figure 1.4 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 18 in Appendix .5 that shows coefficients in every single year. Figure 1.5 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 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 inspect the dynamics in years following every single merger. By inspecting single period changes computed as $\Delta P_{t+i} = ln\,(P_{t+i}) - ln\,(P_{t-1})$, one can see if the effect is stronger in the long or in the short run after the merger. Figure 21 in Appendix .5 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.



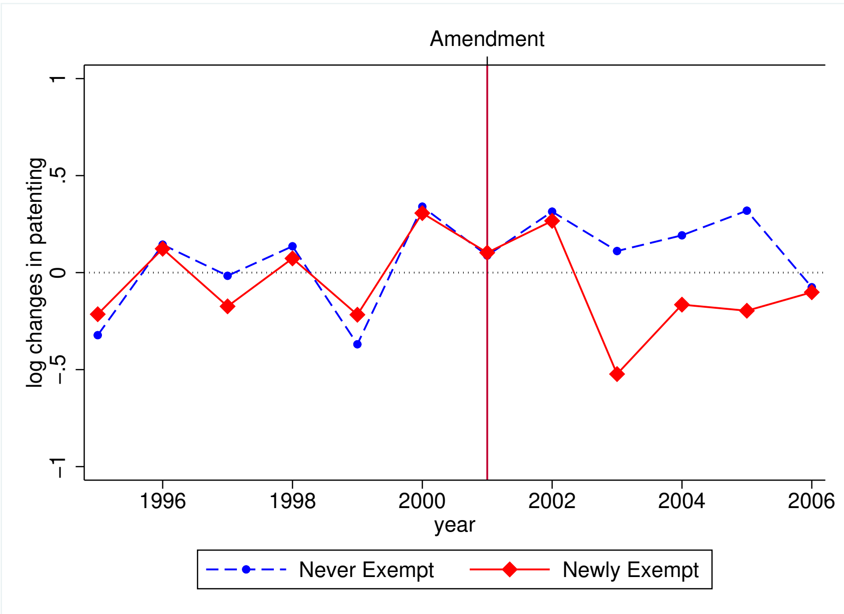
Notes: Coefficients of triple diff-in-diff term $I^{Ex}I^{Hor}$ interacted with two year periods around the Amendment. Appendix .5 reports the same figure for single years and three or four year periods. Results have the same qualitative implications.

Figure 1.4: Coefficients of triple Diff in Diff for years around the Amendment.

1.4.1 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 1.6 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.

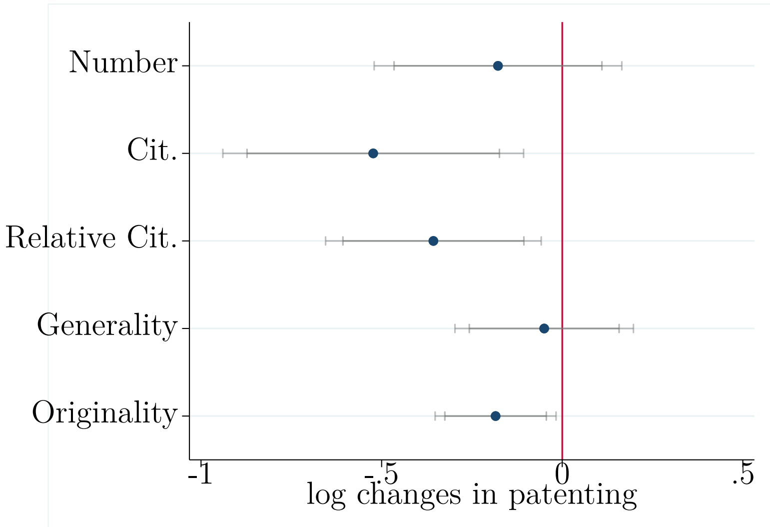
Innovation quality can be measured is several ways. Following Lerner et al. (2011) construct two measures using citations and technological fields. First, I compute Generality of a patent as the dispersion of citing patents among technological fields. This is calculated as $(1 - HHI_{citing})$ where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



Notes: The data points of the figure are constructed by averaging the innovation effect of mergers in various groups. Data are residualized on fixed effects used in column (3) of Table 1.2. Each point reports the difference between horizontal and non horizontal mergers.

Figure 1.5: Time series of the difference in innovation between Horizontal and Non Horizontal Mergers

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 1.6 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.



Notes: Coefficients of triple diff-in-diff equation 1.2 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.

Figure 1.6: Triple difference in differences results for various innovation activity measures.

1.4.2 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). 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, 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.[43]

In order to assess which kind of innovation is most affected by the policy change, I repeat the main analysis described by equation 1.2 but considering only patents classified as one kind of innovation. Table 1.3 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

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

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)
Citations			
$I^{Post} \cdot I^{Ex} \cdot I^{Hor}$	-0.523**	-0.499*	-0.358
	(0.212)	(0.301)	(0.229)
Observations	2,601	1,599	2,242
Year FE	YES	YES	YES
Industry FE	YES	YES	YES
State FE	YES	YES	YES
Cluster SE	SIC4	SIC4	SIC4

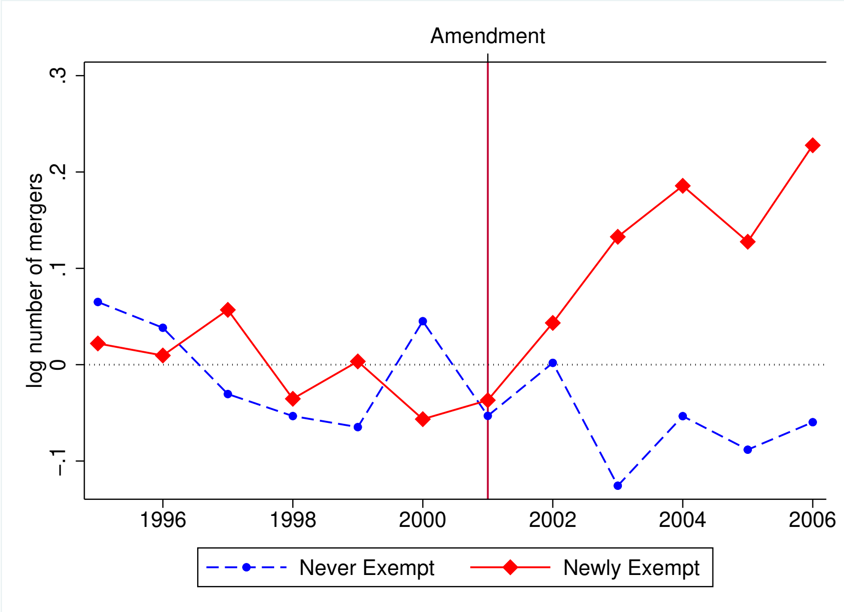
Notes: This table reports coefficients of equation 1.2 applied on process or product innovation. Column (1) shows baseline results. Column (2) shows results computed considering only process patents. Column (3) reports results computed considering only product patents.

Table 1.3: Triple Difference in Differences Results for Product and Process Innovation

to increase their productivity.

1.4.3 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 defines this merger wave as Stealth Consolidation, a series of potentially anticompetitive mergers that escape antitrust scrutiny. In Figure 1.7 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 1.7 shows that they increase by about



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.

Figure 1.7: Time series of the number of mergers

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 12 in Appendix .5 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.

1.5 Discussion and Robustness

A feature of the analysis that is worth emphasizing is the size of the sample. Table 14 in Appendix .6 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.

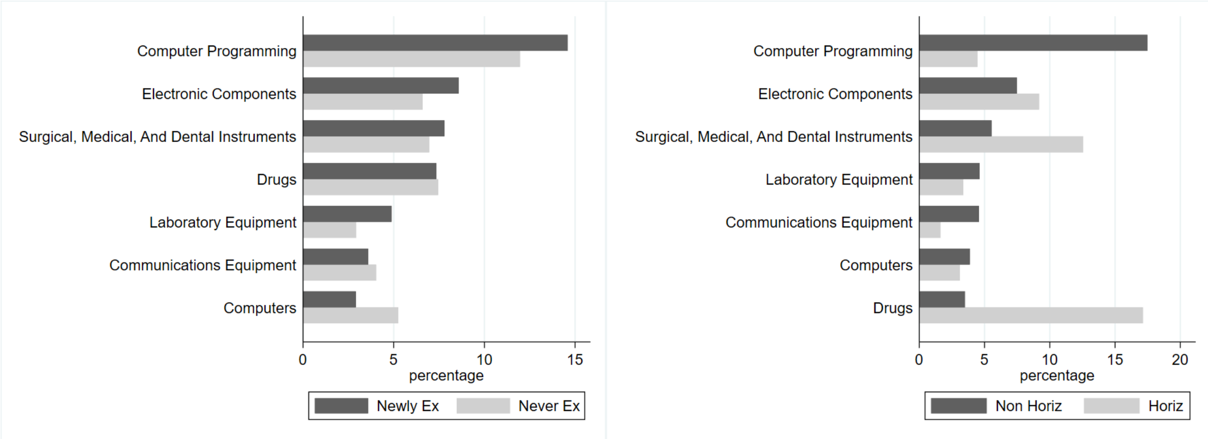
Given the definition of horizontal mergers in Section 1.3.1, the number of non horizontal mergers is much higher than the number of non horizontal ones. Moreover, the number of never exempt horizontal mergers is much lower than the number of newly exempt ones. This is due to two reasons. First, most of the mergers are below the new threshold introduced by the amendment.[44] 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.[45] In panel B and C of Table 14 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.

1.5.1 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 1.8 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 components" and "communication equipment". The second is the Pharma industry, which comprises "drugs", "Surgical, medical and dental equipment" and "laboratory equipment". Figure 1.8a shows that the sector distribution for Newly and Never-Exempt mergers is quite similar. Conversely, Figure 1.8b implies that horizontal mergers are more common in the Pharma industry, while non horizontal

[44] One can see this also in the Figure 17, where it is clear that there are many more mergers below $50 million than above this threshold.

[45] The conditions are summarized in the "size of person test" and the "size of transaction test" described in Section 1.3.2



(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.

Figure 1.8: Sector composition differences between merger categories

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.[46]

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 1.4 shows results of the leave-one-out exercise, and it is clear that results are not affected by removing sectors. Column (4) and

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

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)
$I^{Post} \cdot I^{Ex}$	0.341***	0.294***	0.358**	0.358***	0.352***
	(0.116)	(0.113)	(0.147)	(0.118)	(0.119)
$I^{Post} \cdot I^{Hor}$	0.188	0.157	0.221	0.221*	0.196*
	(0.117)	(0.121)	(0.141)	(0.118)	(0.118)
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 1.2 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".

Table 1.4: Triple difference in differences leaving out particular sectors

(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.

1.5.2 Definition of Horizontal Mergers

The identification of horizontal mergers presented in Section 1.3.1 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 15 to 18 in Appendix .6 show coefficient of the main equation 1.2 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 1.2.

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 1.3 shows that similarity measures Sim can be put in place of the horizontal dummy I^{Hor} for an alternative identification strategy. Table 1.5 shows results of this identification strategy. The panel of relative citation average shows an average decrease in innovation of bout 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 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.

$$\begin{aligned}\Delta P_{it} =& \beta I^{Post} I^{Ex} Sim + \gamma I^{Post} I^{Ex} + \theta I^{Post} Sim + \delta_1 I^{Ex} Sim + \delta_2 I^{Ex} + \delta_3 Sim \\ &+ \alpha_t + \xi X_{it} + \epsilon_{it}\end{aligned} \tag{1.3}$$

1.5.3 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.[47] 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

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

measure of the innovation activity of a company. The triple difference in differences identification strategy explained in Section 1.3 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. Moreover, the results of this paper stand even with measures of innovation that do not rely on patent citations.

1.5.4 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.[48] 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 very limited effects on merger outcome such as prices. In his resent 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

[48] 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 1.9.

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, which 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, while structural remedies and divestitures are pursued less. Therefore, one can conclude that remedies are ineffective at preventing anti-competitive outcomes, and thus they should be also ineffective at preventing detrimental effects on innovation.

1.5.5 Amendment

Some more robustness checks can be constructed based on specifics of the Amendment. First, I test whether the choice of 2001 as year of the amendment affects significantly the results. Since 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 19 in Appendix .6 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 21 in Appendix .6 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.[49]

[49]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.

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 20 in Appendix .6 shows that results are unaffected by this change, the coefficient is still negative and significant and of similar size to the main results.

1.6 Model of Deterrence

To explain the deterrence effect of Antitrust Authorities I build a simple 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 a là 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, and for simplicity of exposition I assume that both have perfect information.[50] 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 an imperfect ability to block the merger, and it exerts an effort to influence the probability that the

[50]The assumption of perfect information of the antitrust authority can be relaxed without changing the implications of the model.

judge will allow the merger $\alpha \in [0,1]$. This is effectively the antitrust policy, which is known to the merging parties.[51]

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 will start to solve the model from period $t = 2$ and then move backwards toward period $t = 0$.

1.6.1 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.[52]. Before the merger, or in case the merger does not realize, each merging firm gains profit π_b as in equation 1.4. In line with Motta and Tarantino (2017) I assume symmetric firms.[53] After a successful merger, each firm internalizes the profit of its competitor in its maximization and earns π_M described in equation 1.5. The merger has a potential to generate efficiencies $\lambda G(x_i, x_k)$ that reduce the fixed cost of innovation, which satisfy $F(x_i) + F(x_k) - \lambda G(x_i, x_k) \geq 0$ to ensure no negative costs. $\lambda \geq 0$ is a scalar that determines the size of merger efficiencies. With $\lambda = 0$ there are no efficiencies and the merger is most anticompetitive.

[51] 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.

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

[53] Motta and Tarantino (2017) show that for the general case of asymmetric firms the qualitative outcomes of the model do not change.

$$\pi_b = \max_{p_i, x_i} (p_i - c(x_i))q_i(p_i, \bar{p}_{-i}) - F(x_i) \tag{1.4}$$

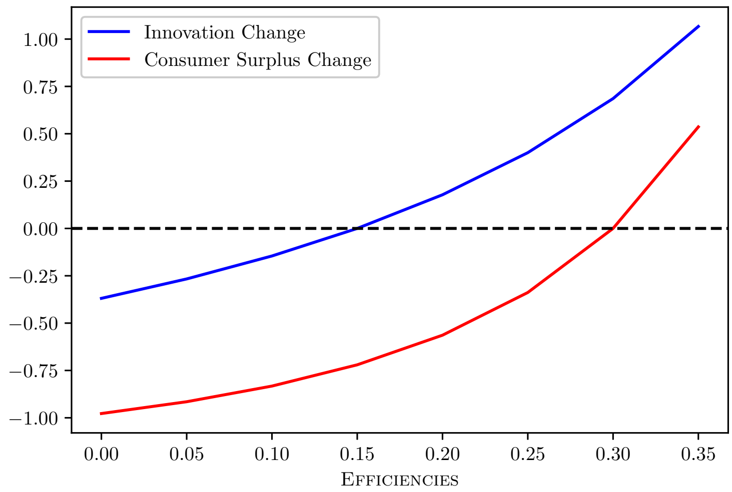
$$\begin{aligned} 2\pi_M = \max_{p_i, x_i, p_k, x_k} & (p_i - c(x_i))q_i(p_i, \bar{p}_{-i}) + (p_k - c(x_k))q_k(p_k, \bar{p}_{-k}) \\ & - F(x_i) - F(x_k) + \lambda G(x_i, x_k) \end{aligned} \tag{1.5}$$

What then is the effect of a merger for varying levels of efficiencies λ? Figure 1.9 shows changes (after the merger minus before the merger) of consumer surplus and innovation depending on the level of efficiencies.[54] 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.

1.6.2 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

[54] Appendix A shows how to solve these kind of models in closed form.



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 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 B. Marginal cost is linear in innovation $c(x_i) = 1 - x_i$. Fixed costs and efficiencies are quadratic $F(x_i) = \frac{1}{2}x_i^2, \quad G(x_i, x_k) = \frac{1}{2}x_i x_k = \frac{1}{2}x_i^2$. 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 1.9: Changes in Consumer Surplus and Innovation after a merger

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 (1.6) describes the resulting problem of the manager, who chooses optimally φ.

$$\varphi^*(\alpha^*) = \underset{\varphi}{\text{argmax}} \quad \hat{\pi}\alpha^*\varphi - \Gamma(\varphi) \tag{1.6}$$

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.

1.6.3 Antitrust Policy (t=0)

If an antitrust authority wants to bock a merger in the US, it has to challenge the merger 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. In Appendix D I relax this assumption and show that it does not affect the results. Similarly to the firms, authorities care about expected consumer surplus. Therefore, their expected payoff is $\hat{CS}\alpha^*\varphi(\alpha^*)$, where $\hat{CS} = CS_M - CS_b$ 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 (1.7) shows the problem of the antitrust authority.

$$\alpha^* = \underset{\alpha}{\operatorname{argmax}} \;\; \hat{CS}\alpha\varphi^*(\alpha) - \Phi(\alpha) \tag{1.7}$$

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 $\hat{CS}$. In Appendix C I give a simple functional form to Γ and Φ deriving a closed form solution for the optimal antitrust policy.

Lemma 1. *Mergers with lower $\hat{CS}$ have a lower chance of being approved α, and through deterrence a lower chance of being proposed φ.*

1.6.4 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.*

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 1.9. 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. One last result of this model is that these same mergers should lead to less consumer surplus, as per Corollary 1. This prediction I cannot verify with the available data, but it warrants future empirical analysis of the price effects of these mergers.

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.*

Corollary 1. *The average consumer surplus change generated by realized mergers is lower when mergers become non-notifiable.*

1.7 Conclusion

In this paper I ask whether antitrust authorities can stimulate innovation by promoting competition. In particular, I study the effect of antitrust policy on the innovation activity of merging firms. To examine this issue I exploit a unique policy change that made hundreds of small M&A exempt from notifying to the authorities. This amendment to existing regulation was so dramatic that the number of pre-merger notifications fell by as much as 70%. This allows me to build a reliable counterfactual of mergers that are not subject to antitrust policy. Given the stealthy and covert nature of these transactions between small and private firms, 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. Then I focus on horizontal mergers, transactions between close competitors that are the most likely to be anticompetitive and to attract the attention of the antitrust authorities. To identify horizontal transactions in my unconventional data-set, I employ a natural language processing algorithm to compare the abstract of patents published by the merging parties. I train a word embeddings machine learning model on the whole corpus of US patents, to account for the specificity of patent jargon. I show that my algorithm performs better than standard industry classification at predicting EC and FTC classification of horizontal mergers.

In my sample mergers lead to an average innovation reduction of about 30%. The results of my difference in differences exercise indicate that horizontal mergers that are not reported to the authorities imply a further 30% less innovation. Moreover, merging firms become less productive after the policy change. Lastly, I find that the number of unreported horizontal mergers increase, in accordance with deterrence being the mechanism behind my results. To explore this avenue, I build a model of deterrence with endogenous merger choice and an optimal antitrust policy. In this model mergers that are more detrimental to consumers are also mergers that lead to decreasing innovation, while 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. A further result of the model is that these mergers will lead to a decrease in consumer surplus.

The policy implications of these results are that the policymaker should not dismiss small mergers as negligible for competition and innovation. Quite on the contrary, a large number of these transactions can have a profound impact in several product markets. Therefore, it is worth to extend antitrust scrutiny even below the currently high notification thresholds. The FTC already cited stealth consolidation when it issued special orders compelling big tech to disclose previously non-reportable deals.[55] 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.[56]

This paper can be the foundation for several future research avenues. Using the measure of firm similarity based on text analysis of patents, one could create a network of competitive relations and demand elasticity 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 my measure would be to encompass also small and private firms, and not only large and public ones. This would allow me 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

[55] "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.

[56] "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

empirical evidence on consumer surplus and welfare more in general.

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 1.3 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}.

Table 1.5: Triple difference in differences results computed using continuous patent similarity.

Bibliography

Acemoglu, D. and Akcigit, U. (2012). INTELLECTUAL PROPERTY RIGHTS POLICY, COMPETITION AND INNOVATION. *Journal of the European Economic Association*, 10(1):1–42.

Affeldt, P., Duso, T., and Szücs, F. (2021). 25 years of European merger control. *International Journal of Industrial Organization*, 76:102720.

Aghion, P., Bloom, N., Blundell, R., Griffith, R., and Howitt, P. (2005). COMPETITION AND INNOVATION: AN INVERTED-U RELATIONSHIP. *QUARTERLY JOURNAL OF ECONOMICS*, page 28.

Arts, S., Hou, J., and Gomez, J. C. (2021). Natural language processing to identify the creation and impact of new technologies in patent text: Code, data, and new measures. *Research Policy*, 50(2):104144.

Ash, E., Chen, D. L., and Naidu, S. (2022). Ideas Have Consequences: The Impact of Law and Economics on American Justice.

Baker, J. B. (1988). Private Information and the Deterrent Effect of Antitrust Damage Remedies. *Journal of Law, Economics & Organization*, 4(2):385–408.

Barrios, J. M. and Wollmann, T. G. (2022). A New Era of Midnight Mergers: Antitrust Risk and Investor Disclosures. page 58.

Besanko, D. and Spulber, D. F. (1989). Antitrust Enforcement Under Asymmetric Information. *The Economic Journal*, 99(396):408–425. Publisher: [Royal Economic Society, Wiley].

Besley, T., Fontana, N., and Limodio, N. (2021). Antitrust Policies and Profitability in Nontradable Sectors. *American Economic Review: Insights*, 3(2):251–265.

Bourreau, M., Jullien, B., and Lefouili, Y. (2021). Mergers and Demand-Enhancing Innovation. SSRN Scholarly Paper 3846118, Social Science Research Network, Rochester, NY.

Breit, W. and Elzinga, K. G. (1973). Antitrust Penalties and Attitudes toward Risk: An Economic Analysis. *Harvard Law Review*, 86(4):693–713. Publisher: The Harvard Law Review Association.

Crandall, R. W. and Winston, C. (2003). Does Antitrust Policy Improve Consumer Welfare? Assessing the Evidence. *Journal of Economic Perspectives*, 17(4):3–26.

Cunningham, C., Ederer, F., and Ma, S. (2019). Killer Acquisitions. page 106.

De Ridder, M. (2020). Market Power and Innovation in the Intangible Economy. page 69.

Decarolis, F. and Giorgiantonio, C. (2022). Corruption red flags in public procurement: new evidence from Italian calls for tenders. *EPJ Data Science*, 11(1):16. Number: 1 Publisher: Springer Berlin Heidelberg.

Denicolò, V. and Polo, M. (2021). Acquisitions, innovation and the entrenchment of monopoly. page 41.

Dickens, D. S., Kozielski, R., Khan, J., Forus, A., and Cripe, T. P. (2002). Cyclooxygenase-2 Expression in Pediatric Sarcomas. *Pediatric and Developmental Pathology*, 5(4):356–364.

Eckbo, B. E. (1992). Mergers and the Value of Antitrust Deterrence. *The Journal of Finance*, 47(3):1005–1029. _eprint: https://onlinelibrary.wiley.com/doi/pdf/10.1111/j.1540-6261.1992.tb04003.x.

Federico, G., Langus, G., and Valletti, T. (2018). Horizontal mergers and product innovation. *International Journal of Industrial Organization*, 61:590–612.

Fumagalli, C., Motta, M., and Tarantino, E. (2020). Shelving or developing? The acquisition of potential competitors under financial constraints. page 31.

Ganglmair, B., Robinson, W. K., and Seeligson, M. (2022). The Rise of Process Claims: Evidence from a Century of U.S. Patents. SSRN Scholarly Paper ID 4069994, Social Science Research Network, Rochester, NY.

Gutiérrez, G. and Philippon, T. (2017). Declining Competition and Investment in the U.S. Technical Report w23583, National Bureau of Economic Research, Cambridge, MA.

Hall, B. H. and Ziedonis, R. H. (2001). The Patent Paradox Revisited: An Empirical Study of Patenting in the U.S. Semiconductor Industry, 1979-1995. *The RAND Journal of Economics*, 32(1):101–128. Publisher: [RAND Corporation, Wiley].

Haucap, J., Rasch, A., and Stiebale, J. (2019). How mergers affect innovation: Theory and evidence. *International Journal of Industrial Organization*, 63:283–325.

Hoberg, G. and Phillips, G. (2016). Text-Based Network Industries and Endogenous Product Differentiation. *Journal of Political Economy*, 124(5):1423–1465. Publisher: The University of Chicago Press.

Iaria, A., Schwarz, C., and Waldinger, F. (2018). Frontier Knowledge and Scientific Production: Evidence from the Collapse of International Science*. *The Quarterly Journal of Economics*, 133(2):927–991.

Jullien, B. and Lefouili, Y. (2018). Horizontal Mergers and Innovation. page 29.

Kang, B. and Motohashi, K. (2015). Essential intellectual property rights and inventors' involvement in standardization. *Research Policy*, 44(2):483–492. Publisher: Elsevier.

Kuhn, J., Younge, K., and Marco, A. (2020). Patent citations reexamined. *The RAND Journal of Economics*, 51(1):109–132.

Kuhn, J. M. and Younge, K. A. (2019). Corrected Measures for Patent Citation Analysis: Accounting for Published Patent Applications. *SSRN Electronic Journal*.

Kwoka, J. (2014). *Mergers, Merger Control, and Remedies*. The MIT Press.

Lande, R. H. and Davis, J. P. (2011). Comparative Deterrence from Private Enforcement and Criminal Enforcement of the U.S. Antitrust Laws. *Brigham Young University Law Review*, 2011(2):315–390.

Lerner, J., Sorensen, M., and Strömberg, P. (2011). Private Equity and Long-Run Investment: The Case of Innovation. *The Journal of Finance*, 66(2):445–477. _eprint: https://onlinelibrary.wiley.com/doi/pdf/10.1111/j.1540-6261.2010.01639.x.

Lerner, J. and Tirole, J. (2015). Standard-Essential Patents. *Journal of Political Economy*, 123(3):547–586. Publisher: The University of Chicago Press.

Levin, R. C., Klevorick, A. K., Nelson, R. R., Winter, S. G., Gilbert, R., and Griliches, Z. (1987). Appropriating the Returns from Industrial Research and Development. *Brookings Papers on Economic Activity*, 1987(3):783.

Mermelstein, B., White, B., Nocke, V., Satterthwaite, M. A., and Whinston, M. D. (2018). Internal versus External Growth in Industries with Scale Economies: A Computational Model of Optimal Merger Policy. page 36.

Mikolov, T., Chen, K., Corrado, G., and Dean, J. (2013a). Efficient Estimation of Word Representations in Vector Space.

Mikolov, T., Sutskever, I., Chen, K., Corrado, G., and Dean, J. (2013b). Distributed Representations of Words and Phrases and their Compositionality. *arXiv:1310.4546 [cs, stat]*. arXiv: 1310.4546.

Miller, N. H. (2009). Strategic Leniency and Cartel Enforcement. *The American Economic Review*, 99(3):750–768. Publisher: American Economic Association.

Morzenti, G. (2022). A Cross Country Analysis of Stealth Consolidation and its effects on Inequality. page 48.

Motta, M. and Tarantino, E. (2017). The Effect of Horizontal Mergers, When Firms Compete in Prices and Investments. page 47.

Pellegrino, B. (2021). Product Differentiation and Oligopoly: a Network Approach. page 61.

Righi, C. and Simcoe, T. (2020). Patenting Inventions or Inventing Patents? Continuation Practice at the USPTO. Technical Report w27686, National Bureau of Economic Research, Cambridge, MA.

Rysman, M. and Simcoe, T. (2008). Patents and the Performance of Voluntary Standard-Setting Organizations. *Management Science*, 54(11):1920–1934. Publisher: INFORMS.

Thatchenkery, S. and Katila, R. (2021). INNOVATION AND PROFITABILITY FOLLOWING ANTITRUST INTERVENTION AGAINST A DOMINANT PLATFORM: THE WILD, WILD. page 65.

Vives, X. (2008). INNOVATION AND COMPETITIVE PRESSURE*: INNOVATION AND COMPETITIVE PRESSURE. *The Journal of Industrial Economics*, 56(3):419–469.

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

Wollmann, T. (2020). How to Get Away with Merger: Stealth Consolidation and Its Effects on US Healthcare. Technical Report w27274, National Bureau of Economic Research, Cambridge, MA.

Wollmann, T. G. (2019). Stealth Consolidation: Evidence from an Amendment to the Hart-Scott-Rodino Act. *American Economic Review: Insights*, 1(1):77–94.

Younge, K. A. and Kuhn, J. M. (2015). Patent-to-Patent Similarity: A Vector Space Model. *SSRN Electronic Journal*.

Zeidler, R., Csanady, M., Gires, O., Lang, S., Schmitt, B., and Wollenberg, B. (2000). Tumor cell-derived prostaglandin E2 inhibits monocyte function by interfering with CCR5 and Mac-1. *The FASEB Journal*, 14(5):661–668.

APPENDIX

.1 Model of Deterrence

A Competition in Price and Innovation

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:

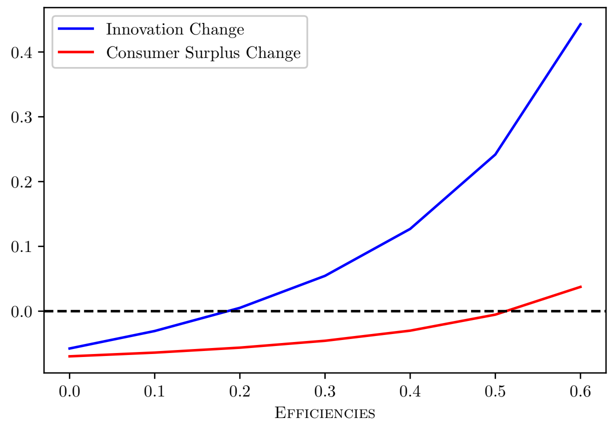
$$\begin{aligned}
2\pi_M = \max_{p_i, x_i, p_k, x_k} &(p_i - c(x_i))q_i(p_i, \bar{p}_{-i}) + (p_k - c(x_k))q_k(p_k, \bar{p}_{-k}) \\
&- F(x_i) - F(x_k) + \lambda G(x_i, x_k)
\end{aligned}$$

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

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

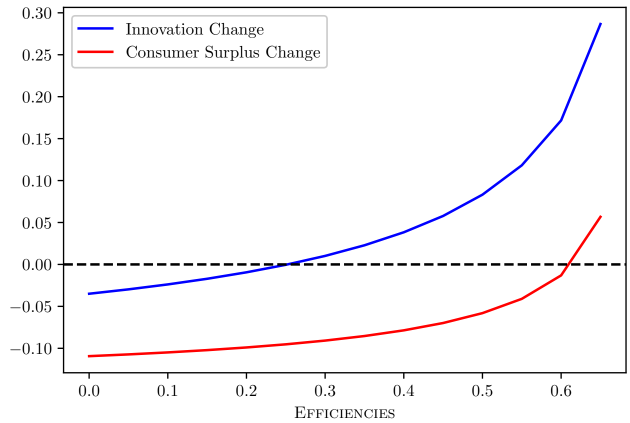
Since consumer surplus is determined by prices, then also consumer surplus increases with efficiency gains λ

B Alternative demand systems



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 linear $q_i(p_i, \bar{p}_{-i}) = \exp((1-p_i)/0.4)/\sum_j \exp((1-p_j)/0.4)$. 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 10: Changes in Consumer Surplus and Innovation after a merger



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 linear $q_i(p_i, \bar{p}_{-i}) = p_i^{-2} / \sum_j p_j^{-1}$. Marginal cost is linear in innovation $c(x_i) = 1 - x_i$. Fixed costs and efficiencies are quadratic $F(x_i) = \frac{1}{2}x_i^2, \quad G(x_i, x_k) = \frac{1}{2}x_i x_k = \frac{1}{2}x_i^2$. 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 11: Changes in Consumer Surplus and Innovation after a merger

C Simple Example and Closed Form Solution

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:

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

D EU Antitrust Policy

The US antitrust authority with perfect information described in the main paper might seem counter-intuitive to those that work with the European Commission. This antitrust authority has the power of both evaluating and eventually blocking mergers. The main problem of the Europen Commission is to discern anticompetitive mergers from pro-competitive ones. Therefore, one can assume thet the authority observes the change in consumer surplus $\Delta CS = CS_M - CS_b$ with an error:

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

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 authoritiy is willing to tolerate. Therefore, equation (8) shows the probability that a merger is allowed.

$$\begin{aligned} \alpha(\Delta CS, \sigma) = P(\widehat{\Delta CS} > H) = P(\Delta CS + \varepsilon > H) = P(\varepsilon > -\Delta CS + H) \\ \stackrel{sym}{=} P(\varepsilon < \Delta CS - H) = \Phi_\sigma(\Delta CS - H) \end{aligned} \tag{8}$$

Given the properties of CDF Φ_σ, then a similar conclusion to the main model arises. The optimal * is increasing in the true consumer surplus outcome ΔCS. Therefore, Mergers with lower ΔCS have a lower chance of being approved α, and through deterrence a lower chance of being proposed φ. This conclusion is the equivalent of Lemma 1.

E Proofs of model propositions

Proof of Lemma 2: 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.

Proof of Proposition 1: 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 $[\varphi(1) - \alpha\varphi(\lambda)] \leq [\varphi(1) - \alpha\varphi(0)] \quad \forall\lambda \in \Lambda$. 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.

Proof of Corollary 1: Call the probability of merger realization $\alpha\varphi(\lambda) = \alpha^*(\hat{CS}(\lambda))\varphi^*(\alpha^*(\hat{CS}(\lambda)))$. Call CS change implied by mergers $\hat{CS}(\lambda) = CS_M(\lambda) - CS_B$.

Call $\bar{\lambda} > 0$ the value of efficiencies such that $\hat{CS}(\bar{\lambda}) = 0$. Since $\hat{CS}$ is increasing in λ, then $\lambda \leq \bar{\lambda} \Rightarrow C\hat{S}(\lambda) \leq 0$. Then:

$$\int_0^{\bar{\lambda}} \hat{CS}(\lambda)dF(\lambda) \leq 0 \Rightarrow \int_0^{\bar{\lambda}} \hat{CS}(\lambda) \overbrace{[\varphi(1) - \alpha\varphi(\lambda)]}^{\geq 0} dF(\lambda) \leq 0$$

Since $\lambda \geq \bar{\lambda} \Rightarrow C\hat{S}(\lambda) \geq 0 \Rightarrow \alpha(\lambda) = 1 \Rightarrow \varphi(1) = \alpha\varphi(\lambda) \Rightarrow \varphi(1) - \alpha\varphi(\lambda) = 0$, then:

$$\Rightarrow \int_0^{\bar{\lambda}} \hat{CS}(\lambda)[\varphi(1) - \alpha\varphi(\lambda)]dF(\lambda) = \int_\Lambda \hat{CS}(\lambda)[\varphi(1) - \alpha\varphi(\lambda)]dF(\lambda) \leq 0$$

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

Q.E.D.

.2 Natural Language Processing

A Doc2Vec

TBD

See also a gentle introduction to Doc2Vec

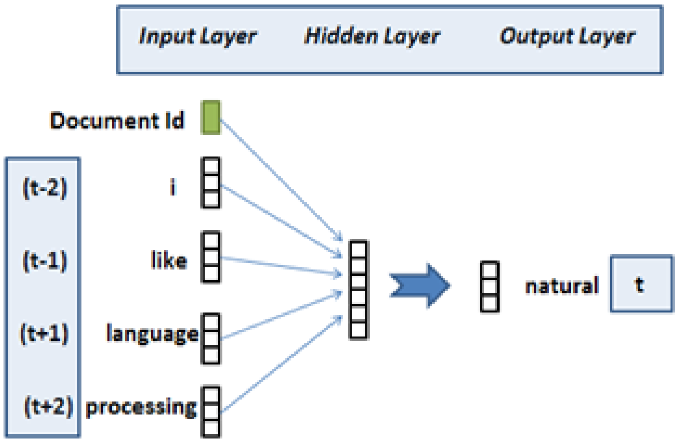


Figure 12: Doc2Vec algorithm on the sentence "I like natural language processing"

B Second Example

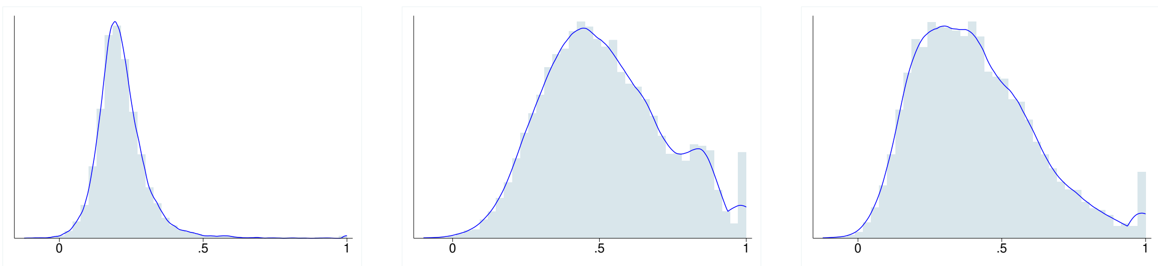
Here I report a further example of patents with high similarity. In particular, this is the couple of patents with the second highest similarity in the Pfizer-Pharmacia merger.

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

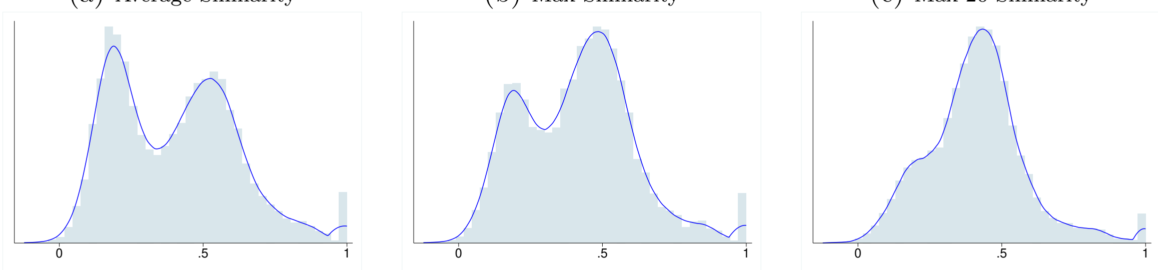
US 6809111 B2 : Prodrugs of COX-2 inhibitors "A compound of... or a pharmaceutically-acceptable salt thereof, suitable for use in the treatment of a cyclooxygenase-2 mediated disease is provided... and a method for treatment of a cyclooxygenase-2 mediated disease..." [Filed: May 15, 2003]

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.

.3 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.

Figure 13: Distribution of Patent Similarity Statistics

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.

Table 6: Correlation Table of Similarity Measures and 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 7: Performance of different cutoff rules on EC Decisions.

A Predict EC decisions

B 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 Commission (hereafter FTC).[57] 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

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

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.

Table 8: Performance of Similarity Statistics with threshold rule $[> pc(75)]$ on EC Decisions.

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 9, 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 concept.

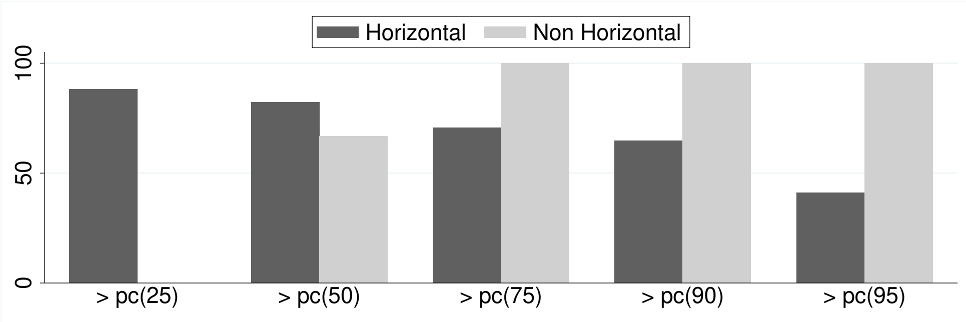
Variables	FTC	SIC	Max	Max 20	Max 1%	Max 2%	Max 5%
FTC Decisions	1.00						
SIC 4 Digits	0.15	1.00					
Max Patent Sim	0.29	0.05	1.00				
Max 20 Patent Sim	0.24	0.03	0.92	1.00			
Max 1% Patent Sim	0.35	0.03	0.86	0.92	1.00		
Max 2% Patent Sim	0.36	0.03	0.87	0.91	0.90	1.00	
Max 5% Patent Sim	0.40	0.03	0.87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

Table 9: Correlation Table of Similarity Measures and FTC Decisions

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 14 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.

Once the cutoff rule is set to the 75^{th} percentile, I compare similarity statistics in Figure 15. Using the SIC industry classification one can predict only 50% of horizontal



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 10

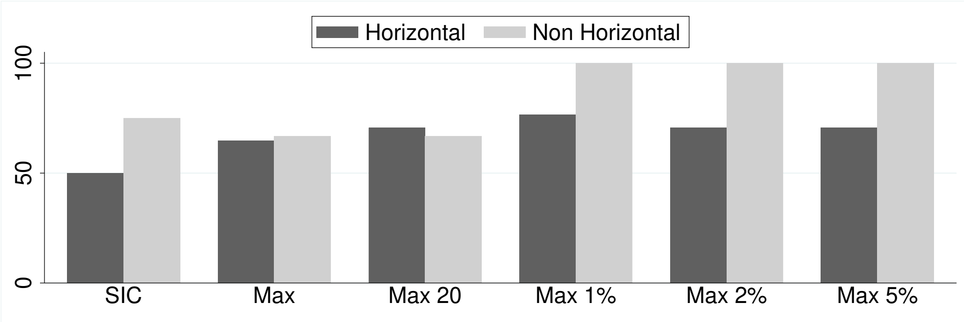
Figure 14: Performance of different cutoff rules on FTC Decisions.

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.

Table 10: Performance of different cutoff rules on FTC Decisions.

mergers, while all patent similarity statistics outperform this measure. Similarly to the correlation results in Table 9, 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.



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 11.

Figure 15: Performance of Similarity Statistics 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

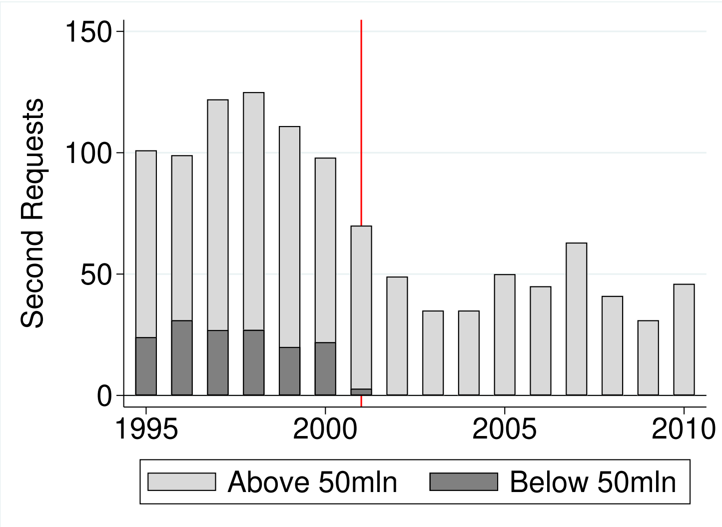
Notes: Percentage of correctly predicted horizontal (dark grey) and non horizontal (light grey) mergers, as defined by the FTC, with different methodologies. SIC stands for horizontal mergers defined using the same 4 digit SIC code, as in Wollmann (2019). The exercise is performed using the $> pc(75)$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

Table 11: Performance of Similarity Statistics with threshold rule $[> pc(75)]$ on FTC Decisions.

.4 Amendment

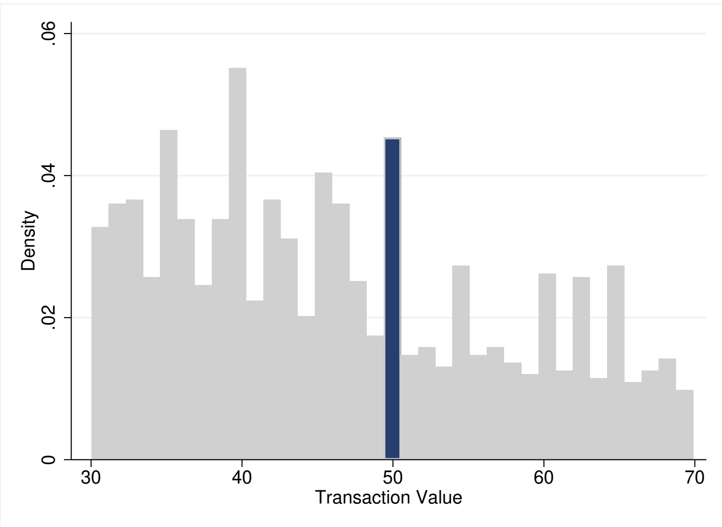
.5 More Results

.6 Robustness



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.

Figure 16: Number of Notifications received by US Antitrust Authorities.



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.

Figure 17: Distribution of Mergers by Transaction Value.

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.

Table 12: Policy changes that relaxed notifications requirements

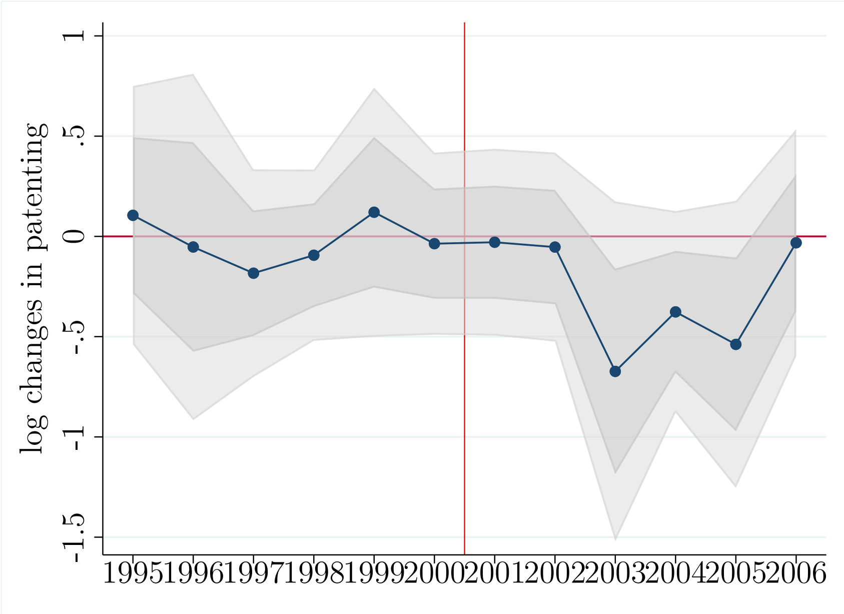


Figure 18: Coefficients of triple Diff in Diff for single years around the Amendment.

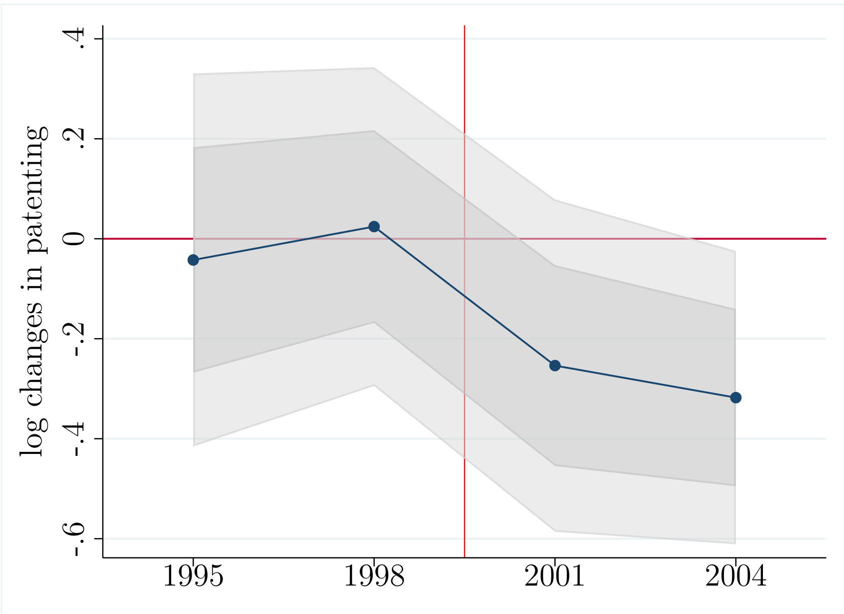


Figure 19: Coefficients of triple Diff in Diff for years in groups of 3 around the Amendment.

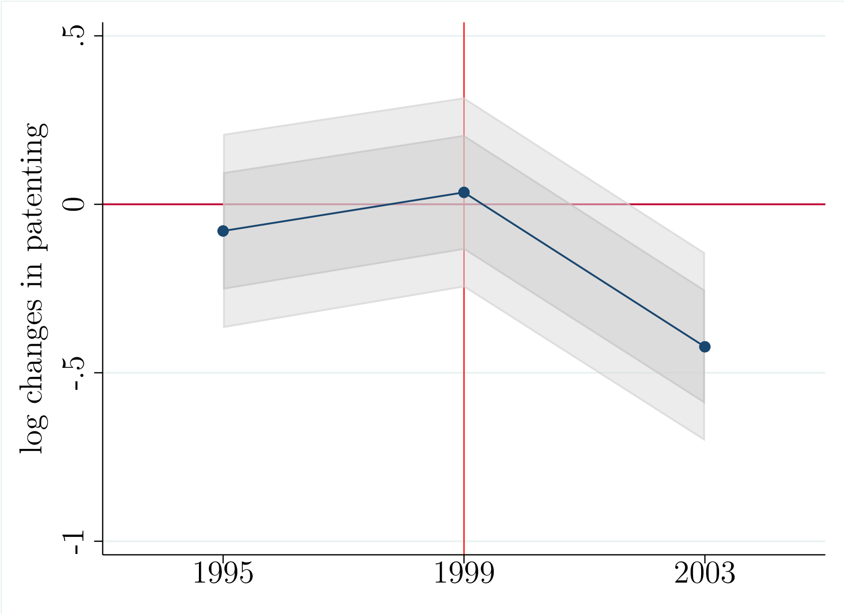
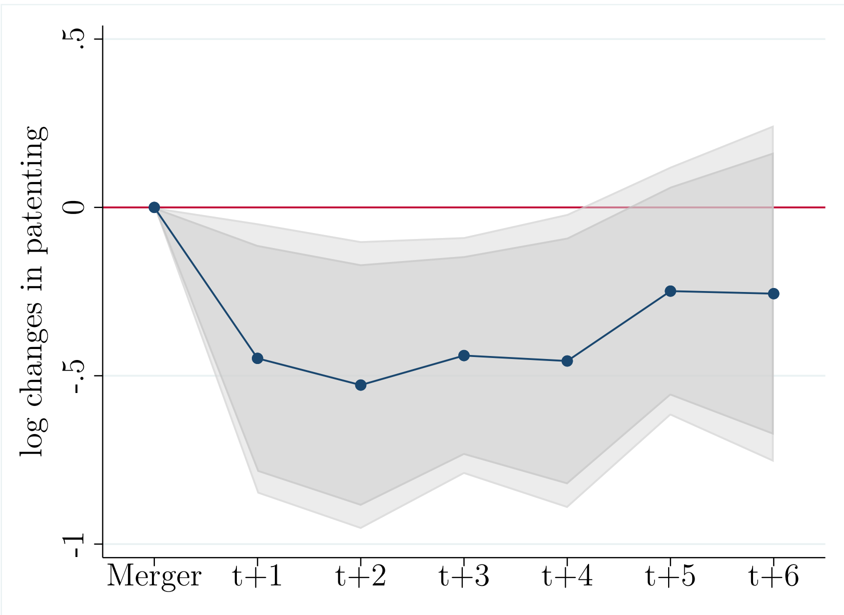


Figure 20: Coefficients of triple Diff in Diff for years in groups of 4 around the Amendment.



Notes: Coefficients of triple diff-in-diff equation 1.2 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.

Figure 21: Coefficients for different time span after the merger.

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 1.2 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 13: Triple difference in differences results for various innovation activity measures.

	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 14: Sample size by categories before and after the Amendment

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 1.2 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 15: 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
$I^{Post} \cdot I^{Ex} \cdot I^{Hor}$	-0.273*	-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 1.2 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 16: 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
$I^{Post} \cdot I^{Ex} \cdot I^{Hor}$	-0.424***	-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 1.2 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 17: 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
$I^{Post} \cdot I^{Ex} \cdot I^{Hor}$	-0.219	-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 1.2 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 18: 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
$I^{Post} \cdot I^{Ex} \cdot I^{Hor}$	-0.285**	-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 1.2 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 $\Delta P^{t+i} = ln\left(P_{t+i}\right) - ln\left(P_{t-1}\right)$.

Table 19: 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
$I^{Post} \cdot I^{Ex} \cdot I^{Hor}$	-0.334**	-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 1.2 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 20: 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
$I^{Post} \cdot I^{Ex} \cdot I^{Hor}$	-0.109	-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 1.2 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 $\Delta P^{t+i} = ln\left(P_{t+i}\right) - ln\left(P_{t-1}\right)$.

Table 21: Triple difference in differences results considering transactions close to the threshold.

Chapter 2

Second paper: A Cross Country Analysis of Stealth Consolidation and its effects on Inequality

2.1 Introduction

Antitrust policy serves a crucial role in modern economies. By limiting the market power of producers it promotes competition in the best interest of innovation and consumer welfare. Market power is defined as the ability of a firm to influence the market outcomes (usually prices and output) so as to raise its profits. By doing so in an anticompetitive way, companies retain a larger share of the production surplus, while harming the production process as a whole. They take a larger slice of the pie, while the pie is even shrinking. There is a recent body of literature that shows how market power is rising in the United States[1] and on a global level[2]. This in itself provides reasons for concern. On top of that, market power has been shown to stifle investment[3] and promote inequality[4], by driving

[1] De Loecker, Eeckhout, and Unger 2020
[2] Dıez, Fan, and Villegas-Sanchez 2019
[3] Gutiérrez and Philippon 2017
[4] Morzenti 2020

surplus away from labor[5] in favor of profits[6]. Therefore, increasing market power can be considered one of the phenomenon leading to the global rise of inequality[7].

If the aim of antitrust policy is to limit market power, but the same is rising on a global scale, then it is natural to ask: is antitrust policy missing something? Causes behind the rise in market power are likely to be varied an compelx, but one phenomenon that concur to the decline in competition is stealth consolidation. This is defined as a plethora of anticompetitive deals that go unnoticed by the antitrust authorities due to their unassuming size. The first instance of stealth consolidation happened in 2000 in United States, where an amendment to existing policy made thousands of small M&A exempt from reporting to the authorities[8]. Small mergers are considered to be trifle and they are expected to have little effect on the market structure. However, they can lead to duopolies or even monopolies in local and segmented product markets. Wollamann (2019) shows that this policy change generated thousands anticompetitive mergers that went under the radar of the authorities[9]. This is evidence that the deterrence effect of antitrust policy is strongest for anticompetitive deals, which are the ones that benefit the most from antitrust scrutiny exemptions.

Over the past two decades, thresholds determining premerger reporting requirements have risen sharply not just in the US, but around the world. As just one example, Italy amended its premerger notification program in 2012, resulting in an abrupt 90% year-over-year decline in merger filings. The first contribution of this paper is to determine whether stealth consolidation took place in several other countries[10] that experienced sim-

[5]Karabarbounis and Neiman 2014

[6]Barkai 2019

[7]Piketty 2014

[8]The amendment implied an increase in the reporting threshold set by the Hart-Scott Rodino Act, a threshold on the transaction value and the amount of assets owned by the parties. For more details on the same, the reader can refer to the Appendix.

[9]The autor shows that the amendment increased the number of horizonatal mergers that are exempt from reporting by about 3200. Horizontal transactions involve firms operating in the same narrowly defined industry and competing in the same product markets. Thus, these are considered to be most likely to be anticompetitive by the Antitrust Authorities.

[10]In particular the list of countries that experienced policy changes that are likely to lead to stealth consolidation is: United States, Germany, Italy, Belgium, Sweden, Spain, Hungary, Canada, Brazil and

ilar antitrust policy changes, which rendered hundreds of mergers exempt from reporting. This is worth investigating by itself, as the previous work of Wollmann (2019) described stealth consolidation in the United States, while leaving uncertainty on whether this is a widespread phenomenon. As a first step, I show that after such policy changes the number of merger notification received by Antitrust Authorities declined sharply in all affected countries. The fact that notification decreased is not surprising, as this was the original intent of regulators. The size of such a decrease is quite starking though, and shows that such Amendments had a significant impact on the merger policy in affected countries.

Furthermore, I show that these policy changes generate a series of anticompetitive mergers in the respective countries, providing evidence of stealth consolidation. In particular I construct a series of event studies, one for each country, in which I compare different categories of mergers before and after the amendments. First, I exploit the difference between mergers that become exempt from reporting and mergers that are never exempt, before and after the policy changes. Second, I exploit the difference between horizontal and non-horizontal mergers[11]. These two dimensions are combined in a triple difference-in-differences design, in which I show that the number of horizontal mergers that are made exempt from reporting increases significantly after the amendments[12]. This is evidence of an increase in the number of anticompetitive deals caused by the amendments, this is evidence of stealth consolidation.

The second contribution of this paper is to study the effect of stealth consolidation on the economy as a whole, and in particular the consequences of such an increase in market power on resource distribution and on inequality. How can reporting thresholds influence income inequality? Such policy changes have an effect on the market structure by incen-

Russia

[11]Horizontal transactions involve firms operating in the same narrowly defined industry and competing in the same product markets. Thus, these are considered to be most likely to be anticompetitive by the antitrust authorities. Indeed, since 1992 the US merger evaluation procedures are formally titled Horizontal Merger Giudelines.

[12]Such increase is computed with respect to the other categories, which serve as controls

tivizing small anticompetitive deals. Although small in size, some of these transactions can lead to duopoly or even to monopoly in their respective markets. This reduction in competition generates an increase in concentration and in market power. Then two are the key channels through which the effects propagates to all agents: profits and labor markets. First, market power allows firms to raise their profits, which will be distributed to firm owners. As these tend to be the richest part of the population, such increase in market power allocates a larger share of resources to the rich. Second, Berger et al. (2019) show that small local labor markets are characterized by higher concentration and more monopsony power. In these markets, low elasticity of labor supply allows firms to charge a low markdown over marginal productivity of labor, and thus to lower labor shares by paying lower wages [13] . The combination of these effects generates a significant transfer of resources from the lower to the higher end of the income distribution, increasing inequality[14].

The second contribution of this paper is composed by four event studies, which can be considered a collection of facts regarding the consequences of stealth consolidation. These event studies are all in the form of difference-in-differences which exploit variation across countries and across industries[15], and rely on policy changes as identification devices. First, I document that the Amendments cause an increase in concentration in affected industries and in affected countries. This shows that better prospect of acquisition do not spur enough entry to compensate for the plethora of anticompetitive mergers. Second, I determine that investment decreases after these policy changes. This shows that the eventual profits are not re-invested in new capital or in innovation. In the Appendix I

[13]On top of that, Manning (2011) argues that wage-posting is most common in low skilled occupations. In models of wage-posting with imperfect competition, concentration directly affects monopsony power, and decreases the level of wages (see Boal et al. 1997). On the other hand, Wozniak (2007) shows that competition in the banking sector decreases managers' wages, while leaving non-managers' compensation unaffected. This suggests that only high skilled workers, such as managers, are able to share rents with the firm.

[14]Morzenti (2020) shows all these mechanisms at play in the US alone, with a different identification strategy. That work exploits exogenous variations in the number of horizontal mergers, and it uses time series methodologies to infer their effect on the US economy.

[15]As a matter of fact, the unit of the analysis is an industry within a country in a given year.

study the effects of Amendments on R&D spending, and I find negative or non significant effects. Third, I report that Amendments cause a decrease in the labor share, which is the portion of surplus that goes to the workers. This provides evidence of increasing monopsony power, and it is a crucial mechanism in determining the distributional consequences of raising market power. Lastly, I verify that Amendments increased income inequality in their respective countries[16]. This is the ultimate effect of an increase in market power and a decrease in the labor share.

The rest of the paper is structured as follows. Firstly, a review of the literature. Section 2 describes the various datasets used in all the analyses. Section 3 details the effect of Amendments on merger dynamics and provides evidence of stealth consolidation in various countries. Section 4 studies the effect of stealth consolidation on Concentration, Investment, labor share and Inequality. Secton 5 provides robustness checkes and alternative specifications for the various event studies. Section 6 provides concluding remarks.

Related Literature

The first contribution of this paper is closely related to the work of Wollmann (2019) on the United States, and it is meant to be an extension of its results to several other countries. The author studies an Amendment to the Hart-Scott-Rodino Act that is very similar to the other ones included in this paper. By using an event study, Wollmann shows that the Amendment increased the number of horizontal mergers, raising concentration in the economy. The autor uses the term stealth consolidation to describe a widespread surge in small mergers that go under the radar of US authorities. Albeit small on paper, these transactions affect many local product markets, and they can increase significantly the level of market power in many sectors.

The second contribution of this paper is related to my previous work Morzenti (2020).

[16]For this last event study the unit of analysis is not an industry within a country, but an entire country. This is due to the fact that inequality itself is computed at the country level. To compensate for that, the dataset features a larger number of countries, so as to have enough decrees of freedom.

Here I show that the same mechanisms realize after a market power shock representing stealth consolidation hits the US economy. Rather than an event study, I use time series methodologies and I leverage a large dimensional dataset. Regardless of the methodology conclusions are the same, as stealth consolidation generates an increase in market power, which transfers resources from firm workers to firm owners generating inequality.

This paper then fits into the literature that relates firms activity to income inequality. Some recent works are descriptive, such as the work of Song et al. (2019) who use a confidential matched employer-employee database to ascertain firm contribution to the rise in earnings inequality in the US. Other papers in the literature analyze this issue from a theoretical point of view. Boar and Midrigan (2019) build a model with heterogeneous entrepreneurs that own heterogeneous firms, and the authors demonstrate that size dependent subsidies can reduce markup dispersion and increase welfare. Colciago and Mechelli (2019) study oligopolistic competition in an heterogeneous agents model, by embedding Cournot and Bertrand competition in an Aiyagary model. Eggertsson, Robbins, and Wold (2018) modify a standard neoclassical model so as to document how rising market power can explain declining interest rates and labor share. There are also papers in the Law and Economics literature on antitrust arguing that market power has an effect on inequality (Elhauge (2015); Khan and Vaheesan (2017)).

This paper contributes to the recent empirical literature on market power (De Loecker, Eeckhout, and Unger (2020); Dıez, Fan, and Villegas-Sanchez (2019)) and its effect on the economy (Karabarbounis and Neiman (2014); Gutiérrez and Philippon (2017); Berger et al. (2019)). This work contributes also to the empirical literature on inequality (Jäntti and Jenkins (2010); Heathcote, Perri, and Violante (2010); Guvenen, Ozkan, and Song (2014); De Giorgi and Gambetti (2017)).

2.2 Data Description

2.2.1 Merger Dynamics

Data on merger notifications come from different sources for any country. For the United States Wollmann (2019) provides an extensive analysis of notifications and investigations following the Amendment. For Italy, Hungary and Japan information comes from written contributions submitted for Item 5 of the 123rd meeting of the OECD Working Party No. 3 on Co-operation and Enforcement on June 2016. For Germany data comes from the Bundeskartellamt's Activity Report 1999/2000[17]. For Spain, merger notification numbers come from Global Merger Control Manual by David J. Laing, Luis A. Gómez and from the fifth edition of Merger Control by Urla Menéndez. For Belgium, information is taken from the Report to the ICN Annual Conference, Kyoto April 2008. For Sweden, number of notifications were privately provided by the Swedish Competition Authority. For Canada, number of notifications comes from the Merger Review Performance Report of April 2012. For Russia, information is taken from Report of the Federal Antimonopoly Service on Competition Policy in 2005.

Transaction level data on Mergers and Acquisitions is provided by Thompson Reuters SDC Platinum. Wollmann (2019) uses the same database to asses the effect of an Amendment to the Hart-Scott-Rodino Act in United States, which raised the threshold under which parties are exempt from reporting their transaction to the authorities. This work extends this analysis to several other countries that experienced similar changes in antitrust policy. For the sake of the analysis, mergers are difined as horizontal if the target and the acquirer operate in the same narrowly defined industry (4 digit SIC code[18], as classified by Thomson Reuters). Moreover, mergers that are affected by changes in an-

[17] The table with number of notifications is at page 205. Link: https://www.bundeskartellamt.de/SharedDocs/Publikation/DE/Taetigkeitsberichte/Bundeskartellamt%20-%20T%C3%A4tigkeitsbericht%202000.pdf?__blob=publicationFile&v=2

[18] This is a common convention in the literature, as one can see from Shahrur (2005)

titrust policy are identified based on revenue or assets of the parties, depending on the relevant jurisdiction. This allows to classify mergers in horizontal and non-horizontal, affected and non-affected, and count their respective numbers each year. Mergers that are attempted but blocked by the antitrust authority are reported in the Thompson Reuters database, but they are not included in the analysis[19]. However, since the vast majority of challenged mergers are restructured rather than abandoned, this omission does not significantly impact results of this paper. For each country, mergers are included into the analysis if either the buyer or the target firm are assigned to that country in the Thomson Reuters database. As a consequence each merger can be attributed to two different countries. This reflects the fact that mergers should be cleared by antitrust authorities in any country in which they have an effect on the economy.

2.2.2 Concentration

In order to construct industry level concentration measures, this work follows Gutiérrez and Philippon (2018), who document how EU markets became more competitive than US ones. This works uses the Herfindahl–Hirschman Index (commonly known as HHI) as a measure of concentration. This index amounts to the squared sum of market shares[20]. The authors construct measures of concentration from two different datasets. For European countries BvD Amadeus[21] provides financial data on public and private companies. For this work all size categories of firms are included[22], so as to provide the widest possible cover of industry dynamics. Industries are classified at the two digit NAICS level, for

[19]Mergers are included into the analysis only if they have an effective date, on top of an announcement date. The same is done in the work of Wollmann (2019)

[20]The HHI_{jt} is defined as $\sum_i s_{ijt}^2$ where s_{ijt} is the market share of firm i in industry j in period t, and the share is computed from firm's turnover.

[21]For this work Bureau van Dijk data were accessed through WRDS. The recent work of Kalemli-Ozcan et al. (2019) studies the representativeness of Amadeus micro data as accessed from WRDS. They conclude that data downloaded from WRDS lack coverage for several financial variables, but turnover is not one of them. For the purpose of this work, then, the WRDS database suffices.

[22]In particular the categories are: Very Large Companies, Large Companies, Medium-sized Companies, Small Companies.

a total of 23 industries[23]. All countries present in the Amadeus dataset are included in the analysis[24]. Out of these 44 countries 7 experienced changes in Antitrust Policies and thus they are considered treated: Germany, Italy, Belgium, Sweden, Spain, Hungary and Russia.

The last treated country is United States, which experienced a change in policy in December 2000. Since US is not included in the BvD Amadeus Dataset, concentration measures are computed on the Compustat[25] dataset, which collects financial information on publicly traded companies. Given the different level of coverage in the two databases, one can expect HHI indexes to be higher in US. In the analysis, country fixed effects will account for this. In accordance to the BvD Amadeus dataset, industries are classified as two digit NAICS, so as to have a correspondence between the two datasets. Given concerns over data reliability, both datasets are included from 1990 up to 2018, although Amadues data are available from 1985 and Compustat data are available from 1955.

2.2.3 Investment

Data on Investment come from the Capital Input Files of the September 2017 (Revised Juyly 2018) EU KLEMS release[26]. The dataset reports industry level aggregates for investments and capital stocks. Investment levels are computed as the ratio of Real Total Non-residential Investment and Real Total Assets[27], for each country, each industry and each year. Investment is divided by total assets so as to make values comparable across industries and countries. Data are included from 1995 up to 2015, because later years

[23]NAICS industries included in the analysis correspond to codes: 11, 21, 22, 23, 31, 32, 33, 42, 44, 45, 48, 49, 51, 52, 53, 54, 55, 56, 61, 62, 71, 72, 81, 92

[24]Countries included in the BvD Amadeus dataset: AL, AT, BA, BE, BG, BY, CH, CY, CZ, DE, DK, EE, ES, FI, FR, GB, GI, GR, HR, HU, IE, IS, IT, KV, LI, LT, LU, LV, MC, MD, ME, MK, MT, NL, NO, PL, PT, RO, RS, RU, SE, SI, SK, TR, UA

[25]Compustat North America was accessd through WRDS as well.

[26]The most recent release of the EU KLEMS database can be found at the following link: http://www.euklems.net/

[27]For the stock variables, investment is computed as Kq_OCon/Kq_GFCF. For the flow variables, investment is computed as Iq_OCon/Iq_GFCF.

are not available in the KLEMS release. The capital database comprises 26 European countries[28] and the US. Belgium and Croatia are not included in the dataset because capital information is not available for these countries. Out of these 27 countries six experienced a policy change that is relevant for this work: Germany, Italy, Sweden, Spain, Hungary and United States. All available countries and industries are used in the analysis. In particular, the KLEMS database follows the ISIC Rev. 4 industry classification. For the purpose of this work, 29 industries are included for each country, based on a mixed two digit ISIC classification[29].

2.2.4 Labor Share

Data on labor shares come from the Basic Files of the September 2017 (Revised Juyly 2018) EU KLEMS release[30]. The dataset reports industry level aggregates for several components of gross output and value added. Gutiérrez and Philippon (2018) use it, among others, to document decreasing levels of competition in Unites States, as compared to European countries. Labor shares are computed as the ratio of labor compensation and value added, for each country, each industry and each year. Data are included from 1995 up to 2015, because later years are not available in the KLEMS release. The database comprises 28 European countries[31] and the US. Out of these 29 countries seven experienced a policy change that is relevant for this work: Germany, Italy, Belgium, Sweden, Spain, Hungary and United States. All available countries and industries are used in the analysis. In particular, the KLEMS database follows the ISIC Rev. 4 industry classification. For the purpose of this work, 29 industries are included for each country, based on a mixed

[28] AT, BG, CY, CZ, DE, DK, EE, EL, ES, FI, FR, HU, IE, IT, LT, LU, LV, MT, NL, PL, PT, RO, SE, SI, SK and UK

[29] Industries included in this work are: 10-12, 13-15, 16-18, 19, 20-21, 22-23, 24-25, 26-27, 28, 29, 30, 31-33, 45, 46, 47, 49-52, 53, 58-60, 61, 62-63, A, B, D-E, F, I, K, L, M-N, O, P, Q, R, S

[30] The most recent release of the EU KLEMS database can be found at the following link: http://www.euklems.net/

[31] All member states of the EU as of 1 September 2013, namely: AT, BE, BG, HR, CY, CZ, DK, EE, FI, FR, DE, EL, HU, IE, IT, LV, LT, LU, MT, NL, PL, PT, RO, SK, SI, ES, SE, and UK

two digit ISIC classification[32].

2.2.5 Income Inequality

Data on income inequality are taken from the Gini index database constructed by the Development Research Group at the World Bank[33]. The dataset provides yearly Gini index observation for several countries. Gini index measures the extent to which the distribution of income[34] among individuals or households within an economy deviates from a perfectly equal distribution[35]. Data are available with varying levels of completeness among countries. The panel structure, thus, is not balanced. Some countries were excluded from the analysis, due to lack of observations in the dataset[36]. More information on the 66 countries included in the dataset are available in the Appendix. Out of those, 10 countries experienced changes in antitrust policy that can be exploited for the analysis: United States, Germany, Italy, Belgium, Sweden, Spain, Hungary, Canada, Brazil and Russia. Most of these countries are classified as high income countries by the World Bank, while Russia and Brazil are classified as upper middle income.

2.3 Evidence of Stealth Consolidation

This section replicated the event study of Wollmann (2019), extending it to several countries that experienced changes in merger reporting thresholds. The aim of this exercise is to ascertain whether stealth consolidation took place after such antitrust policy changes. Stealth consolidation is defined as a plethora of anticompetitive deals that go unnoticed by the antitrust authorities due to their unassuming size. In United States the Hart-

[32]Industries included in this work are: 10-12, 13-15, 16-18, 19, 20-21, 22-23, 24-25, 26-27, 28, 29, 30, 31-33, 45, 46, 47, 49-52, 53, 58-60, 61, 62-63, A, B, D-E, F, I, K, L, M-N, O, P, Q, R, S

[33]The database can be accessed at the following link: https://data.worldbank.org/indicator/SI.POV.GINI

[34]In some cases the World Bank relies on consuption expenditure, when other data is not available

[35]A Gini index of 0 represents perfect equality, while an index of 100 implies perfect inequality (all income in the hands of a single person).

[36]In particular Japan was excluded because it did not have have a minimum of 9 Gini Index observations.

Scott-Rodino Act defines a threshold under which merging parties are not required to report their transaction to the authorities[37]. The rationale behind such a threshold is to leave resources for larger merger investigations, since small mergers are considered to be trifle and they are expected to have little effect on the market structure.

Wollmann exploits an Amendment in 2000 that raised this threshold as an identification device[38]. The effect of this Amendment was to make thousands of mergers exempt from reporting to the authorities. The author shows that this policy change generated about 3200 horizontal mergers that go under the radar of the authorities in 2001-2011, and defines this process as stealth consolidation. Horizontal transactions involve firms operating in the same narrowly defined industry and competing in the same product markets. Thus, these are considered to be most likely to be anticompetitive by the Antitrust Authorities[39].

As a first step, Table 2.1 reports changes in merger notifications to the authorities, whenever data are available. As one would expect, in all countries the number of notifications decreased after the amendments. In some countries the drop in notifications was quites starking, in United States for instance they dropped by about 70%, while in Italy they fell by 90%. This table is meant to show that amendments considered in this analysis had a significant impact on merger dynamics, as they were intended to be. Were all mergers affected equally? Were the authorities able to prevent an increase in the number of anticompetitive deals following these amendments? This section will show a significant increase in the number of anticompetitive deals that were made exempt from the amendments.

[37] The Hart-Scott Rodino Act sets a threshold on the transaction value and the amount of assets owned by the parties. For more details on the same, the reader can refer to the Appendix.

[38] For more datails on the 2000 Amendment to the Hart-Scott-Rodino Act the reader can refer to the Appendix.

[39] Since 1992 the US merger evaluation procedures are formally titled Horizontal Merger Giudelines.

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

Table 2.1: Change in merger notifications received by antitrust authorities after amendments to notification thresholds. Only countries with available infrmation are included.

2.3.1 Country by Country Analysis

Following Wollmann (2019) this analysis exploits differences between the number of mergers that become exempt from reporting after the amendments (Newly-Exempt), and mergers that are never exempt, before or after the amendments (Never-Exempt). The reason behind such an identification strategy is that increases in thresholds affects only Newly-Exempt mergers, while Never-Exempt mergers are not affected. As a consequence Never-Exempt mergers are a reliable control group for Newly-Exempt ones. How are Newly-Exempt and Never-Exempt mergers defined? Each country faced a different change in its antitrust policy, and in the Appendix I provide a list that explains which mergers were affected.

On top of that, these amendments affect Horizontal mergers, which are the ones that fall under the scrutiny of antitrust authorities because they are most likely to be anti-competitive[40]. Horizontal mergers are difined as transactions between firms operating in the same narrowly defined industry[41]. As anticompetitive deals are the most likely to be restructured or even blocked by the anthoruties, the deterrence effect of antitrust policy is strongest for them. Therefore one might expect that reporting exemptions encourage more horizontal mergers with respect to non-horizontal ones. Therefore, Non-Horizontal mergers provide a suitable control group, which can account for trends in market structure and merger intensity. As a matter of fact, this exercise is a triple Difference-in-Differences that confronts the difference between the number of Newly-Exempt Horizontal mergers and the number of Never-Exempt Horizontal mergers with the same difference between Non-Horizontal ones. This is accomplished by the regression in equation (2.1), where β is the coefficient of interest.

[40] In United States the guidelines for merger control are named *Horizontal Merger Guidelines* since 1992

[41] Following Wollmann (2019) an industry is defined as 4 digit SIC code, as classified by Thomson Reuters.

$$lnMergers_{ist} = \beta I_i^H I_s^{Ex} I_t^{Post} + \gamma I_i^H + \lambda I_s^{Ex} + \mu I_i^H I_s^{Ex} + \alpha_t + \nu_{ist} \qquad (2.1)$$

The dependent variable is the logarithm of the number of mergers in each category. The dummy variable I_i^H is equal to 1 for Horizontal Mergers, the variable I_s^{Ex} is equal to 1 for Newly-Exempt mergers, and it is equal to 0 for Never-Exempt Mergers, while I_t^{Post} is equal to 1 for years after the policy change[42]. Lastly, α_t represents year fixed effects, so as to control for general trends in the number of mergers. Germany and Spain do not allow for such an analysis, since the data do not allow to identify Newly-Exempt and Never-Exempt mergers. However, one can conduct a less refined analysis, a Difference-in-Differences between the total number of Horizontal and Non-Horizontal mergers before and after the Amendments [43] . Such analysis yields a positive and significant treatment effect after the policy change.

Results of this identification strategy are reported in Table 2.2. In all countries except for Hungary and Canada treatment effects are positive and significant. For the case of Canada, it is arguable that the policy change was quite modest, if confronted to the ones seen in other countries. This is a possible explanation for the negative but insignificant coefficient. For the case of Hungary, data quality is lower with respect to other countries, and the number of transactions present in the database is limited. This might contribute to the non significant coefficient.

A further way to visualize the effects of such changes in antitrust policy is to plot the number of mergers before and after these Amendments. Figure 2.1 shows a graphical representation of the regression (2.1), the triple Diff-in-Diff design. Figure 2.2 represents

[42]The year of the policy change is not included into the treated years, but it is included in the untreated years. This is a conservative identification choice: as one can see from Figure 2.1 the effect would likely be even higher if one includes the year of policy change in the treated years.

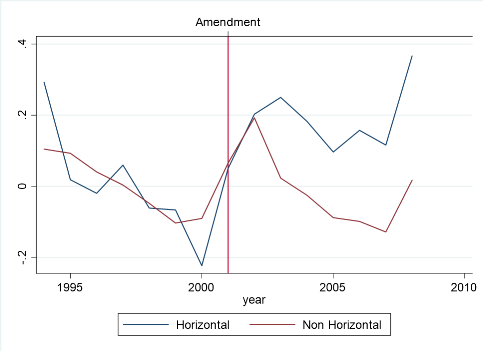
[43]For the simple Difference in Difference analysis the equation takes the following form:

$$lnMergers_{it} = \beta I_i^H I_t^{Post} + \gamma I_i^H + \alpha_t + \nu_{it}$$

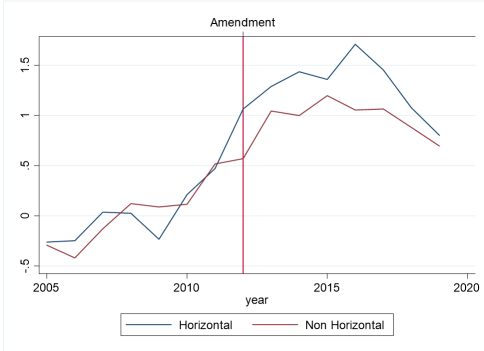
the simpler Diff-in-Diff analysis conducted for Germany and Spain. In all the countries considered in these figures it is clear that the number of horizontal mergers increases with respect to the number of non horizontal ones after the Amendments.

VARIABLES	(1) US	(2) Italy	(3) Belgium	(4) Sweden	(5) Hungary	(6) Canada	(7) Japan	(8) Russia	(9) Brazil	(10) Germany	(11) Spain
Ih_Iex_Ipost	0.184***	0.777***	0.744***	0.804**	0.453	-0.414	0.392**	0.578*	0.748***		
	(0.0587)	(0.155)	(0.209)	(0.296)	(0.582)	(0.305)	(0.159)	(0.288)	(0.265)		
Ih_Ipost										0.124**	0.223***
										(0.0514)	(0.0637)
Observations	72	76	28	32	59	40	44	43	79	22	26
R-squared	0.980	0.894	0.964	0.960	0.855	0.936	0.941	0.936	0.893	0.994	0.985
Year FE	YES	YES	YES	YES	YES	YES	YES	YES	YES	YES	YES

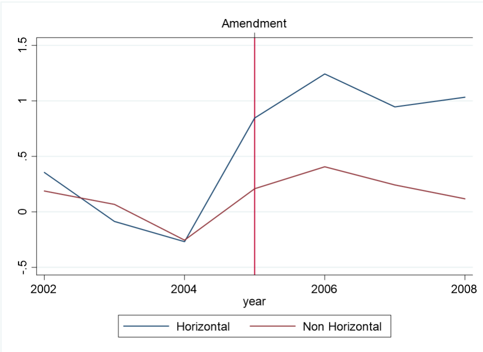
Table 2.2: Results of various event studies conducted in countries that experienced amendments to notification thresholds. All coefficients except the on for Germany and Spain, refer to triple difference in differences between horizontal and non-horizontal, newly-exempt and never-exempt mergers. Coefficients for Robust standard errors in parentheses.



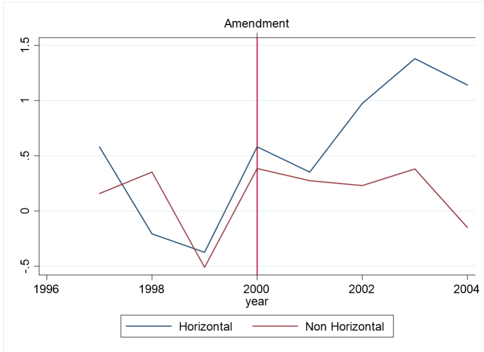
(a) United States



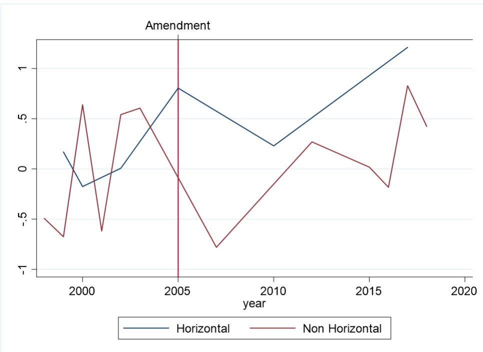
(b) Italy



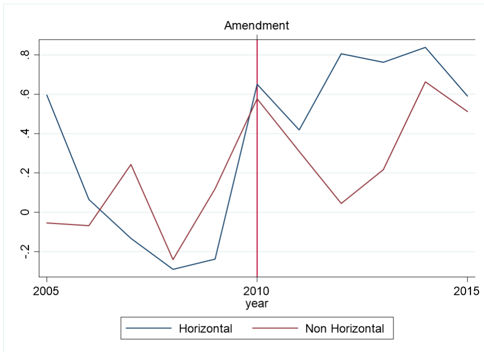
(c) Belgium



(d) Sweden

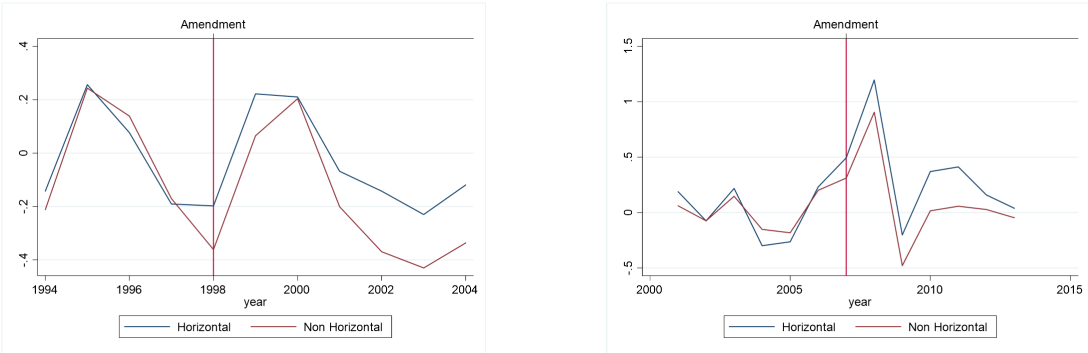


(e) Hungary



(f) Japan

Figure 2.1: Graphical representation of the triple Diff-in-Diff analysis described by equation (2.1). Each line represents the difference between the log of number of Newly-Exempt deals and the log of number of Never-Exempt Deals. The blue line represents Horizontal deals, which are expected to increase more than Non-Horizontal Deals, reported by the red line.



(a) Germany (b) Spain

Figure 2.2: Graphical representation of the simple Diff-in-Diff analysis implemented for Germany and Spain. Each line represents the log of number of mergers in each year. The blue line represents Horizontal deals, which are expected to increase more than Non-Horizontal Deals, reported by the red line.

2.3.2 Cross Country Analysis

An alternative way to inspect the results is to consider all treated countries in a single dataset. This allows to have a higher number of observations, and allows the researcher to control even better for global trends in merger markets. Equation (2.2) describes the triple difference-in-differences strategy used for this analysis, which is very similar to equation (2.1) with the addition of country fixed effects I_j. In this analysis the researcher exploits differences between Newly-Exempt and Never-Exempt mergers, as well as the difference between horizontal and non horizontal ones[44]. One important caveat for this analysis is that Newly-Exempt mergers are defined differently in every country, and this is captured by the coefficient η of the interaction between country fixed effect I_j and the dummy for Newly-Exempt mergers I_s^{Ex}. Table 2.3 reports results of this aggregate analysis, column (1) with robust errors and column (4) with clustered errors. The resulting coefficient is positive and significant, meaning that following amendments the number of Newly-Exempt horizontal deals increased.

[44] As a consequence, this analysis can be done only for countries in which it is possible to identify Newly-Exempt and Never-Exempt mergers (Italy, Belgium, Sweden, Hungary, Canada, Japan, Russia, Brazil).

$$lnMergers_{isjt} = \beta I_i^H I_s^{Ex} I_t^{Post} + \gamma I_i^H + \lambda I_s^{Ex} + \mu I_i^H I_s^{Ex} + \xi I_j + \eta I_j I_s^{Ex} + \alpha_t + \nu_{isjt} \quad (2.2)$$

Considering the aggregate sample, a further way to slice the results is to consider only Newly-Exempt mergers or only Never-Exempt mergers, and perform a simple difference in differences between horizontal and non horizontal mergers[45]. In Table 2.2 columns (2) and (5) report coefficients for Newly-Exempt deals, and as one would expect the number of such deals increases significantly after the amendments. On the other hand, columns (3) and (6) show negative coefficients for Never-Exempt deals, and this coefficient is even significant when standard errors are clustered at the country level. This can be interpreted as a substitution effect between horizontal Never-Exempt deals and horizontal Newly-Exempt deals. After the amendments, it becomes more convenient to engage in horizontal deals below the thresholds, and this diverts resources from larger deals. An alternative explanation is that after the amendments antitrust authorities are more effective in deterring large anticompetitive horizontal deals, as they divert more resources to them.

2.4 Effects of Stealth Consolidation on the Economy

2.4.1 Effect on Concentration

This section shows the results of a Cross-Country Diff-in-Diff analysis on the effect of changes in antitrust policy on industry level concentration. Concentration is defined as the Herfindahl–Hirschman Index (HHI), which is the sum of squared market shares.

[45] This is accomplished with the following equation:

$$lnMergers_{it} = \beta I_i^H I_t^{Post} + \gamma I_i^H + \xi I_j + \alpha_t + \nu_{it}$$

VARIABLES	(1) lnDeals	(2) Ex logDeals	(3) N-Ex logDeals	(4) lnDeals	(5) Ex logDeals	(6) N-Ex logDeals
Ih_Iex_Ipost	0.508***			0.508***		
	(0.128)			(0.148)		
Ih_Ipost		0.329**	-0.126		0.329**	-0.126**
		(0.140)	(0.110)		(0.135)	(0.0540)
Observations	473	222	251	473	222	251
R-squared	0.920	0.936	0.898	0.920	0.936	0.898
Year FE	YES	YES	YES	YES	YES	YES
Country FE	YES	YES	YES	YES	YES	YES
Cluster				Country	Country	Country

Table 2.3: Column (1) and (4) report results of a cross-country triple difference-in-differences between horizontal and non-horizontal, newly-exempt and never-exempt mergers. Other columns report results of a cross-country difference in differences between horizontal and non-horizontal mergers, conducted on a subsample of the data. Only countries for which it is possible to identify newly-exempt transactions are included. Robust standard errors in parentheses.

This is a widely used measure in the literature, and it is exploited by several Antitrust Authorities in evaluating persepctive mergers. One would expect that after an increase in the merger reporting threshold, the series of horizontal mergers brought by stealth consolidation results in an increase in concentration. Only an increase in the number of entrants could balance such an increase in mergers. The prospect of easier acquisition could spur the creation of new companies, with the aim of being acquired by existing incumbents, reducing concentration. Results of this section clarify which of the two effects prevails.

This analysis exploits differences between countries, as well as differences between sectors. Each observation represent HHI in an industry, in a specific country in a given year. Among the 44 countries considered in the analysis, eight experienced an Amendment to their policy that resulted in stealth consolidation: United States, Germany, Italy, Sweden, Belgium, Spain, Hungary and Russia. Equation (2.3) describes the Diff-in-Diff methodology. The dummy variable $I_{ij}^{Treated}$ is 1 for treated industries in treated countries. In this analysis all industries are considered as treated. The dummy variable I_t^{Post} is 1 for years after the Amendments in these countries. The coefficient of interest is β. The coefficient α_t represents year fixed effects, the coefficient θ_i stands for country fixed effects and γ_j stands for industry fixed effects. Results are reported in Table (2.4).

$$HHI_{ijt} = \beta I_{ij}^{Treated} I_t^{Post} + \alpha_t + \theta_i + \gamma_j + \eta X_{ijt} + \epsilon_{it} \tag{2.3}$$

The same exercise is repeated considering alternative sets of control variables X_{ijt}. If one is concerned that trends in industry level concentration are driving the result, then one can add the interaction between industry an year fixed effects, as in column (4) of Table 2.4. In this case, only the variation between countries is identifying the estimated coefficient. The same reasoning can be applied to country trends in Concentration, and column (3) shows the coefficient identified by including the interaction of country and year fixed effects. Another concern might be that country specific levels of HHI in some industries are driving the results. This can be accounted for by interacting country and industry fixed effects, as one can see in column (2). In all instances the coefficient β is estimated to be positive and significant, which implies that analyzed policy changes resulted in an increase in concentration. This proves that the entry of new competitors does not compensate for the increase in the number of horizontal mergers. After stealth consolidation, industries are more concentrated and firms have lessi incentive to compete, gaining more market power.

In order to understand whether the result is driven by just one or few countries, one can repeat the same exercise considering only one country as treated, while excluding all other treated countries from the analysis. Results are reported in Table 2.5. In all countries where the β coefficient is significant it is also positive. However, this coefficient is negative but non significant for Belgium and Sweden. Overall, one can conclude that across treated countries the analyzed change in Antitrust Policies resulted in an increase in concentration, but in some countries these effects were stronger than in others.

VARIABLES	(1) HHI	(2) HHI	(3) HHI	(4) HHI
Itreat_Ipost	0.194***	0.207***	0.767***	0.195***
	(0.0182)	(0.0184)	(0.136)	(0.0180)
Observations	17,573	17,573	17,573	17,573
R-squared	0.517	0.635	0.641	0.535
Cluster SE	C.try-Ind	C.try-Ind	C.try-Ind	C.try-Ind
Year FE	YES	YES	YES	YES
Country FE	YES	YES	YES	YES
Industry FE	YES	YES	YES	YES
Country-Industry FE		YES		
Year-Industry FE				YES
Year-Country FE			YES	

Table 2.4: Results of a difference-in-differences on concentration levels (measured as the HHI) between industries in countries that experienced amendments to reporting thresholds and countries that did not. Robust Standard Errors are clustered at Country-Industry level.

VARIABLES	(1) US	(2) DE	(3) IT	(4) SE	(5) BE	(6) ES	(7) HU	(I
Itreat_Ipost	0.549***	0.0291	0.111***	-0.0317	-0.0527	0.404***	0.128***	0.3
	(0.0164)	(0.0576)	(0.0261)	(0.0467)	(0.0442)	(0.0241)	(0.0319)	(0.0
Observations	13,760	13,678	13,682	13,619	13,624	13,772	13,611	13
R-squared	0.522	0.533	0.531	0.533	0.536	0.527	0.525	0.
Cluster SE	C.try-Ind	C.try-Ind	C.try-Ind	C.try-Ind	C.try-Ind	C.try-Ind	C.try-Ind	C.tr
Year FE	YES	YES	YES	YES	YES	YES	YES	Y
Country FE	YES	YES	YES	YES	YES	YES	YES	Y
Industry FE	YES	YES	YES	YES	YES	YES	YES	Y

Table 2.5: Results of a difference-in-differences on concentration levels (measured as the HHI) between industries in a single country that experienced amendments to reporting thresholds and all countries that did not. Robust Standard Errors are clustered at Country-Industry level.

2.4.2 Effect on Investment

This section shows the results of a Cross-Country Diff-in-Diff analysis on the effect of changes in antitrust policy on firms' investment. Higher levels of concentration and market power allow firms to gain more profits. Companies often argue that these profits are essential to sustain investment and innovation. As a way to test this claim, results of this section will inquire whether these profits are reinvested in the company. In the Appendix a separate section analyzes the effect of increases in reporting thresholds on research and development spending, documenting that R&D expenses do not increase in response to stealth consolidation.

This analysis exploits differences between countries, as well as differences between sectors. Each observation represent investment in an industry, in a specific country in a given year. Among the 27 countries considered in the analysis, six experienced an Amendment to their policy that resulted in stealth consolidation: United States, Germany, Italy, Sweden, Spain, Hungary. Equation (2.4) describes the Diff-in-Diff methodology. The dependent variable is the ratio between the stock of non-residential investment and the stock of total assets[46]. The dummy variable $I_{ij}^{Treated}$ is 1 for treated industries in treated countries. In this analysis all industries are considered as treated. The dummy variable I_t^{Post} is 1 for years after the Amendments in these countries. The coefficient of interest is β. The coefficient α_t represents year fixed effects, the coefficient θ_i stands for country fixed effects and γ_j stands for industry fixed effects. Results are reported in Table (2.6).

$$Inv_{ijt} = \beta I_{ij}^{Treated} I_t^{Post} + \alpha_t + \theta_i + \gamma_j + \eta X_{ijt} + \epsilon_{it} \tag{2.4}$$

The same exercise is repeated considering alternative sets of control variables X_{ijt}.

[46]In the robustness section the main results are shown to hold also with the ratio between flow of investment and flow of total assets.

If one is concerned that trends in industry level investment are driving the result, then one can add the interaction between industry an year fixed effects, as in column (4) of Table 2.6. In this case, only the variation between countries is identifying the estimated coefficient. The same reasoning can be applied to country trends in investment, and column (3) shows the coefficient identified by including the interaction of country and year fixed effects. Another concern might be that country specific levels of investment in some industries are driving the results. This can be accounted for by interacting country and industry fixed effects, as one can see in column (2). In all instances the coefficient β is estimated to be negative and significant, which implies that analyzed policy changes resulted in a decrease in the level of investment. This is proof that firms in affected industries do not reinvest profits gained thanks to the loosening in competition brought by stealth consolidation.

VARIABLES	(1) All Countries	(2) All Countries	(3) All Countries	(4) All Countries
Itreat_Ipost	-0.0412***	-0.0412***	-0.192***	-0.0442***
	(0.0107)	(0.0110)	(0.0420)	(0.00956)
Observations	9,778	9,778	9,778	9,778
R-squared	0.583	0.896	0.593	0.597
Cluster SE	C.try-Ind	C.try-Ind	C.try-Ind	C.try-Ind
Year FE	YES	YES	YES	YES
Country FE	YES	YES	YES	YES
Industry FE	YES	YES	YES	YES
Country-Industry FE		YES		
Year-Industry FE				YES
Year-Country FE			YES	

Table 2.6: Results of a difference-in-differences on firms investment (measured as the stock of non-residential investment divided by the stock of total assets) between industries in countries that experienced amendments to reporting thresholds and countries that did not. Robust Standard Errors are clustered at Country-Industry level.

In order to understand whether the result is driven by just one or few countries, one can repeat the same exercise considering only one country as treated, while excluding all other treated countries from the analysis. Results are reported in Table 2.7. In almost all countries the β coefficient is negative and significant. It is not significant for Germany and Italy. Overall, one can conclude that across treated countries the analyzed change in Antitrust Policies resulted in a decrease of investment, but in some countries these effects were stronger than in others.

	(1)	(2)	(3)	(4)	(5)	(6)
VARIABLES	US	DE	IT	SE	ES	HU
Itreat_Ipost	-0.0820***	-0.0208	0.0214	-0.0474**	-0.0481**	-0.0835***
	(0.0218)	(0.0208)	(0.0134)	(0.0218)	(0.0210)	(0.0264)
Observations	7,221	7,263	7,131	7,171	7,116	6,831
R-squared	0.546	0.565	0.568	0.563	0.539	0.573
Cluster SE	C.try-Ind	C.try-Ind	C.try-Ind	C.try-Ind	C.try-Ind	C.try-Ind
Year FE	YES	YES	YES	YES	YES	YES
Country FE	YES	YES	YES	YES	YES	YES
Industry FE	YES	YES	YES	YES	YES	YES

Table 2.7: Results of a difference-in-differences on firms investment (measured as the stock of non-residential investment divided by the stock of total assets) between industries in a country that experienced amendments to reporting thresholds and countries that did not. Robust Standard Errors are clustered at Country-Industry level.

2.4.3 Effect on Labor Share

This section shows the results of a Cross-Country Diff-in-Diff analysis of the effect of stealth consolidation on labor share. This is computed in each industry as the ratio between labor compensation and value added. As monopsony power increases, firms are able to extract a larger portion of the surplus (represented by value added), converting labor compensation into profits. A decrease in labor share could be accomplished by lowering wages or laying off employees, or a combination of the two. Regardless, the final result is a redistribution of resources from workers to firm owners. The empirical results of this section determine whether stealth consolidation can generate this mechanism.

This analysis exploits differences between countries, as well as differences between sectors. Each observation represent labor share in an industry, in a specific country in a given year. Among the 28 countries considered in the analysis, seven experienced an Amendment to their policy that resulted in stealth consolidation: United States, Germany, Italy, Sweden, Belgium, Spain, Hungary. Equation (2.5) describes the Diff-in-Diff methodology. The dummy variable $I_{ij}^{Treated}$ is 1 for treated industries in treated countries. In this analysis all industries are considered as treated. The dummy variable I_t^{Post} is 1 for years after the Amendments in these countries. The coefficient of interest is β. The coefficient α_t represents year fixed effects, the coefficient θ_i stands for country fixed effects and γ_j stands for industry fixed effects. Results are reported in Table (2.8).

$$LS_{ijt} = \beta I_{ij}^{Treated} I_t^{Post} + \alpha_t + \theta_i + \gamma_j + \eta X_{ijt} + \epsilon_{it} \tag{2.5}$$

The same exercise is repeated considering alternative sets of control variables X_{ijt}. If one is concerned that trends in industry level labor shares are driving the result, then one can add the interaction between industry an year fixed effects, as in column (3) of Table 2.8. In this case, only the variation between countries is identifying the estimated coefficient. Another concern might be that country specific levels of labor share in some industries are driving the results. This can be accounted for by interacting country and industry fixed effects, as one can see in column (2). In all instances in which it is significant, the coefficient β is estimated to be negative[47], which implies that analyzed policy changes resulted in a decrease of the labor share. Therefore, one can conclude that the increase

[47]The coefficient is not significant when interaction between year and country fixed effects is included in the regression, though.

in market power brought by stealth consolidation resulted in monopsony power sufficient to decrease labor compensation. Such a reduction in labor share is key to understand the redistribution channel that eventually can lead to an increse in inequality.

	(1)	(2)	(3)	(4)
VARIABLES	All Countries	Labor Share	Labor Share	Labor Share
Itreat_Ipost	-0.0212**	-0.0218**	0.0244	-0.0212**
	(0.00887)	(0.00917)	(0.0242)	(0.00895)
Observations	16,052	16,052	16,052	16,052
R-squared	0.453	0.754	0.468	0.470
Cluster SE	C.try-Ind	C.try-Ind	C.try-Ind	C.try-Ind
Year FE	YES	YES	YES	YES
Country FE	YES	YES	YES	YES
Industry FE	YES	YES	YES	YES
Country-Industry FE		YES		
Year-Industry FE				YES
Year-Country FE			YES	

Table 2.8: Results of a difference-in-differences on firms labor share (measured as the ratio between labor compensation and value added) between industries in countries that experienced amendments to reporting thresholds and countries that did not. Robust Standard Errors are clustered at Country-Industry level.

In order to understand whether the result is driven by just one or few countries, one can repeat the same exercise considering only one country as treated, while excluding all other treated countries from the analysis. Results are reported in Table 2.9. In all countries the β coefficient is negative although it is not significant for Italy, Sweden and Begium. Overall, one can conclude that across treated countries the analyzed change in Antitrust Policies resulted in a decrease of the labor share, but in some countries these effects were stronger than in others.

	(1)	(2)	(3)	(4)	(5)	(6)	(7)
VARIABLES	US	DE	IT	SE	BE	ES	HU
Itreat_Ipost	-0.0292**	-0.0404**	-0.0111	-0.0171	-0.0201*	-0.0242	-0.0957***
	(0.0131)	(0.0163)	(0.0427)	(0.0115)	(0.0117)	(0.0173)	(0.0332)
Observations	12,652	12,589	12,694	12,652	12,689	12,694	12,214
R-squared	0.437	0.438	0.437	0.442	0.438	0.433	0.435
Cluster SE	C.try-Ind	C.try-Ind	C.try-Ind	C.try-Ind	C.try-Ind	C.try-Ind	C.try-Ind
Year FE	YES	YES	YES	YES	YES	YES	YES
Country FE	YES	YES	YES	YES	YES	YES	YES
Industry FE	YES	YES	YES	YES	YES	YES	YES

Table 2.9: Results of a difference-in-differences on firms labor share (measured as the ratio between labor compensation and value added) between industries in a country that experienced amendments to reporting thresholds and all countries that did not. Robust Standard Errors are clustered at Country-Industry level

2.4.4 Effect on Income Inequality

This section shows the results of a cross-country Diff-in-Diff in which I study the effects of stealth consolidation on inequality. Previous results showed how stealth consolidation resulted in more concentrated industries, where firms are able to extract more surplus at the expenses of their employees. A decrease in the labor share and an increase in profits concentrates resources in the hands of few wealthy agents who own large shares of productive companies. This results in higher levels of income inequality, as it is documented by the following results.

In the event study I consider as treaded those countries in which there was a change in antitrust policy that resulted in stealth consolidation. The treatment, therefore, is the antitrust policy change. The sample includes 66 countries, out of which 10 experienced changes in antitrust policy. The Appendix reports an extensive list of countries included in the analysis. Equation (2.6) describes the Diff-in-Diff methodology. The dependent variable is Gini index computed on the income distribution of a given country. The dummy variable $I_i^{Treated}$ is 1 for treated countries. The dummy variable I_t^{Post} is 1 for years after the Amendments in these countries. The coefficient of interest is β. The coefficient α_t represents year fixed effects, and θ_i stands for country fixed effects.

$$Gini_{it} = \beta I_i^{Treated} I_t^{Post} + \alpha_t + \theta_i + \epsilon_{it} \tag{2.6}$$

The same exercise is done considering only advanced economies (US, Germany, Italy, Belgium, Sweden, Spain, Hungary, Canada) as treated, and dropping developing economies (Brazil, Russia) from the dataset. Then it is repeated again considering only developing economies as treated and dropping advanced economies from the dataset. Results are reported in Table 2.10, showing that income inequality increases significantly after stealth consolidation. Such increase is particularly strong for advanced economies, while it is not significant for developing economics[48]. Overall, one can conclude that a rise in monopsony power that lead to a decrease in labor share resulted in increasing inequality.

A further way to inspect this result is to plot the data before and after the Amendments, as it is shown in Figure 2.3. This analysis presents a staggered Diff-in-Diff, in which treatment is assigned in different dates to different countries. Therefore treated countries are aligned on the time dimension, so that it represents years before and after the treatment. Year 0 represents the year of the treatment. The control group is constructed by bootstrap, assigning randomly treatment dates of treated countries to non treated ones. The figure shows that countries in the control group experienced no variation in Gini index, on average. Treated countries, on the other hand, show an increase in the years following the treatment. Such increase in Gini indexes becomes statistically significant

[48]In the robustness section of this work, I show that including more controls, such as regional-year fixed effects, results in significant inequality increases also for developing countries.

	(1)	(2)	(3)
VARIABLES	AllCountries	AdvCountries	DevCountries
Itreat_Ipost	1.774**	2.423***	0.693
	(0.873)	(0.785)	(1.648)
Observations	1,029	985	928
R-squared	0.940	0.938	0.938
Year FE	YES	YES	YES
Country FE	YES	YES	YES
Cluster SE	Country	Country	Country

Table 2.10: Results of a difference-in-differences on income inequality (measured as the Gini Index) between countries that experienced amendments to reporting thresholds and countries that did not. The sample is further divided into Developing Countries and Advanced Countries. Robust Standard Errors are clustered at Country level.

only several years after the treatment. Since inequality is a complex phenomenon, one might expect it to move rather slowly in response to such events.

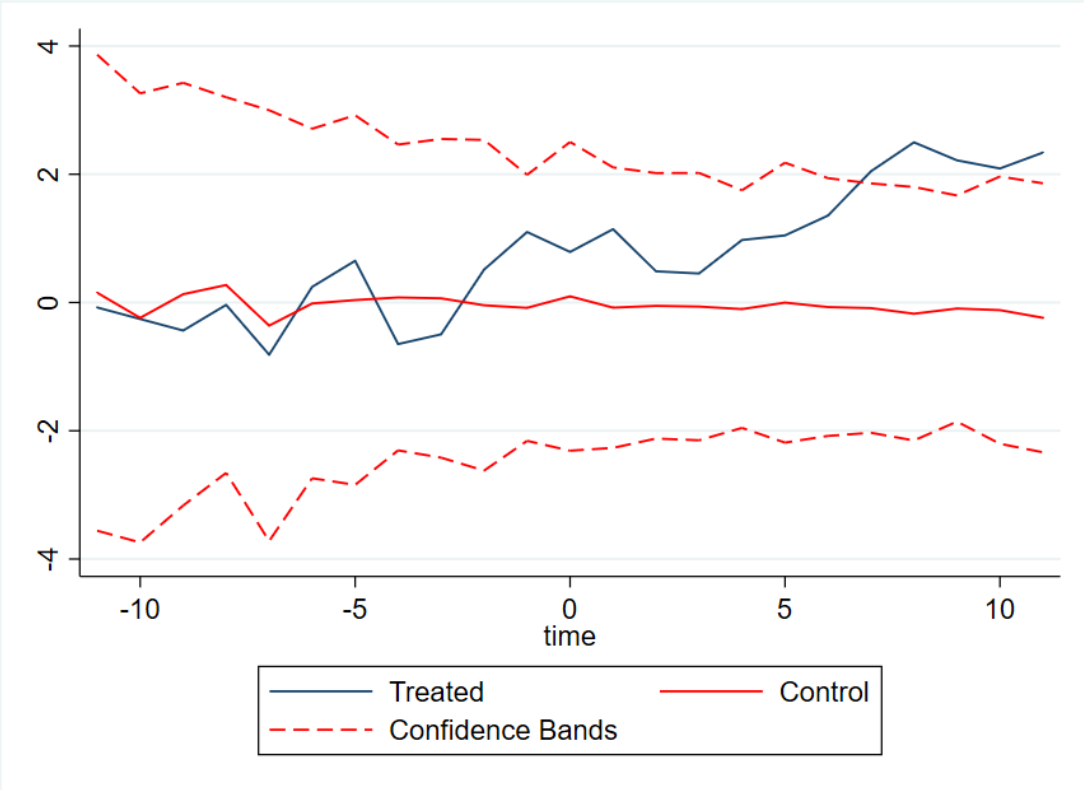


Figure 2.3: Average Gini index for the treatment and control group, only considering Advanced countries as treated. Numbers reported are residual with respect to time and country fixed effects. The horizontal axis represents years before and after the Amendments, and time 0 is the year of the Amendment. The control group is constructed by bootstrap, assigning at random the treatment date of a treated country to non-treated ones. The two series are aligned at time 0, so as to have the same value. Confidence bands represent one standard deviation of the bootstrap sample.

2.5 Robustness

2.5.1 Concentration

Policy changes considered in the analysis are heterogeneous by definition, given that they are implemented by different countries in different years. On top of that, each industry will be affected differently by these changes. As a consequence, it is natural to expect treatment effects to be heterogeneous across industries in various countries. Coefficients estimated with the Difference in Differences estimator will be a weighted average of single treatment effects. Chaisemartin and D'Haultfoeuille (2019) show that weights used to compute the DID estimator can be negative. Consequently the resulting coefficient can be of opposite sign with respect to the Average Treatment Effects (ATE), provided that there is enough heterogeneity between ATE. Computing the weights as in Chaisemartin and D'Haultfoeuille (2019) shows that out of 111 ATE, none receive a negative weight. Therefore there is no reason for concern that the DID coefficient is of a different sign with respect to the ATE[49].

As a way to further exploit differences between industries, one can consider a triple difference in differences design, excluding some industries from the treatment sample. In each treated country I consider all industries except for "Financial and Insurance Activities" as treated. Therefore this industry act as a control with respect to all others. The choice of "Financial and Insurance Activities" is justified by the fact that transactions in this industry tend to be much larger, and as such they are not affected by stealth consolidation. Table 2.11 shows the average transaction value of mergers in financial an non financial sectors, and clearly in every country financial mergers are significantly larger. On top of that, this industry is heavily regulated in any country, and in several cases mergers are revised by industry authorities together with the Antitrust Authority[50]. As an example, in United States following the Federal Deposit Insurance Act, the Bank Holding Company Act, and the Change in Bank Control Act, firms that provide banking services are always subject to at least one antitrust review, and oftentimes overlapping reviews, depending on their size and affiliation.

Equation (2.7) describes the triple Diff-in-Diff methodology. The dummy variable $I_{ij}^{Treated}$ is 1 for treated industries in treated countries. The dummy variable I_t^{Post} is 1 for years after the Amendments in these countries. The coefficient of interest is β. The

[49]The robust estimator proposed by Chaisemartin and D'Haultfoeuille (2019) is positive and significant when computed on the dataset of this analysis. However, there is no need to use it, since the classical DID estimator is negative and significant and negative weighting is not an issue.

[50]As was the case for Italy. Until 2005 the central bank (Banca d'Italia) supervised instances of concentration. Then after 2005 this supervision was assigned to the antitrust authority (AGCM).

COUNTRY	Sector	Share	Transaction_Value
Belgium	Non Financial	95.25	161.6
	Financial	4.746	1052
Germany	Non Financial	97.10	238.8
	Financial	2.899	537.9
Hungary	Non Financial	93.83	58.45
	Financial	6.169	65.32
Italy	Non Financial	94.26	161.6
	Financial	5.739	272.3
Spain	Non Financial	94.94	126.8
	Financial	5.064	140.1
Sweden	Non Financial	97.20	107.7
	Financial	2.802	259.3
United_States	Non Financial	92.54	210.3
	Financial	7.459	310.8

Table 2.11: Statistics on mergers in Financial and Non-Financial sectors. "Share" represents the share in terms of numbers of mergers in the dataset. "Transaction Value" represents the average transaction value of mergers, meaning the amount paid to the target company for the acquisition.

coefficient α_t represents year fixed effects, the coefficient θ_i stands for country fixed effects and γ_j stands for industry fixed effects. Results are reported in Table (2.12).

$$HHI_{ijt} = \beta I_{ij}^{Treated} I_t^{Post} + \alpha_t + \theta_i + \gamma_j + \eta X_{ijt} + \epsilon_{it} \tag{2.7}$$

The same exercise is repeated considering alternative sets of control variables X_{ijt}. As in the main analysis, the same sets of controls are included in column (2) and (3) and (4). In all instances the coefficient β is estimated to be positive, which implies that analyzed policy changes resulted in an increase in concentration. The coefficient is not significant when interaction between year and country fixed effects is included in the regression.

	(1)	(2)	(3)	(4)
VARIABLES	HHI	HHI	HHI	HHI
Itreat_Ipost	0.176***	0.202***	0.00721	0.177***
	(0.0183)	(0.0193)	(0.0384)	(0.0181)
Observations	17,573	17,573	17,573	17,573
R-squared	0.516	0.634	0.641	0.533
Cluster SE	C.try-Ind	C.try-Ind	C.try-Ind	C.try-Ind
Year FE	YES	YES	YES	YES
Country FE	YES	YES	YES	YES
Industry FE	YES	YES	YES	YES
Country-Industry FE		YES		
Year-Industry FE				YES
Year-Country FE			YES	

Table 2.12: Results of a triple difference-in-differences on concentration levels (measured as the HHI) between industries and countries that experienced amendments to reporting thresholds and industries and countries that did not. Robust Standard Errors are clustered at Country-Industry level.

In order to understand whether these results are driven by just one or few countries, one can repeat the same exercise considering only one country as treated, while excluding all other treated countries from the analysis. Results are reported in Table 2.13. Results are very similar to the main analysis, with negative coefficients for Sweden and Spain. As in the main analysis, one can conclude that across treated countries the analyzed change in Antitrust Policies resulted in an increase in concentration, but in some countries these effects were stronger than in others.

VARIABLES	(1) US	(2) DE	(3) IT	(4) SE	(5) BE	(6) ES	(7) HU	(8) RU
Itreat_Ipost	0.517***	0.0235	0.115***	-0.0162	-0.0633	0.392***	0.123***	0.281***
	(0.0326)	(0.0462)	(0.0261)	(0.0387)	(0.0420)	(0.0263)	(0.0315)	(0.0393)
Observations	13,760	13,678	13,682	13,619	13,624	13,772	13,611	13,639
R-squared	0.520	0.533	0.531	0.533	0.536	0.526	0.525	0.526
Cluster SE	C.try-Ind	C.try-Ind	C.try-Ind	C.try-Ind	C.try-Ind	C.try-Ind	C.try-Ind	C.try-Ind
Year FE	YES	YES	YES	YES	YES	YES	YES	YES
Country FE	YES	YES	YES	YES	YES	YES	YES	YES
Industry FE	YES	YES	YES	YES	YES	YES	YES	YES

Table 2.13: Results of a triple difference-in-differences on concentration levels (measured as the HHI) between industries in a country that experienced amendments to reporting thresholds and industries and countries that did not. Robust Standard Errors are clustered at Country-Industry level.

2.5.2 Investment

Policy changes considered in the analysis are heterogeneous by definition, given that they are implemented by different countries in different years. On top of that, each industry will be affected differently by these changes. As a consequence, it is natural to expect treatment effects to be heterogeneous across industries in various countries. Coefficients estimated with the Difference in Differences estimator will be a weighted average of single treatment effects. Chaisemartin and D'Haultfoeuille (2019) show that weights used to compute the DID estimator can be negative. Consequently the resulting coefficient can be of opposite sign with respect to the Average Treatment Effects (ATE), provided that there is enough heterogeneity between ATE. Computing the weights as in Chaisemartin and D'Haultfoeuille (2019) shows that out of 64 ATE, none receive a negative weight. Therefore there is no reason for concern that the DID coefficient is of a different sign with respect to the ATE[51].

[51] The robust estimator proposed by Chaisemartin and D'Haultfoeuille (2019) is not significant when computed on the dataset of this analysis. However, this is likely due to a lack of power, since the classical DID estimator is negative and significant in this case, and negative weighting is not an issue.

As a way to further exploit differences between industries, one can consider a triple difference in differences design, excluding some industries from the treatment sample. In each treated country I consider all industries except for "Financial and Insurance Activities" as treated. Therefore this industry act as a control with respect to all others. The choice of "Financial and Insurance Activities" is justified by the fact that transactions in this industry tend to be much larger, and as such they are not affected by stealth consolidation[52]. Results are reported in Table (2.14). The same exercise is repeated considering alternative sets of control variables. As in the main analysis, the same sets of controls are included in column (2) and (3) and (4). The coefficient β is estimated to be negative in all instances in which it is significant, which implies that analyzed policy changes resulted in a decrease in investment levels. The coefficient is not significant when interaction between year and country fixed effects is included in the regression, however.

[52] Table 2.11 shows the average transaction value of mergers in financial an non financial sectors, and clearly in every country financial mergers are significantly larger

VARIABLES	(1) All Countries	(2) All Countries	(3) All Countries	(4) All Countries
Itreat_Ipost	-0.0342**	-0.0327***	0.0359	-0.0424***
	(0.0154)	(0.0109)	(0.120)	(0.0154)
Observations	9,778	9,778	9,778	9,778
R-squared	0.582	0.896	0.593	0.597
Cluster SE	C.try-Ind	C.try-Ind	C.try-Ind	C.try-Ind
Year FE	YES	YES	YES	YES
Country FE	YES	YES	YES	YES
Industry FE	YES	YES	YES	YES
Country-Industry FE		YES		
Year-Industry FE				YES
Year-Country FE			YES	

Table 2.14: Results of a triple difference-in-differences on firms investment (measured as the stock of non-residential investment divided by the stock of total assets) between industries and countries that experienced amendments to reporting thresholds and countries and industries that did not. Robust Standard Errors are clustered at Country-Industry level.

Another possible dimension on which to test the robustness of main results is the dependent variable. The main section analyzes the change in the stock of non-residential investment[53] in response to stealth consolidation. One then can study the response of the flow of non-residential investment[54], meaning the annual change in the stock. At a constant depreciation rate, if the flow is increasing over time, also the stock should increase. On the contrary, an increase in the flow of investment could just be a response

[53]as measured by Kq_OCon/Kq_GFCF in the KLEMS capital input files.
[54]as measured by Iq_OCon/Iq_GFCF in the KLEMS capital input files.

to rising depreciation rate brought on by technological change. If stock and flow move together, then this means that the effect is not coused by a change in the depreciation rate. Table 2.15 shows the response of flow investment to stealth consolidation. In all specifications the response is negative, as in the main results. The coefficient is not significant when one includes year-country fixed effects, though. The fact that both the stock and the flow of non-residential investment is decreasing after episodes of stealth consolidation means that firms have less incentive to invest. This, in turn, can be evidence of increased market power, which allows firms to compete less fiercely.

	(1)	(2)	(3)	(4)
VARIABLES	All Countries	All Countries	All Countries	All Countries
Itreat_Ipost	-0.0334***	-0.0336***	-0.0119	-0.0352***
	(0.0102)	(0.0104)	(0.0359)	(0.00971)
Observations	12,057	12,057	12,057	12,057
R-squared	0.252	0.403	0.283	0.289
Cluster SE	C.try-Ind	C.try-Ind	C.try-Ind	C.try-Ind
Year FE	YES	YES	YES	YES
Country FE	YES	YES	YES	YES
Industry FE	YES	YES	YES	YES
Country-Industry FE		YES		
Year-Industry FE				YES
Year-Country FE			YES	

Table 2.15: Results of a difference-in-differences on firms investment (measured as the flow of non-residential investment divided by the flow of total assets) between industries in countries that experienced amendments to reporting thresholds and countries that did not. Robust Standard Errors are clustered at Country-Industry level.

2.5.3 Labor Share

Policy changes considered in the analysis are heterogeneous by definition, given that they are implemented by different countries in different years. On top of that, each industry will be affected differently by these changes. As a consequence, it is natural to expect treatment effects to be heterogeneous across industries in various countries. Coefficients estimated with the Difference in Differences estimator will be a weighted average of single treatment effects. Chaisemartin and D'Haultfoeuille (2019) show that weights used to compute the DID estimator can be negative. Consequently the resulting coefficient can be of opposite sign with respect to the Average Treatment Effects (ATE), provided that there is enough heterogeneity between ATE. Computing the weights as in Chaisemartin and D'Haultfoeuille (2019) shows that out of 73 ATE, only 3 receive a negative weight. Therefore there is no reason for concern that the DID coefficient is of a different sign with respect to the ATE[55].

As a way to further exploit differences between industries, one can consider a triple difference in differences design, excluding some industries from the treatment sample. In each treated country I consider all industries except for "Financial and Insurance Activities" as treated. Therefore this industry act as a control with respect to all others. The choice of "Financial and Insurance Activities" is justified by the fact that transactions in this industry tend to be much larger, and as such they are not affected by stealth consolidation. Table 2.11 shows the average transaction value of mergers in financial an non financial sectors, and clearly in every country financial mergers are significantly larger.

Equation (2.8) describes the triple Diff-in-Diff methodology. The dummy variable $I_{ij}^{Treated}$ is 1 for treated industries in treated countries. The dummy variable I_t^{Post} is 1 for years after the Amendments in these countries. The coefficient of interest is β. The

[55]The robust estimator proposed by Chaisemartin and D'Haultfoeuille (2019) is not significant when computed on the dataset of this analysis. However, this is likely due to a lack of power, since the classical DID estimator is negative and significant in this case, and negative weighting is not an issue.

VARIABLES	(1) Labor Share	(2) Labor Share	(3) Labor Share	(4) Labor Share
Itreat_Ipost	-0.0255***	-0.0215**	-0.0605	-0.0271***
	(0.00922)	(0.00928)	(0.0381)	(0.00947)
Observations	16,052	16,052	16,052	16,052
R-squared	0.454	0.754	0.468	0.470
Cluster SE	C.try-Ind	C.try-Ind	C.try-Ind	C.try-Ind
Year FE	YES	YES	YES	YES
Country FE	YES	YES	YES	YES
Industry FE	YES	YES	YES	YES
Country-Industry FE		YES		
Year-Industry FE				YES
Year-Country FE			YES	

Table 2.16: Results of a triple difference-in-differences on firms labor share (measured as the ratio between labor compensation and value added) between industries and countries that experienced amendments to reporting thresholds and countries and industries that did not. Robust Standard Errors are clustered at Country-Industry level.

coefficient α_t represents year fixed effects, the coefficient θ_i stands for country fixed effects and γ_j stands for industry fixed effects. Results are reported in Table (2.16).

$$LS_{ijt} = \beta I_{ij}^{Treated} I_t^{Post} + \alpha_t + \theta_i + \gamma_j + \eta X_{ijt} + \epsilon_{it} \tag{2.8}$$

The same exercise is repeated considering alternative sets of control variables X_{ijt}. As in the main analysis, the same sets of controls are included in column (2) and (3). In all instances the coefficient β is estimated to be negative, which implies that analyzed policy changes resulted in a decrease of the labor share. Again, the coefficient is not significant when interaction between year and country fixed effects is included in the regression.

In order to understand whether these results are driven by just one or few countries, one can repeat the same exercise considering only one country as treated, while excluding all other treated countries from the analysis. Results are reported in Table 2.17. In all countries the β coefficient is negative although it is not significant for Italy, Sweden and Begium and Spain. As in the main analysis, one can conclude that across treated countries the analyzed change in Antitrust Policies resulted in a decrease of the labor share, but in some countries these effects were stronger than in others.

VARIABLES	(1) US	(2) DE	(3) IT	(4) SE	(5) BE	(6) ES	(7) HU
Itreat_Ipost	-0.0390**	-0.0613***	-0.0116	-0.0133	-0.0186	-0.0259	-0.141***
	(0.0154)	(0.0214)	(0.0442)	(0.0121)	(0.0117)	(0.0174)	(0.0325)
Observations	12,652	12,589	12,694	12,652	12,689	12,694	12,214
R-squared	0.437	0.438	0.437	0.442	0.438	0.433	0.435
Cluster SE	C.try-Ind	C.try-Ind	C.try-Ind	C.try-Ind	C.try-Ind	C.try-Ind	C.try-Ind
Year FE	YES	YES	YES	YES	YES	YES	YES
Country FE	YES	YES	YES	YES	YES	YES	YES
Industry FE	YES	YES	YES	YES	YES	YES	YES

Table 2.17: Results of a triple difference-in-differences on firms labor share (measured as the ratio between labor compensation and value added) between industries in a country that experienced amendments to reporting thresholds and countries that did not. Robust Standard Errors are clustered at Country-Industry level.

2.5.4 Inequality

Policy changes considered in the analysis are heterogeneous by definition, given that they are implemented by different countries in different years. On top of that, each industry will be affected differently by these changes. As a consequence, it is natural to expect treatment effects to be heterogeneous across various countries. Coefficients estimated with the Difference in Differences estimator will be a weighted average of single treatment effects. Chaisemartin and D'Haultfoeuille (2019) show that weights used to compute the DID estimator can be negative. Consequently the resulting coefficient can be of opposite sign with respect to the Average Treatment Effects (ATE), provided that there is enough heterogeneity between ATE. Computing the weights as in Chaisemartin and D'Haultfoeuille (2019) shows that out of 81 ATE, only 8 receive a negative weight. Therefore there is no reason for concern that the DID coefficient is of a different sign with respect to the ATE[56].

Confounding regional effects could be an alternative explanation of the main results. If some regions experienced increasing Gini index, and treated countries happen to be more frequently in these regions, then the main result would not reflect the effect of the Amendments, but only these regional trends. In order to account for this alternative explanation, one can add an interaction between time fixed effects and regional fixed effects to the original regression:

$$Gini_{it} = \beta I_i^{Treated} I_t^{Post} + \alpha_t + \theta_i + \gamma_{j,t} + \epsilon_{it}$$

Where $\gamma_{j,t}$ is an time t and region j fixed effect. Some example of region can be "Europe and Central Asia", "North America", "East Asia and Pacific". Table 2.18 shows that coefficients remain significant, although the one for Advanced Economies decreases substantially.

VARIABLES	(1) AllCountries	(2) AdvCountries	(3) DevCountries
Itreat_Ipost	1.557***	1.443**	1.848**
	(0.487)	(0.627)	(0.755)
Observations	1,029	985	928
R-squared	0.958	0.957	0.957
Year FE	YES	YES	YES
Country FE	YES	YES	YES
Region-Year FE	YES	YES	YES

Table 2.18: Results of a difference-in-differences on income inequality (measured as the Gini Index) between countries that experienced amendments to reporting thresholds and countries that did not. The interaction between region and year fixed effect is included in the controls. The sample is further divided into Developing Countries and Advanced Countries. Robust Standard Errors are clustered at Country level.

[56] The robust estimator proposed by Chaisemartin and D'Haultfoeuille (2019) is not significant when computed on the dataset of this analysis. However, this is likely due to a lack of power, since the classical DID estimator is negative and significant in this case, and negative weighting is not an issue.

Another concern is that a different trend in inequality in advanced countries is driving the main result. Similarly to regional trends, this can be controlled for by the interaction between time fixed effect and income group (as defined by the World Bank) fixed effects. The resulting regression is:

$$Gini_{it} = \beta I_i^{Treated} I_t^{Post} + \alpha_t + \theta_i + \gamma_{k,t} + \epsilon_{it}$$

Where $\gamma_{j,t}$ is an time t and income group k fixed effect. Income groups are "High income", "Upper middle income" and "Lower middle income". Table 2.19 shows that coefficients remain significant, although the one for Advanced Economies decreases substantially.

	(1)	(2)	(3)
VARIABLES	AllCountries	AdvCountries	DevCountries
Itreat_Ipost	1.451***	1.448**	1.455*
	(0.520)	(0.668)	(0.851)
Observations	1,029	985	928
R-squared	0.945	0.944	0.944
Year FE	YES	YES	YES
Country FE	YES	YES	YES
IG_Year FE	YES	YES	YES

Table 2.19: Results of a difference-in-differences on income inequality (measured as the Gini Index) between countries that experienced amendments to reporting thresholds and countries that did not. The interaction between income group and year fixed effect is included in the controls. The sample is further divided into Developing Countries and Advanced Countries. Robust Standard Errors are clustered at Country level.

2.6 Conclusion

This paper documents instances of stealth consolidation in several developed and developing countries and it describes their effects on resource distribution within the economy. stealth consolidation is defined as a series of anticompetitive mergers that go unnoticed by the Antitrust Authorities. In several jurisdictions mergers that fall below certain thresholds are not required to report to the authorities. These threshold have been raised in various countries in sudden and significant ways, making thousands of mergers exempt from reporting. These Amendments provide discontinuities that can be used as reliable identification tools for event studies on their effects on the market structure and on the economy.

The contribution of the paper is twofold. First, following the work of Wollmann (2019), I show that these policy changes resulted in a decrease in merger notifications and in an increase in the number of anticompetitive deals that go unnoticed by the authorities. Therefore I provide evidence that stealth consolidation is a widespread phenomenon and that it is relevant also outside of the US. Second, I describe a series of facts that follow instances of stealth consolidation. I document that industries in affected countries show higher level of concentration, as one would expect after a plethora of anticompetitive deals. Then I show that investment decreases in these industries, as proof that eventual profits are not invested in technological change or in innovation[57]. Thereafter I document a decline of labor share in affected industries, which implies that surplus is diverted from workers to firm owners. Lastly, I describe an increase in income inequality following the Amendments, which is the results of rising profits and decreasing workers' compensation.

Policy implications of these results are clear. Starting from the early 2000s, many antitrust authorities introduced or raised notification thresholds for merging parties. The rationale behind such policy was to divert resources to larger merger investigations, under the assumption that small transactions have no effect on the market structure. This paper makes a case for considering these thresholds carefully and extend antitrust scrutiny to smaller transactions. This should be done in order to prevent stealth consolidation, a series of anticompetitive mergers with the potential to decrease innovation, lower labor share and increase inequality. It is worth noting, though, that more work is needed to

[57]In the Appendix I also report that the Amendments have a negative or insignificant effect on R&D expenditure.

assess the overall welfare impact of Stealth Consolidation, as this would be a first extension of the present paper.

The empirical results of this work can be the basis of several new strands of research. The acquisition of highly innovative start-ups by the tech giants has attracted the attention of antitrust authorities around the world. Since these transaction involve small parties, many tend to fall below reporting threshold, and can concur to stealth consolidation. One possible extension of this paper can focus on innovation effort by entrants, incumbents and acquired firms, by exploiting the existing threshold or variation of the same. This could shed some light on these high tech acquisitions are anticompetitive and detrimental to innovation. The recent work of Cunningham, Ederer, and Ma (2019) describes the phenomenon of Killer Acquisitions. The authors document that in the pharmaceutical sector incumbents acquire potential entrants with the sole purpose of discontinuing competing products. This phenomenon is particularly accentuated below reporting threshold. Studies on changes of these thresholds could provide further evidence on the practice of acquisition aimed at eliminating rivals. Lastly, the empirical findings of this work justify the study of a theoretical model featuring both heterogeneous agent and an antitrust authority, so as to study the effects that merger policy can have on resource allocation and redistribution.

REFERENCES

Barkai, S. (2019) 'Declining Labor and Capital Shares', SSRN Electronic Journal. doi: 10.2139/ssrn.3489965.

Berger, D., Herkenhoff, K. and Mongey, S. (2019) Labor Market Power. w25719. Cambridge, MA: National Bureau of Economic Research, p. w25719. doi: 10.3386/w25719.

Boal, W. M., Ransom, M. R. and Robinson, J. (1997) 'Monopsony in the Labor Market', p. 28.

Boar, C. and Midrigan, V. (2019) Markups and Inequality. w25952. Cambridge, MA: National Bureau of Economic Research, p. w25952. doi: 10.3386/w25952.

Chaisemartin, C. and D'Haultfoeuille, X. (2019) Two-way Fixed Effects Estimators with Heterogeneous Treatment Effects. w25904. Cambridge, MA: National Bureau of Economic Research, p. w25904. doi: 10.3386/w25904.

Colciago, A. and Mechelli, R. (2019) Competition and Inequality: Aiyagari meets Bertrand and Cournot. SSRN Scholarly Paper ID 3312919. Rochester, NY: Social Science Research Network. Available at: https://papers.ssrn.com/abstract=3312919 (Accessed: 24 May 2019).

Conley, T. G. and Taber, C. R. (2011) 'Inference with "Difference in Differences" with a Small Number of Policy Changes', The Review of Economics and Statistics, 93(1), pp. 113–125. doi: 10.1162/REST$_{a0}$0049.

Cunningham, C., Ederer, F. and Ma, S. (2019) 'Killer Acquisitions', p. 106.

De Giorgi, G. and Gambetti, L. (2017) 'Business cycle fluctuations and the distribution of consumption', Review of Economic Dynamics, 23, pp. 19–41. doi: 10.1016/j.red.2016.07.005.

De Loecker, J., Eeckhout, J. and Unger, G. (2020) 'The Rise of Market Power and the Macroeconomic Implications*', The Quarterly Journal of Economics, 135(2), pp. 561–644. doi: 10.1093/qje/qjz041.

Dıez, F. J., Fan, J. and Villegas-Sanchez, C. (2019) 'Global Declining Competition',

p. 41.

Eggertsson, G. B., Robbins, J. A. and Wold, E. G. (2018) Kaldor and Piketty's Facts: The Rise of Monopoly Power in the United States. Working Paper 24287. National Bureau of Economic Research. doi: 10.3386/w24287.

Elhauge, E. (2015) 'Horizontal Shareholding', Harvard Law Review, 129, p. 1267.

Gutiérrez, G. and Philippon, T. (2017) Declining Competition and Investment in the U.S. w23583. Cambridge, MA: National Bureau of Economic Research, p. w23583. doi: 10.3386/w23583.

Gutiérrez, G. and Philippon, T. (2018) How EU Markets Became More Competitive Than US Markets: A Study of Institutional Drift. w24700. Cambridge, MA: National Bureau of Economic Research, p. w24700. doi: 10.3386/w24700.

Guvenen, F., Ozkan, S. and Song, J. (2014) 'The Nature of Countercyclical Income Risk', journal of political economy, p. 40.

Heathcote, J., Perri, F. and Violante, G. L. (2010) 'Unequal we stand: An empirical analysis of economic inequality in the United States, 1967–2006', Review of Economic Dynamics, 13(1), pp. 15–51. doi: 10.1016/j.red.2009.10.010.

Jäntti, M. and Jenkins, S. P. (2010) 'The impact of macroeconomic conditions on income inequality', The Journal of Economic Inequality, 8(2), pp. 221–240. doi: 10.1007/s10888-009-9113-8.

Kalemli-Ozcan, S. et al. (2019) How to Construct Nationally Representative Firm Level Data from the Orbis Global Database: New Facts and Aggregate Implications. w21558. Cambridge, MA: National Bureau of Economic Research, p. w21558. doi: 10.3386/w21558.

Karabarbounis, L. and Neiman, B. (2014) 'The Global Decline of the labor share*', The Quarterly Journal of Economics, 129(1), pp. 61–103. doi: 10.1093/qje/qjt032.

Khan, L. and Vaheesan, S. (2017) 'Market Power and Inequality: The Antitrust Counterrevolution and Its Discontents', Policy Review, 11, p. 61.

Manning, A. (2011) 'Imperfect Competition in the Labor Market', in Handbook of Labor Economics. Elsevier, pp. 973–1041. doi: 10.1016/S0169-7218(11)02409-9.

Piketty, T. (2014) Capital in the twenty-first century. Cambridge, Mass.: Belknap Press of Harvard Univ. Press.

Shahrur, H. (2005) 'Industry structure and horizontal takeovers: Analysis of wealth effects on rivals, suppliers, and corporate customers', Journal of Financial Economics, 76(1), pp. 61–98. doi: 10.1016/j.jfineco.2004.01.001.

Song, J. et al. (2019) 'Firming Up Inequality*', The Quarterly Journal of Economics, 134(1), pp. 1–50. doi: 10.1093/qje/qjy025.

Wollmann, T. G. (2019) 'stealth consolidation: Evidence from an Amendment to the Hart-Scott-Rodino Act', American Economic Review: Insights, 1(1), pp. 77–94. doi: 10.1257/aeri.20180137.

Wozniak, A. K. (2007) 'Product Markets and Paychecks: Deregulation's Effect on the Compensation Structure in Banking', ILR Review, 60(2), pp. 246–267. doi: 10.1177/00197939070600020

A: Antitrust Policy Changes

Here I report a list of policy changes in countries included in the analysis. Moreover, I include the definition of Newly-Exempt and Never-Exempt mergers for every country.

- *United States*: In December 2000 an Amendment to the Hart-Scott-Rodino Act raised the threshold under which merging parties are exempt from reporting to the authorities (Wollmann (2019)). Before the Amendment, the threshold exempted all transactions in which assets of the acquired company were below 10 million USD. After the Amendment, all mergers whose transaction value was below 50 million USD were made exempt from reporting. As a consequence of this Amendment premerger notifications fell by 70%. For the context of this analysis, Newly-Exempt mergers are identified as those transaction with a value below 50 million USD, or with target sales between 10 and 50 million USD, or with target assets between 10 and 50 million USD. Never-Exempt mergers have a value greater than 50 million USD, or target sales above 50 million USD or target assets above 50 million USD.

- *Italy*: In 2012 requirements for merger filing were amended. Before 2012, a filing was required whenever acquirer sales exceeded 474 million euros or target sales exceeded 47 million euros. After the amendment both requirements must be met. Premerger notifications fell nearly 90%[58] in the subsequent year. Newly-Exempt mergers are identified as those transactions that satisfy one requirement, but not the other. Never-Exempt mergers, on the other hand, are those that satisfy both requirements.

- *Germany*: In May 1998 Amendments to the German Act Against Restraints of Competition raised the premerger reporting threshold. This threshold applies to the affected market turnover, and it was raised from 10 million DM to 30 million

[58]Notifications decreased from 400 to just 50.

DM. These changes were expected to lead to a two-third decline in premerger notifications, but they resulted in actual drop of 37%, likely due to an overall increase in merger activity[59]. The number of notifications dropped from 1888 in 1998 to 1182 in 1999[60]. For the purpose of this analysis it is not possible to identify affected markets, and as such a simplified analysis is conducted for the case of Germany, comparing only horizontal and non-horizontal mergers.

- *Spain*: In 2007 the National Commission was formed, and new regulations regarding notification were introduced. The two thresholds on global turnover were not affected by the policy change. Merging parties are required to report their transactions if global combined turnover is above 240 million euros, and at least two of the parties have a turnover of 60 million euros. What changed was the threshold on relevant product market shares, which was raised from 25% to 30%. The number of notifications to the antitrust authority steadily declined from 132 in 2006[61] to 58 in 2013[62] (a 55% decrease over seven years). For the purpose of this analysis it is not possible to identify affected markets, and as such a simplified analysis is conducted for the case of Spain, comparing only horizontal and non-horizontal mergers.

- *Belgium*: The merger reporting threshold was amended three times[63], and in order to avoid overlapping treatments, only the last amendment is considered in this analysis. In 2006 two merging reporting threshold were raised. The threshold on combined global revenues was raised from 40 million euros to 100 million euros, while the threshold on individual revenues of the parties was raised from 15 million euros to 40 million euros. In 1997 the number of notifications to the Belgian merger authority reached an all time high of 60. In 2007, at the end of its reforming period,

[59]See: Rudo, Joachim. "The 1999 Amendments to the German Act Against Restraints of Competition."
[60]See Bundeskartellamt's Activity Report 1999/2000, page 205.
[61]See Global Merger Control Manual by David J. Laing, Luis A. Gómez (2011)
[62]See "Merger Control" Fifth Edition by Uria Menéndez
[63]The amendments took place in 1995, 1999 and 2006

the notifications were 17 (a drop of about 70%)[64]. For the purpose of this analysis, Newly-Exempt mergers are those for which combined revenues were between 40 and 100 million, or target revenue were between 15 and 40 million, or buyer revenue were between 15 and 40 million. Never-Exempt transactions are defined as those that satisfy new requirements imposed by the amendments.

- *Sweden*: Before the 2000 pre-merger filing required that combined global revenue of merging parites exceeded 425 million euros. In 2000 a new threshold was introduced on top of the existing one. The target and buyer must also individually have domestic revenue of at least 11 million euros. The number of notified mergers decreased from 168 in 1999 to 84 in 2001, a decrease of 50%. For the sake of this analysis, Newly-Exempt mergers are those for which target revenue were below 11 million or buyer revenue were below 11 million. Never-Exempt mergers are defined as those for which target and buyer revenue are above 11 million euros.

- *Hungary*: Before 2005 premerger notification was required whenever combined global turnover of mergin parties exceede 10 billion HUF and at least two of the parties had turnover exceeding 500 million HUF. In 2005 the 10 billion HUF threhold was raised to 15 billion HUF. Between 2000 and 2005 the Hungarian authority received an average of 70 notifications per year, while it received an average of 42 notifications between 2006 and 2010 (a drop of about 40%). For the purpose of this analysis, then, mergers in which global combined turnover of the parties is between 10 and 15 billion HUF are considered Newly-Exempt. On the other hand, mergers that satisfy both requirements are considered Never-Exempt.

- *Canada*: In 2009 The threshold for mandatory merger notification on asset or revenue of the acquired party was raised from 50 million USD to 70 million USD for most kinds of transactions. The number of notifications to the Competition Bureau

[64]From "SETTING NOTIFICATION THRESHOLDS FOR MERGER REVIEW", a Report to the ICN Annual Conference, Kyoto April 2008

of Canada decresed from 236 in 2008-2009 to 216 in 2009-2010[65] (a decrease of about 9%). Newly-Exempt mergers are defined as those in which the target firm maximum between asset and revenue lays between 50 million USD to 70 million USD. On the other hand, Never-Exempt transactions are those for which the maximum between target firm asset and revenue is above 70 million USD.

- *Japan*: Before 2010 pre-merger notification was required whenever the buyer's revenue exceeds 10 billion Yen and the target's revenue exceeds 1 billion Yen. In 2010, the Amendment to the Antimonopoly Act revised the 1 billion Yen threshold upwards 400% to 5 billion Yen and revised the 10 billion Yen threshold upwards by 100% to 20 billion Yen. Notifications to the authorities dropped by 70% from 1000 in 2009 to 300 in the following years. Thereafter, Newly-Exempt transactions are those in which buyer's revenue is between 10 billion and 20 billion Yen or target's revenue is between 1 billion and 5 billion. Conversely, Never-Exempt mergers are transaction in which buyer's revenue exceeds 20 billion Yen and the target's revenue exceeds 5 billion Yen.

- *Russia*: In March 2005, amendments to antitrust law raised the pre-merger reporting thresholds 15-fold[66]. This was increased from 200 million RUB of combined asset value to 3 billion RUB of combined asset value. The number of considered applications declined by 48% to 6265 in the year 2005[67]. For the purpose of this analysis mergers in which the combined asset value of both parties is between 200 million RUB and 3 billion RUB are defined as Newly-Exempt. Mergers which combined asset value is above 3 billion RUB are considered Never-Exempt.

- *Brazil*: Before 2011 pre-merger filing requirements were based on whether the merg-

[65] According to the "Merger Review Performance Report" of April 12, 2012 this decrease was due to the Amendment to the reporting threshold.

[66] See: OECD. "OECD Economic Surveys: Russian Federation 2006." Vol. 2006/17. Paris, France.

[67] As reported in the "Report on the federal antimonopoly service on competition policy in 2005", which cites the amendment as the likely couse of the drop in applications.

ing parties' combined revenues would exceed 400 million Reais. In 2011 these requirements were amended so that a filing is required only when the larger party's revenues exceed 750 million Reais and the smaller party's revenues exceed 75 million Reais. For the sake of this analysis, Newly-Exempt mergers are those in which combined revenues is greater than 400 million Reais, but the larger party's revenues is smaller than 750 million Reais or the smaller party's revenues is lower than 75 million Reais. On the other hand, mergers in which larger party's revenues exceed 750 million Reais and the smaller party's revenues exceed 75 million Reais are considered Never-Exempt.

B: Effect of the Amendment on Research & Development

This section shows the results of a Cross-Country Diff-in-Diff analysis on the effect of changes in antitrust policy on Firms R&D expenditure. This analysis exploits differences between countries, as well as differences between sectors. Each observation represent R&D in an industry, in a specific country in a given year. Data on R&D come from the Capital Input Files of the September 2017 (Revised Juyly 2018) EU KLEMS release[68]. R&D levels are computed as the ratio of Real Total Research and Development and Real Total Assets[69], for each country, each industry and each year. R&D is divided by total assets so as to make values comparable across industries and countries. Data are included from 1995 up to 2015, because later years are not available in the KLEMS release. Among the 27 countries considered in the analysis, six experienced an Amendment to their policy that resulted in stealth consolidation: United States, Germany, Italy, Sweden, Spain, Hungary.

[68] The most recent release of the EU KLEMS database can be found at the following link: http://www.euklems.net/

[69] For the stock variables, Research and development is computed as Kq_RD/Kq_GFCF.

Equation (2.9) describes the Diff-in-Diff methodology. The dependent variable is the ratio between the stock of research and development and the stock of total assets. The dummy variable $I_{ij}^{Treated}$ is 1 for treated industries in treated countries. In this analysis all industries are considered as treated. The dummy variable I_t^{Post} is 1 for years after the Amendments in these countries. The coefficient of interest is β. The coefficient α_t represents year fixed effects, the coefficient θ_i stands for country fixed effects and γ_j stands for industry fixed effects. Results are reported in Table (2.20).

$$Inv_{ijt} = \beta I_{ij}^{Treated} I_t^{Post} + \alpha_t + \theta_i + \gamma_j + \eta X_{ijt} + \epsilon_{it} \tag{2.9}$$

The same exercise is repeated considering alternative sets of control variables X_{ijt}. If one is concerned that trends in industry level investment are driving the result, then one can add the interaction between industry an year fixed effects, as in column (4) of Table 2.20. In this case, only the variation between countries is identifying the estimated coefficient. The same reasoning can be applied to country trends in investment, and column (3) shows the coefficient identified by including the interaction of country and year fixed effects. Another concern might be that country specific levels of investment in some industries are driving the results. This can be accounted for by interacting country and industry fixed effects, as one can see in column (2). In no instance the coefficient β is estimated to be positive and significant, which implies that analyzed policy changes did not result in an increase of R&D expenditure.

C: List of countries in Gini Index dataset

Hereafter I report a table detailing all countries included in the study and their treatment status. Only countries that have at least 9 year of Gini coefficients are included in the analysis (9 is the number of available Gini observations for US).

VARIABLES	(1) All Countries	(2) All Countries	(3) All Countries	(4) All Countries
Itreat_Ipost	0.000196	0.000196	-0.0361**	0.000373
	(0.00296)	(0.00303)	(0.0161)	(0.00309)
Observations	8,779	8,779	8,779	8,779
R-squared	0.680	0.964	0.684	0.686
Cluster SE	C.try-Ind	C.try-Ind	C.try-Ind	C.try-Ind
Year FE	YES	YES	YES	YES
Country FE	YES	YES	YES	YES
Industry FE	YES	YES	YES	YES
Country-Industry FE		YES		
Year-Industry FE				YES
Year-Country FE			YES	

Table 2.20: Results of a difference-in-differences on firms R&D expenditure (measured as the stock of R&D investment divided by the stock of total assets) between industries in countries that experienced amendments to reporting thresholds and countries that did not. Robust Standard Errors are clustered at Country-Industry level.

TargetNation	Region	IncomeGroup	treat	treatment_year
Belgium	Europe and Central Asia	High income	1	2005
Canada	North America	High income	1	2009
Germany	Europe and Central Asia	High income	1	1998
Hungary	Europe and Central Asia	High income	1	2005
Italy	Europe and Central Asia	High income	1	2012
Spain	Europe and Central Asia	High income	1	2007
Sweden	Europe and Central Asia	High income	1	2000
United States	North America	High income	1	2000
Brazil	Latin America and Caribbean	Upper middle income	1	2011
Russian Federation	Europe and Central Asia	Upper middle income	1	2005
Austria	Europe and Central Asia	High income	0	.
Chile	Latin America and Caribbean	High income	0	.
Cyprus	Europe and Central Asia	High income	0	.
Czech Republic	Europe and Central Asia	High income	0	.
Denmark	Europe and Central Asia	High income	0	.
Estonia	Europe and Central Asia	High income	0	.
Finland	Europe and Central Asia	High income	0	.
France	Europe and Central Asia	High income	0	.
Greece	Europe and Central Asia	High income	0	.
Iceland	Europe and Central Asia	High income	0	.

Ireland	Europe and Central Asia	High income	0	.
Israel	Middle East and North Africa	High income	0	.
Latvia	Europe and Central Asia	High income	0	.
Lithuania	Europe and Central Asia	High income	0	.
Luxembourg	Europe and Central Asia	High income	0	.
Malta	Middle East and North Africa	High income	0	.
Netherlands	Europe and Central Asia	High income	0	.
Norway	Europe and Central Asia	High income	0	.
Panama	Latin America and Caribbean	High income	0	.
Poland	Europe and Central Asia	High income	0	.
Portugal	Europe and Central Asia	High income	0	.
Slovak Republic	Europe and Central Asia	High income	0	.
Slovenia	Europe and Central Asia	High income	0	.
Switzerland	Europe and Central Asia	High income	0	.
United Kingdom	Europe and Central Asia	High income	0	.
Uruguay	Latin America and Caribbean	High income	0	.
Bolivia	Latin America and Caribbean	Lower middle income	0	.
El Salvador	Latin America and Caribbean	Lower middle income	0	.
Honduras	Latin America and Caribbean	Lower middle income	0	.
Indonesia	East Asia and Pacific	Lower middle income	0	.
Kyrgyz Republic	Europe and Central Asia	Lower middle income	0	.
Moldova	Europe and Central Asia	Lower middle income	0	.
Mongolia	East Asia and Pacific	Lower middle income	0	.
Pakistan	South Asia	Lower middle income	0	.
Ukraine	Europe and Central Asia	Lower middle income	0	.
Vietnam	East Asia and Pacific	Lower middle income	0	.
Zambia	Sub-Saharan Africa	Lower middle income	0	.
Argentina	Latin America and Caribbean	Upper middle income	0	.
Armenia	Europe and Central Asia	Upper middle income	0	.
Belarus	Europe and Central Asia	Upper middle income	0	.
Bulgaria	Europe and Central Asia	Upper middle income	0	.
China	East Asia and Pacific	Upper middle income	0	.
Colombia	Latin America and Caribbean	Upper middle income	0	.
Costa Rica	Latin America and Caribbean	Upper middle income	0	.
Dominican Republic	Latin America and Caribbean	Upper middle income	0	.
Ecuador	Latin America and Caribbean	Upper middle income	0	.
Georgia	Europe and Central Asia	Upper middle income	0	.
Iran, Islamic Rep.	Middle East and North Africa	Upper middle income	0	.
Kazakhstan	Europe and Central Asia	Upper middle income	0	.
Kosovo	Europe and Central Asia	Upper middle income	0	.
Malaysia	East Asia and Pacific	Upper middle income	0	.

Mexico	Latin America and Caribbean	Upper middle income	0	.
Montenegro	Europe and Central Asia	Upper middle income	0	.
Paraguay	Latin America and Caribbean	Upper middle income	0	.
Peru	Latin America and Caribbean	Upper middle income	0	.
Romania	Europe and Central Asia	Upper middle income	0	.
Thailand	East Asia and Pacific	Upper middle income	0	.
Turkey	Europe and Central Asia	Upper middle income	0	.
Venezuela, RB	Latin America and Caribbean	Upper middle income	0	.

Chapter 3

Third paper: Stealth Consolidation, Market Power and Income Inequality

3.1 Introduction

Antitrust authorities foster competition so as to ensure that society surplus is shared among all agents in the economy. In the absence of competition, firms can exercise their market power so as to gain a larger share of this surplus. This redirects resources in the hands of few, at the expenses of many, and thus it generates an increase in inequality. There is, however, a lack of empirical evidence on the relationship between market power and income inequality, and this paper aims at filling this gap. In order to do so, this work devises a methodology that can be applied on publicly available data to identify exogenous variations in market power. This identification strategy exploits stealth consolidation in a dynamic factor model. Stealth consolidation, a concept introduced by Wollmann (2019), is defined as a plethora of anticompetitive deals that go unnoticed by antitrust authorities. These authorities tend to focus their attention and resources on large mergers. The rationale behind such policy is that small mergers are thought to have minor effects on market structure. Despite their unassuming size, however, these deals affect local or segmented markets where they can lead to duopolies or even monopolies. Stealth consolidation, then, provides a mean to identify exogenous variations in market power and their effect on the whole economy.

This work analyzes a determinant of income inequality that has not received much consideration in the empirical literature, despite the attention that inequality gathered recently (see Piketty (2014)). This paper shows that an increase in market power causes an increase in income inequality. Market power is defined as the ability of a firm to influence the market (e.g. by raising markups or lowering quantities) so as to gain more profits. It is therefore a chief candidate for increasing inequality. By increasing profits, firms are taking a larger portion of the wealth produced by society and they are redistributing it only to their shareholders. Since the richer part of the population owns a disproportionate share of firms capital, an increase in market power benefits the rich at the expenses of the poor. At the same time, monopsony power in the labor markets allows firms to lower wages and earnings, and this is particularly true for small and local labor markets, where earnings and productivity are already low[1]. These two mechanisms generate an increase in income inequality. The effect of stealth consolidation on market power, then, could explain part of the trend of increasing markups (De Loecker, Eeckhout, and Unger (2020)), raising profit share and falling labor share (Barkai (2019)), as well as the trend of increasing inequality (Heathcote, Perri, and Violante (2010)). This paper, therefore, fits into the debate on whether merger policy should be more or less restrictive, by showing that stealth consolidation can have far reaching consequences on the whole economy.

The lack of empirical work on the relationship between market power and income inequality is in part due to the lack of a unified dataset covering the whole economy. This paper contributes to the existing literature by providing a novel identification strategy to overcome such limitation and by showing evidence on the effect of market power on

[1] Berger et al. (2019) show that monopsony power is higher in small and concentrated labor markets. Manning (2011) claims that in low skilled labor markets it is more common wage-posting, a model of imperfect competition in which concentration decreases wages.

income inequality. Given a product market, an exogenous increase in market power affects the income of workers and owners of competing firms through different channels. It is not possible, however, to identify these workers and owners for any product market and any industry in the US. Consequently, one cannot construct a reliable control group from the population. As a way to deal with such data limitation, this paper proposes a novel method to identify exogenous variation in market power by using stealth consolidation. Then it leverages a large dimension dataset in a dynamic factor model so as to infer the effect of these exogenous changes on the whole economy, and ultimately on income inequality.

How can one identify a macroeconomic shock that captures stealth consolidation? This paper exploits variations over time between the number of horizontal and non horizontal mergers that are exempt from reporting under the Hart-Scott-Rodino Act. These are mergers whose transaction value is below a defined threshold and as such they will be called stealth mergers, as authorities have no way of detecting them. Horizontal mergers are defined as transactions between companies operating in the same industry, and thus they are most likely to decrease incentives to compete and to increase market power[2]. This interpretation is in line with the US Antitrus Authorities: since 1992 merger evaluation procedures are titled "Horizontal Merger Guidelines". The identification scheme relies on exogenous variation between horizontal and non horizontal mergers, and this variation over time is defined as a market power shock[3]. As a consequence a positive shock can be interpreted as an unexpected increase in the number of stealth horizontal mergers with respect to non horizontal ones[4]. Since horizontal mergers are the ones that have the potential to increase market power, this can be considered a macroeconomic shock

[2]As Wollmann (2019) reports, horizontal mergers are more likely to reduce rivalry.

[3]A market power shock is identified as the residual portion of the number of stealth horizontal mergers which is not explained by non horizontal ones. In particular, the market power shock is identified as the shock that moves only the number of stealth horizontal mergers on impact, but it does not move the number of stealth non horizontal mergers contemporaneously.

[4]Non horizontal mergers act as a control, by taking into account any factor that might influence merger activity (e.g. credit market conditions, expectations on future economic performance)

to market power. This is accomplished in a recursive identification scheme applied on stealth M&A activity and factors estimated in a Dynamic Factor Model[5], a methodology commonly referred to as Factor Augmented Vector Auto Regression (FAVAR).

In order to understand the effects of an increase in market power on the whole economy one needs to take into account all relationships between macroeconomic quantities, and this is possible through a Dynamic Factor Model[6] (DFM). These models, which can be considered an extension of the Vector Autoregressive (VAR) methodology to large datasets[7], condense common co-movement of variables by estimating unobservable factors[8]. Then standard time series methodologies are applied to these factors. Although short run fluctuations in inequality are relevant to understand part of the social cost of the business cycle, one is typically interested in long run changes in income distribution. Following the work of Barigozzi, Lippi, and Luciani (2016), long run features of the data are explicitly modeled in a Vector Error Correction (VEC) framework, which accounts for cointegration. This allows to draw conclusions on the short and on the long run consequences of an increase in market power.

In practical terms, this is an empirical work that exploits time sieries methods applied to large dimensional datasets, so as to provide evidence that an increase in market power generates an increase in inequality. Data on firms come from Compustat, a database of publicly traded firms in the US, with the aim of covering the largest cross section of firms available[9]. On top of that, data on mergers come from Thompson and Reuters SDC Platinum, which features transaction level data, as in Wollmann (2019). On the household side, census level data[10] allow to reconstruct the distribution of income, earnings, working

[5]Similarly to the work of De Giorgi and Gambetti (2017) on productivity and uncertainty shocks

[6]see Forni et al. (2009)

[7]By using DFM one can replicate existing results that were obtained using VAR methodologies and extend them to larger datasets. As an example, this work shows how to identify a technology shock a là Galí and Rabanal (2004) and shows that it increases inequality in the short run.

[8]These are linear combinations of variables built by using principal components.

[9]Recent development in the literature allow to compute measures of market power at the firm level, and reveal that aggregation obscures a large portion of the variability in such measures, as shown by De Loecker, Eeckhout, and Unger (2020).

[10]Data for households are taken from the Consumption and Expenditure Survey (CEX), a periodical

hours and consumption, and to study their evolution over time[11]. This gives much more insight than simply looking at inequality measures such as Gini indexes.

The identified shock raises M&A activity and lowers GDP on impact, as one would expect after an increase in market power, if efficiencies do not compensate for it. Moreover, the shock decreases TFP for several years, showing that mergers efficiencies are not generated immediately. The shock increases firms' markups, a key measure of market power, and high markup firms show a stronger rise[12]. Income and labor earnings decreases for poor households, but they increase for rich ones, which are able to share rents with firms. This directly generates an increase in inequality, which is reflected by the Gini index response.

Even more interesting is the effect of a shock to M&A activity in the long run. Output eventually increases, together with TFP, showing that mergers create efficiencies that need time to fully realize. This shows that antitrust policies should take into account both short and long run effects of perspective mergers, rather than focusing only on immediate outcomes. The shock increases also the share of output that goes into profits in the long run. Households income increases in the long run, driven by the increase in output, but rich households gain much more than poor ones. A similar pattern can be seen in earnings. This generates a permanent and significant increase in inequality, several years after the shock hit the economy. As a consequence, antitrust policy should take into account also the distributional implications of increases in market power.

The model of this work can be used also to quantify effects of a sudden increase in horizontal mergers. Wollmann (2019) shows that the Amendment to the Hart-Scott-Rodino Act in December 2000 caused an increase of about 320 horizontal M&A deals per year. This work shows that a shock of similar magnitude increases the income Gini index

survey conducted by the Bureau of Labor Statistics in US.

[11]see Heathcote, Perri, and Violante (2010) for a thorought description of the dataset

[12]In accordance with De Loecker, Eeckhout, and Unger (2020), who observe that high markup firms are the ones responding the most to macroeconomic shocks

by 1 Gini points in the long run[13]. In accordance with this result, it is shown through Error Variance Decomposition that the identified shock to market power accounts for 20% of the forecast error variance of Gini indexes in the long run. Moreover, through an historical decomposition of Gini index, this work shows that the identified market power shock accounted for an increase of about 0.4 Gini points between 2001 and 2006.

Changes brought at the macroeconomic level by the identified shock are shown to be relevant also at the indsutry level. In 20 out of 23 industries the shock increases the number of mergers, and in 15 industries the level of markups followes the response of M&A activity. Overall, the identified shock to market power concentrates a large portion of analyzed industries and it generates a widespread increase in markups. This shows that the shock has the potential to propagate throughout the majority of analyzed industries, and as such it can affect the entire supply side of the economy.

Lastly, this work provides several robustness checks of the main results. Alternative orderings and control variables are explored for the recursive identification strategy. The main results of this paper are shown to be robust to alternative ways of measuring merger activity[14] and alternative ways of measuring markups[15]. Moreover, this work explores a more agnostic identification procedure, based on Antolín-Díaz and Rubio-Ramírez (2018) narrative sign restrictions. This alternative identification scheme produces results that are very similar to the ones that obtain in the recursive framework. Overall, this sensitivity analysis provides evidence for a unique conclusion: the identified shock to market power increases income inequality, both in the short and in the long run.

[13] 1 Gini point is roughly equivalent to 2.5% of income Gini in 2001

[14] Measuring M&A activity with the number of deals, or with the value of such deals.

[15] Results hold when markups are computed with or without fixed costs in the production function. Results are robust even for markups computed using the Lerner Index.

Related Literature

This paper is clearly related to the work of Wollmann (2019) and it is meant to be an extension of its results to the whole macroeconomy, and eventually to income inequality. Wollmann (2019) studies an Amendment to the Hart-Scott-Rodino Act that raised the threshold under which merging parties were exempt from reporting their transactions to US authorities. By using a Diff-in-Diff identification strategy, Wollmann shows that the Amendment increased the number of horizontal mergers, raising concentration in the economy. The autor uses the term stealth consolidation to describe a widespread surge in small mergers that go under the radar of US authorities. Albeit small on paper, these transactions affect many local product markets, and they can increase significantly the level of market power in many sectors.

Furthermore, this work fits into the emerging literature that tries to apply rigorous time series methods to the evolution of inequality. The closest paper in this literature is De Giorgi and Gambetti (2017), in which the authors use the same data and a DFM to study the effect of a shock to productivity and uncertainty on consumption inequality at business cycle frequencies. The authors show that the identified shocks reduce consumption inequality on impact. Mumtaz and Theophilopoulou (2017) apply a VAR methodology, combined with sign restrictions, in order to study the effect of a monetary policy on income and consumption inequality in the short run. They use UK data an find that a contractionary monetary policy shock increases inequality. Olivier Coibion et al. (2017) answer a similar question for the US using a narratively identified monetary shock and local projections a là Jordà (2005). Anderson, Inoue, and Rossi (2016) study fiscal shocks in a VAR framework using narrative identification, and find that government spending shocks decrease consumption inequality in the US.

This paper contributes to this strain of literature on two dimensions. First, rather than studying the effect of standard macroeconomic shocks already identified in the literature, it provides a new strategy to identify a shock to market power. To the best of my knowledge, no other work uses data on M&A activity to identify a market power shocks in a time series framework. Second, rather than focusing on business cycle frequencies, my work models explicitly the cointegration structure of the data and provides significant evidence on long run effects on inequality. To the extent of my knowledge, no other paper applies cointegration methods to income inequality.

Other works try to relate firms activity to income inequality. Some are descriptive, such as the work of Song et al. (2019) who use a massive matched employer-employee database to ascertain firm contribution to the rise in earnings inequality. Others face the question from a theoretical point of view. Boar and Midrigan (2019) build a model with heterogeneous agents that act as entrepreneurs for heterogeneous firms, and show that size dependent subsidies can reduce markup dispersion and increase welfare[16]. Colciago and Mechelli (2019) build an heterogeneous agents model with oligopolistic competition, by embedding Cournot and Bertrand competition in an Aiyagary model. They find that lowering competition increases profits and inequality. Eggertsson, Robbins, and Wold (2018) modify a standard neoclassical model to show how an increase in market power can explain declining interest rates and labor share[17]. There are also papers in the Law and Economics literature on antitrust arguing that market power has an effect on inequality[18].

Notwithstanding the early work of Parker (2000) on why panel methods should not be applied to inequality, the previous empirical literature focused on panel data methods[19] and tried to find determinants of inequality in standard macroeconomic variables or in macroeconomic volatility. The empirical literature describing trends and features of inequality is flourishing, both for the US (Heathcote, Perri, and Violante (2010); Guve-

[16]Their model, however, relies on large publicly traded firms that redistribute profits to the whole population.

[17]Although Eggertsson, Robbins, and Wold (2018) work in a representative agent framework, the authors argue that market power can have sizable effects on inequality.

[18]See for instance Elhauge (2015) or Khan and Vaheesan (2017)

[19]See as an example Iyigun and Owen (2004), Breen and Garcia-Penalosa (2005) and Jäntti and Jenkins (2010)

nen, Ozkan, and Song (2014); Song et al. (2019)) and for other countries (Blundell and Etheridge (2010); Jappelli and Pistaferri (2010); Krueger et al. (2010)). The teoretical litearature on inequality is flourishing as well, thanks to the development of Heterogeneous Agents New Keynesian models (Bhandari et al. (2018); Kaplan et al. (2018)). On the other hand, the literature on markups and market power expands both theoretically (Edmond, Midrigan, and Xu (2018); Gutiérrez, Jones, and Philippon (2019)) and empirically (Nekarda and Ramey (2013); Blonigen and Pierce (2016); Galı, Gertler, and Lopez-Salido (2017); De Loecker, Eeckhout, and Unger (2020); Dıez, Fan, and Villegas-Sanchez (2019)). This work fits also into the literature of structural Dynamic Factor Models (Giannone, Reichlin, and Sala (2005), Forni et al. (2009)) and it applies the recently developed methodologies of Barigozzi, Lippi, and Luciani (2016) to estimate DFM on non stationary data.

The rest of the paper is structured as follows. Section 2 describes the dataset, comprising households, firms and macroeconomic variables. Section 3 covers the empirical strategy, and in particular it describes Factor Error Correction Models and it discusses identification. Section 4 reports the main results, as well as robustness checks. Section 5 concludes.

3.2 Data

Given the nature of this work, the dataset on which the analysis is conducted is large and diversified. Overall the time series dimension spans from 1980Q1 up to 2006Q4, and the data are on quarterly frequency[20].

[20]Current time series limitations are due mainly to availability of data on household variables. The main results hold qualitatively for an analysis run on an alternative dataset constructed from raw CEX data from 1980Q1 to 2012Q4.

3.2.1 Households

Data on households Income, Earnings, Consumption and Labor Hours are gathered from the Consumption and Expenditure Survey (CEX)[21]. It is a rotating panel of households that are selected to be representative of the US population[22]. Each household is interviewed for a maximum of four consecutive quarters. This survey reports, for the cross section of households interviewed, detailed demographic characteristics for all household members, detailed information on consumption expenditures for the three-month period preceding the interview, and information on income, labor earnings, hours worked, and taxes paid over a yearly period.

Overall the sample varies from about 2500 to about 4000 households interviewed in each quarter, and it is built to be representative of the whole US population. This allows for the computation of standard measures of inequality, such as the Gini coefficient, for each variable considered. Moreover, households are divided in income deciles, so that each decile contains a minimum of 240 households in 1980 and a minimum of 400 households in 2006. Then for each income decile one can compute the average Income, Earnings, Labor Hours and Consumption. On top of that. households can be divided by education level of the head[23]. Households are divided into those that have an education level lower or equal to High School, and those who have an education level greater or equal than College. Then for each group one can compute average Income, Earnings, Labor Hours and Consumption. These decile averages, together with averages by education and Gini indexes, will be the main dependent variables of the analysis.

[21]The dataset is thoroughly described in Heathcote et al. (2010) and it was used by previous work on Income and Consumption distributions such as Coibion et al. (2017), De Giorgi and Gambetti (2017) and Anderson et al. (2016).

[22]Continuous and reliable data are available only from the first quarter of 1980, which is the start of the analyzed sample.

[23]the oldest male, or the oldest female if no male is present

3.2.2 Firms

If income inequality is the main dependent variable of this study, then firms' market power is the main explanatory variable. There are several ways to measure market power, and this work explores a variety of them.

Disaggregated Markups

A first measure of market power comes from markups, defined as the ratio between sales price and marginal cost. Prices and sales are relatively easy to measure, but the challenge faced by researchers stands in the computation of marginal costs. This work follows the methodology of De Loecker, Eeckhout, and Unger (2020)[24] applied to Compustat data, which covers publicly traded firms starting in 1950[25]. The choice of the dataset is directed mainly at covering the widest possible array of firms and industries over the longest time horizon. Although publicly traded firms are not the majority of operating firms in US, they account for 41% of private sales and 29% of private US employment.

Appendix A reports details on both data and methodology used to compute markups. The researcher has to make explicit assumptions on the production function of firms, and this work assumes a translog poduction function. The production function is estimated under a further assumption of constant sector level elasticity of output to variable inputs θ_s^v. Markups are computed as the product between such elasticity and the ratio of output to variable input[26]:

$$\mu_{it} = \theta_s^v \frac{\text{Output}_{it}}{\text{Variable Input}_{it}}$$

Markups are computed at the firm level, but for the empirical analysis they are aggregated in two ways. First, they are divided in deciles, similarly to what is done for household data, and then for each decile it is computed the harmonic average weighted by sales.

[24]which was previously developed by De Loecker and Warzynski (2012)
[25]although reliable quarterly data are available only from the 80s'
[26]for more details, refer to the Appendix

Secondly, they are aggregated at the sector level, using 2-digit NAICS sectors so as to be consistent with the sector level elasticities.

Aggregated Measures

The recent literature produced also refined measures of aggregate market power and markups for the US economy. One of the first candidates is the Profit Share, which can be obtained from FRED and is defined as profit per unit of real gross value added of non-financial corporate business.

Another determinant of the level of market power in US is M&A Activity[27], which is captured by the dataset on Mergers and Acquisitions provided by Thompson Reuters SDC Platinum[28]. Wollmann (2019) uses the same database to asses the effect of an Amendment to the Hart-Scott-Rodino Act, which raised the threshold under which parties are exempt from reporting their transaction to the authorities. This work focuses on transactions that are exempt under the Hart-Scott-Rodino Act[29]. These transactions are labeled as stealth mergers, since merging parties are not required to report them to the authorities. Two main measures are constructed from this dataset: the number of stealth horizontal M&A deals, and the number of stealth non-horizontal M&A deals. Horizontal deals are defined as deals between firms in the same four-digit SIC industry, and are meant to capture deals regarding the same product markets. As such, horizontal deals represent anticompetitive increases in concentration, and they will be confronted with non horizontal deals in the identification of market power shock.

[27]Gutiérrez, Jones, and Philippon (2019) identify an Entry Cost shock that raises profits and concentration, and show that it correlates with M&A activity as measured from SDC. The shock Gutiérrez, Jones, and Philippon (2019) identify is conceptually very close to the one analyzed in this work, as the authors interpret it as decreasing competition.

[28]Blonigen and Pierce (2016) use the same dataset to show that M&As are associated with increases in average markups, but find little evidence for effects on productivity.

[29]meaning those transactions whose value is below 50 million dollars (in 2000 USD)

3.2.3 US Economy

As a control for the rest of the economy, several other macroeconomic variables are introduced into the dataset. In particular this work uses 101 macroeconomic series in levels provided by Barigozzi, Lippi, and Luciani (2016)[30]. Moreover, the dataset includes measures of Total Factor Productivity (TFP) provided by Fernald (2014) at quarterly frequency, and adjusted for utilization.

3.3 Empirical Strategy

In order to take full advantage of the extensive dataset available for this work a Large Dimensional Dynamic Factor Model (DFM) is the natural choice. These models became popular in the econometric and macroeconomic literature in the early 2000s and have been successfully used for policy analysis based on impulse response function[31]. DFMs represent the intuition that all variables in the economy are driven by few common macroeconomic shocks, with their residual component being idiosyncratic.

DFMs allow the researcher to infer causal relationships from a very large pool of time series data, exploiting the common movement of macroeconomic series. In doing so, they rely on a small set of assumptions that are clearly stated in the definition of the model. Although DFMs can be considered an extension of Vector Auto Regressive methods (VAR), they do not suffer from the issue of non-fundamentalness, which arises when agents form choices based on expectations of future variables, and makes VAR impossible to estimate[32]. Besides, DFM can be used to replicate VAR analysis[33].

[30] This wide dataset well approximates the information set available to a large institution, for instance a central bank.

[31] see as reference Giannone, Reichlin, and Sala (2005), Stock and Watson (2005), Forni et al. (2009), Forni and Gambetti (2010), Barigozzi, Conti, and Luciani (2014) and De Giorgi and Gambetti (2017), whose methodology is the closest to this work

[32] see Forni, Gambetti, and Sala (2014)

[33] Appendix D shows that one can identify a technology shock a là Galí and Rabanal (2004) using a DFM applied to the dataset of this paper. In accordance with the work of De Giorgi and Gambetti (2017), a technology shock increases income inequality. On top of that, the technology shock increases

Moreover, aternative methods such as panel regressions are affected by the issue of spurious correlation[34] that arises whenever dependent and independent variables are not stationary, which is the case for most of the variables describing income inequality. On the other hand, recent results of Barigozzi, Lippi, and Luciani (2016) show that DFMs can be successfully applied to non-stationary data by imposing and error correction structure on the factors. This allows the researcher to make reliable inference at all frequencies, and especially in the long run.

3.3.1 Factor Error Correction Models

At the core of a DFM there is the assumption that the data x_{it} can be decomposed into the sum of two unobservable components, the common component χ_{it} and the idiosyncratic component ξ_{it}

$$x_{it} = \chi_{it} + \xi_{it}$$

Where $i \in \{1...N\}$ represents the cross section and $t \in \{1...T\}$ represents the time series dimension. A further assumption of any DFM is that the common component of each variable i is a linear combination of r common factors $F_t = (F_{1t}, ..., F_{rt})'$:

$$\chi_{it} = \lambda_{i1} F_{1t} + ... + \lambda_{ir} F_{rt} = \lambda_i F_t$$

Where the vector $\lambda_i = (\lambda_{i1}, ..., \lambda_{ir})$ is called factor loading of variable i. These represent the weight that is given to each factor in determining the common component of variable i.

Most of the literature on DFM considers stationary data, and thus it imposes a simple VAR structure on the factors $C(L)F_t = Ru_t$, where L represents the lag operator, and u_t are the structural shocks that drive the whole system. Given the non-stationary nature of

firms' markup, and high markup firms show a stronger increase.

[34]see Parker (2000)

x_{it}, which implies non-stationarity of the factors F_t, the results of Sims, Stock, and Watson (1990) show that the parameters of VAR in levels on the factors are consistently estimated. Nonetheless, Phillips (1998) shows that in the presence of cointegration long run IRF are consistently estimated only if one models the long run properties of the system, i.e. within a Vector Error Correction Model (VECM). Therefore this work imposes a VECM specification on the factor dynamics, so as to reliably estimate IRF also in the long run:

$$G(L)\Delta F_t = h + \alpha\beta' F_{t-1} + Ru_t \tag{3.1}$$

Where $G(L)$ is a matrix of lag polynomials, h is a vector of constants, α is a r x c matrix of weights and β is the r x c matrix describing c cointegrating relationships between factors. Lastly R is a r x r rotation matrix that rotates the shocks u_t so as to achieve identification.

By defining:

$$x_t = (x_{1t}, ..., x_{Nt})', \ \chi_t = (\chi_{1t}, ..., \chi_{Nt})', \ \xi_t = (\xi_{1t}, ..., \xi_{Nt})', \ \Lambda = (\lambda_1, ..., \lambda_N)'$$

One can write the Factor Error Correction Model (FECM) that is used in this work:

$$\begin{aligned} x_t &= \chi_t + \xi_t = \Lambda F_t + \xi_t \\ G(L)\Delta F_t &= h + \alpha\beta' F_{t-1} + Ru_t \end{aligned} \tag{3.2}$$

Barigozzi, Lippi, and Luciani (2016) show that factors F_t and factor loadings Λ of non stationary data can be consistently estimated by using principal components, the standard tool of DFM. Given the sample convariance $\hat{\Gamma} = T^{-1}\Delta x \Delta x'$, one can compute the n x r matrix $\hat{W}$ containing the r eigenvectors of $\hat{\Gamma}$ corresponding to the r largest eigenvalues of $\hat{\Gamma}$. Then the estimated factors and factor loadings are:

$$\hat{\Lambda} = \sqrt{N}\hat{W}, \quad \hat{F}_t = \frac{1}{\sqrt{N}}\hat{W}x_t = \frac{1}{N}\hat{\Lambda}'x_t$$

With regard to the estimation of the Impulse Response Function of a VECM one can refer to Lütkepohl (2006). Suffices to say that, given the number of lags p, one has to estimate the matrix polynomial[35]:

$$\hat{A}^{VECM}(L) = I_r - \sum_{k=1}^{p+1} \hat{A}_k^{VECM} L^k$$

So that the IRF of the VECM in the factors is:

$$IRF^{VECM} = \left[\hat{A}^{VECM}(L)\right]^{-1} R$$

The matrix R is an orthogonal rotation matrix with the properties $RR' = I$ and $det(R) = 1$ that serves as identification matrix. As such it is chosen by the researcher to identify a desired shock.

Once IRF for the factors F_t has been estimated, one can easily construct IRF for the variables $x_t = \Lambda F_t + \xi_t$ thanks to the factor loadings Λ:

$$IRF^{FECM} = \hat{\Lambda}\left[\hat{A}^{VECM}(L)\right]^{-1} R$$

Therefore the Impulse Response of valriable i to shock j, as identified by column r_j of

[35]With coefficients given by:

$$\begin{aligned}
\hat{A}_1^{VECM} &= \hat{G}_1 - \alpha\beta' + I_r \\
\hat{A}_k^{VECM} &= \hat{G}_k - \hat{G}_{k-1}, \quad k = 2, ..., p \\
\hat{A}_{p+1}^{VECM} &= -\hat{G}_p
\end{aligned}$$

matrix R, can be written as:

$$IRF_{ij}^{FECM} = \hat{\lambda}_i \left[\hat{A}^{VECM}(L)\right]^{-1} r_j$$

Proposition 1 of Barigozzi, Lippi, and Luciani (2016) proves that, as $N, T \to \infty$ such Impulse Response is consistent. The authors prove consistency also for an IRF computed using VAR in levels on the factors, but note that such consistency holds only for finite horizons, and as such long run IRF based on VAR in levels are no longer consistent.

One last step that is needed for the estimation of the model is to determine the number of static factors r and the number of unit roots τ for the VECM dynamics. With regard to r, Bai and Ng (2002) devise the standard test that is commonly used in the literature. The test applied to the extended dataset indicates a number of static factors $r = 7$. Barigozzi, Lippi, and Luciani (2016) extend the test of Hallin and Liška (2007) and devise a test for the number of shocks driving the data[36]. Similarly to the results of Barigozzi, Lippi, and Luciani (2016), this works finds $\tau = 1$[37].

For the results presented in this work, $r = 7$ static factors and factor loadings are estimated using principal components on data in levels. Relevant variables, including GDP and deciles for income, earnings, hours and consumption, are detrended using a linear trend[38]. Then, the estimation proceeds onto a VECM with $\tau = 1$ unit roots. Identification is conducted as described in the following section.

3.3.2 Identification

The problem of identification for DFMs amounts to finding an appropriate rotation matrix R for the shocks u_t, similarly to identification in standard VAR settings. This allows to

[36] The test is based on the number of diverging eigenvalues of the spectral density of Δx, called $\Sigma(\theta)$. When one considers the eigenvalues of $\Sigma(0)$, the spectral density at long run frequency 0, then the test selects the number of common trends τ.

[37] For more details on the testing procedure, refer to Appendix B.

[38] As recommended in Barigozzi, Lippi, and Luciani (2016). Variables that are not detrended include Gini indexes as well as markups.

identify the structural shocks $\epsilon_t = Ru_t$. Before such rotation, the shocks u_t have little interpretation, and the choice of R determines the shape of Impulse Response Functions. The aim of this work is to identify a shock to market power, and describe its effects on income inequality.

Stealth M&A

Mergers and Acquisitions (M&A) are commonly associated with an increase in market power. The merger between two firms operating in the same sector increases concentration and reduces incentive to compete, as well as increasing the ability of incumbents to prevent new entry. Blonigen and Pierce (2016) show that M&As increase markups without increasing productivity or efficiency, a clear sign of increasing market power. Gutiérrez and Philippon (2017) use discrete jumps in concentration following large M&A to identify changes in competition. In their recent work Gutiérrez, Jones, and Philippon (2019) use a calibrated DSGE model together with maximum likelihood methods to estimate a shock to entry costs. The authors claim that such a shock represents variations in competition, and show that it has had a significant effect on macroeconomic dynamics over the past 30 years. As further evidence, the authors show that their shock correlates well with M&A activity.

By treating M&A activity as a signal of increasing market power, one can use it as a known factor[39] and study how it influences the dynamics of the other factors driving the economy. This can be done by using a standard recursive identification scheme. This is a well known methodology in the literature on structural DFMs, as one can see in De Giorgi and Gambetti (2017), and usually it takes the name of Factor Augmented Vector Acutoregression (FAVAR).

M&A Activity correlates with many macroeconomic variables, and as such there is a clear endogeneity issue whenever one tries to use it to identify a shock to market

[39]Using a measure of market power external to the model is reminiscent of narrative identification methods used in empirical VAR on monetary and fiscal policy (see Romer and Romer (2010) for example).

power. In order to control for anything that might affect merger dynamics, such as credit market conditions, this work exploits the difference between horizontal and non-horizontal mergers. Horizontal M&A are defined as transactions between companies operating in the same narrowly defined industry, and as such they increase concentration and can raise market power in a determined product market[40]. The identified market power shock is defined as a change in the number of horizontal mergers with respect to the number of non horizontal ones. Therefore, this identification scheme relies on exogenous variation over time between horizontal and non horizontal mergers, and this variation is defined as a market power shock.

On top of that, this work will focus on those mergers that usually go under the radar of the US authority, and as such they are referred to as Steath Consolidation by Wollmann (2019). These are transaction whose value is below the threshold defined by the Hart-Scott-Rodino Act. This threshold was revised in December 2000 up to 50 million USD by and Amendment to the same act. Mergers that are small enough to be under this threshold are not required to report to the US authorities, and as such they are not screened for potential anticompetitive outcomes[41]. However, several of these transactions affect local markets, and in several cases they are mergers to duopoly or even monopoly in the relevant product markets. This work intends to show that such mergers can have a significant effect on the whole economy, and ultimately on inequality.

This identification strategy can be implemented by considering both the number of Stealth Horizontal Mergers and Stealth Non-Horizontal Mergers as known factors. Then one can use the number of Stealth Non-Horizontal Mergers (NH_t) as a control for the number of Stealth Horizontal Mergers (H_t). In practice this amounts to constructing a vector containing the two variables and the static factors $FF_t = (F_t, NH_t, H_t)$ and

[40]The US competition authorities recognize the possible anticompetitive nature of horizontal transactions. So much so that they have formally tiled their merger evaluation procedures "Horizontal Merger Guidelines".

[41]Due to their size, these mergers are considered to be harmless by the Hart-Scott-Rodino Act.

running a VECM on this vector[42]. As a consequence the Impulse Response Functions will be:

$$\begin{aligned} IRF^{VECM} &= \left[\hat{A}^{VECM}(L)\right]^{-1} R \qquad \text{for M\&A Deals} \\ IRF^{FECM} &= \hat{\Lambda}\left[\hat{A}^{VECM}(L)\right]^{-1} R \qquad \text{for all other variables} \end{aligned}$$

Where the identification matrix R is the Cholesky factor of the covariance matrix of residuals:

$$R = chol(u_t u_t') = \begin{bmatrix} x_{11} & 0 & 0 & \dots & 0 \\ x_{21} & x_{22} & 0 & \dots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ x_{r1} & x_{r2} & x_{r3} & \dots & x_{rr} \end{bmatrix}$$

This identification scheme relies on imposing restrictions on the contemporaneous response of certain variables to the identified shocks. In particular the researcher imposes zero restrictions, which imply some variables do not react contemporaneously to some shocks. Each row of the rotation matrix R corresponds to a variable, while each column corresponds to a shock. The last column of the matrix describes the identified market power shock, which is defined as the shock that affects only the number of Stealth Horizontal Mergers H_t contemporaneously. All zeros in the last column represent the fact that no variable other than H_t responds contemporaneously to the identified shock. As a consequence all other variables will react in the subsequent periods in response to movements in H_t caused by the shock itself.

[42] The linear equations of this VECM can be described by:

$$G(L)\Delta FF_t = h + \alpha\beta' FF_{t-1} + Ru_t$$

$$H_t = h_m + \alpha_{1,m} NH_t + \alpha_{2,m} F_t + \beta_m H_{t-1} + \gamma_m NH_{t-1} + \delta_m F_{t-1} + ... + \underbrace{u_{1,t} + u_{2,t} + ... + u_{M,t}}_{\epsilon_t} \tag{3.3}$$

The last row of R describes the crucial equation for identification of the market power shock, the equation of Horizontal M&A Activity as described in (3.3). This identification structure allows to decompose the residual ϵ_t into several shocks $u_{i,t}$. The market power shock $u_{M,t}$ is part of the residual of a regression of current H_t on current NH_t, the factors F_t and past variables. One can then interpret the shock $u_{M,t}$ as the part of Horizontal $M\&A$ Activity which is not explained by Non Horizontal $M\&A$ Activity or other shocks in the economy, with the addition of further controls in the form of past realizations of $M\&A$ variables and all other variables represented by the factors F. The presence of other shocks $u_{1,t}, ..., u_{8,t}$ in the equations ensures that the identified market power shock $u_{M,t}$ is not capturing other sources of variation, such as technology shocks[43]. Therefore this shock represents unexpected and unforcastable variations in the number of horizontal mergers. An example of such exogenous variations is the Amendment to the Hart-Scott-Rodino Act of December 2000, which raised the number of horizontal mergers by about 320 per year, as it is shown by Wollmann (2019). This methodology in a FAVAR setting has been already applied in the literature, as one can see from the identification of an uncertainty shock in De Giorgi and Gambetti (2017), and it can be applied also in a Factor Vector Error Correction setting.

Discussion

This identification strategy exploits exogenous variations in M&A activity over the available time span and correlates them with all variables in the dataset, so as to infer their effects on the whole economy. One then might ask: there is a readily available exogenous

[43]The identified shock $u_{M,t}$ is orthogonal to all other shocks $u_{1,t}, ..., u_{8,t}$ identified in this model. This is a property fo Cholesky identification schemes.

change, the Amendment of December 2000, why not use that as an identification device? Why not use a Diff-in-Diff methodology such as Wollmann (2019)? In order to apply such identification strategy to income inequality, one would need to devise a control and a treatment group, dividing the population in those that are affected by the Amendment and those that are not. Unfortunately this is not possible because census data on households provides no information regarding the company or the sector in which interviewed people are employed. Moreover, there is no information regarding the type of assets held by these households.

An exogenous increase in the number of horizontal mergers raises concentration and market power, and as a consequence it increase profits for owners of affected companies and decreases earnings for workers. Unfortunately it is not possible to identify such owners and workers, and therefore it is not possible to devise a treatment and a control group. Instead, this work identifies a series of exogenous changes in M&A activity and shows what effect they have on the whole households distribution of earnings and income, and ultimately on inequality. This identification is clearly weaker than an event study which exploits the Amendment, but it is a way to leverage the whole time series and the large dimensional dataset to overcome limitations in the data.

Given the aforementioned limitations, this work cannot directly infer the effects of the Amendment. This antitrust policy change provides an important source of variation, though. Indeed the identified market power shock is affected by the Amendment, especially in the first quarter of 2001 and then in the years following it. This work, however, can only infer the effects of an exogenous change in the number of Stealth Horizontal M&A, which is what followed the Amendment in December 2000.

One of the main concerns regarding the identified shock is that it is not capturing an increase in market power, but something else. Part of the increase in markups that is documented in the recent literature has been explained by technological changes[44].

[44] I will not go into details regarding such changes, so as to leave the argument as broad as possible

These changes are likely to affect also the market structure, and thus firms' incentive to merge. On the other hand, they can also be linked with income inequality. Therefore, one is lead to think that technological changes could explain the co-movement of markups, mergers and inequality. The identified shock, however, captures a change in the number of horizontal mergers with respect to non-horizontal ones. Therefore, it is not enough that the proposed technological change affects merger incentives. It should also affect horizontal mergers differently with respect to non-horizontal ones, which act as a control for overall merger activity.

Alternatively, the identified shock might be capturing horizontal mergers reacting to changes in non-horizontal ones. In particular, a reduction in non-horizontal merger activity might provide incentive for more horizontal transactions. If this were the case, then the identified shock would not capture stealth merger activity, but exactly the opposite: standard merger activity. In equation (3.3) that identifies the shock, however, the lagged term NH_{t-1} ensures that the shock is not driven by past realizations of non-horizontal mergers. This alternative explanation could hold only if horizontal mergers reacted to non-horizontal ones within one quarter[45], which is quite unrealistic.

With regard to other macroeconomic shocks that might be driving the results, the Cholesky identification strategy comes into play. The residual ϵ_t of equation (3.3) is decomposed into orthogonal shocks $u_{1,t}, ..., u_{M,t}$. Out of these, shocks $u_{1,t}, ..., u_{7,t}$ are the ones affecting the factors contemporaneously, and thus they represent any macroeconomic shock driving all variables represented by the factors. Since factors F_t are a good approximation of all the information available in the system, these shocks capture any movement other than merger activity. This is captured by $u_{8,t}$, which is the shock affecting only horizontal and non-horizontal mergers contemporaneously. The last one, $u_{M,t}$, is the identified market power shock affecting only horizontal mergers, which is residual with respect to

[45]The frequency of the data is quarterly. This argument would be stronger if one were using yearly data.

all the others[46].

The properties of the factors allow to answer a further concern regarding the identification equation (3.3): omitted variable bias. There are many determinants of inequality, and market power is just one of them. Therefore one might think that some key variables explaining the evolution of inequality are missing from equation (3.3). As explained before, however, factors are meant to represent all available information in the economy, and thus they account for any possible variation that might drive explained variables. As an example one could take outsourcing to lower-wage countries, which is likely to increase inequality by compressing unskilled labor earnings. The dataset from which factors are built contains all components of GDP including export and import balance, industrial production by product category, as well as labor market indicators such as unemployment and number of employed by industry. These variables, and the relative factors, carry the information representing the process of outsourcing. Therefore, the factors F_t act as an effective control for outsourcing to lower-wage countries.

3.4 Results

The main results of this work concern the identification strategy based on shocks to M&A Deals. Variations of such identification strategy, as well as the identification strategy based on narrative sign restrictions are included in the robustness section.

3.4.1 Main Results

In order to better understand the identified shocks to M&A Deals, it is useful to inspect its effect on some relevant macroeconomic variables. The first panel of Figure 3.1 shows that the shock has a positive and long lasting effect on the number of Stealth Horizontal M&A deals. With regard to the size of the initial shock, standard procedure in the

[46]Morever, the shock $u_{M,t}$ is orthogonal to all the other shocks

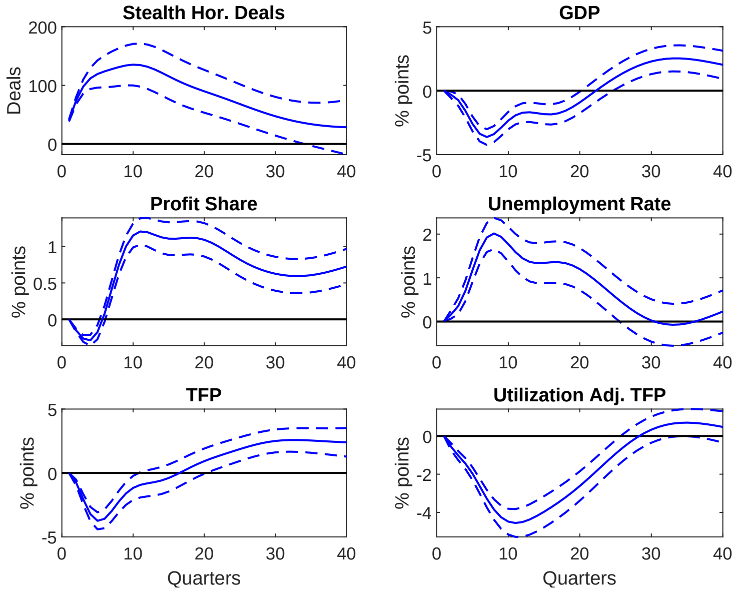


Figure 3.1: Impulse Response Function of macro variables to a shock of M&A Deals. Confidence bands are computed by bootstrap methods on the process generating factors. Bands represent one standard deviation (or 68%) of the distribution of bootstrapped IRFs.

literature is to set the initial shock at time 0 equal to the standard deviation of the relevant variable[47]. An alternative and more conservative procedure is to use the size of a known exogenous shock. Wollmann (2019) estimates that the Amendment increased the number of Horizontal Mergers by about 3200 mergers in 10 years. Such an effect can be attained with an initial shock of about 40 mergers at time 0, which will be the normalization used for the following results. This shock, however, cannot be interpreted as the direct effect of the Amendment, but simply as the effect of an exogenous occurrence of 40 Stealth Horizontal Mergers. The identified shock is self reinforcing, and it is strongly persistent, as one can see from the first panel of Figure 3.1. This persistence will then drive most of the long run results of this work.

The shock has a negative and significant effect on the GDP[48], as one would expect after an increase in market power absent immediate mergere efficiencies. The negative impact on GDP ensures that the shock is not identifying a merger wave which is due

[47]In this case, the standard deviation is about 65 mergers per quarter.

[48]GDP is measured in log points, so the effect of the shock is to reduce GDP by as much as 3% in tow years after the shock hits the economy.

to favorable economic conditions. About five years after the shock, its effect on output become positive, which can mean that mergers eventually result in an increase in output, but require a long time to realize efficiency gains that offset increases in market power. The third panel of Figure 3.1 shows that the share of output that goes to profit responds positively in the long run. On impact it does not move much, but it increases steadily after one year. In particular the shock raises the profit share by about 1% in the long run. Unemployment follows a path that is specular with respect to the one of GDP. When the shock hits, there is an increase in unemployment. However, following the increase in output, the positive effect on unemployment becomes smaller and eventually non-significant.

As a way to get further insights into possible merger efficiencies resulting from the identified shock, one can inspect the IRF of Total Factor Productivity (TFP). From Figure 3.1 it is clear that TFP follows a path that is very similar to the one of GDP, decreasing in the short run and increasing afterwards. This provides further evidence that merger efficiencies take several years to realize. The conclusion changes, however, if one looks at utilization adjusted TFP (last panel of Figure 3.1). If one takes into account the change in utilization of factors of production, the negative effect on productivity is even stronger, and it never turns significantly positive[49]. Therefore the output increase in the long run is due to the use of spare capacity rather than productivity gains. These results show that antitrust policy should consider not only contemporaneous effects of mergers, but also its effects in the years following the transactions, as efficiencies do not realize immediately.

One key measure of market power is firm level markup. A shock to market power is expected to increase markups, and Figure 3.2 shows the response of markup distribution. In this figure, IRF of all 10 deciles are reported, so as to appreciate the eventual spreading of the distribution in response to the shock. Color grading starts with darker lines for the

[49]Such a negative impact on productivity is in accordance with the previous result of Blonigen and Pierce (2016), who show that mergers increase markups without generating productivity gains.

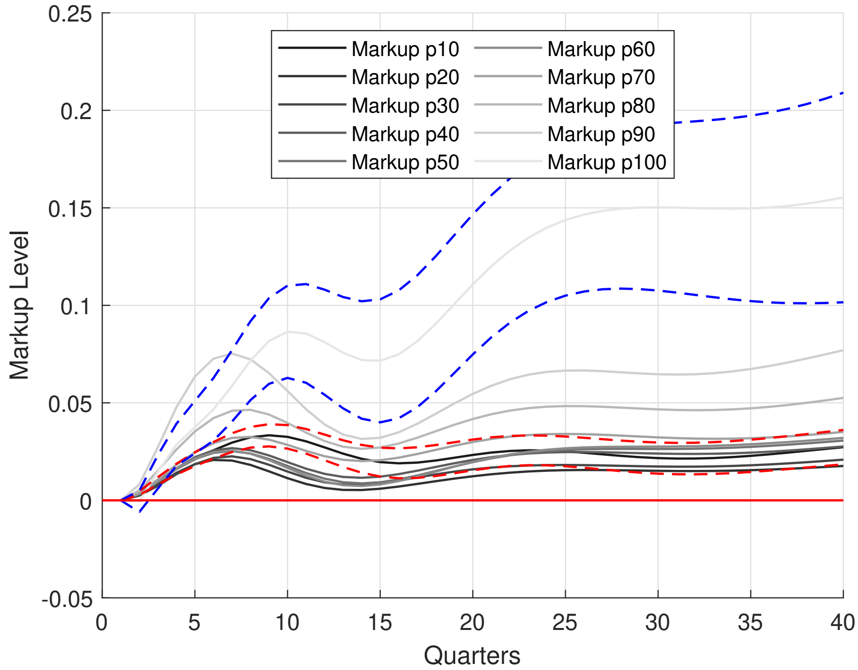


Figure 3.2: Impulse Response Function of the distribution of firms markups to a shock of M&A Deals. Confidence bands for the 10th decile are reported in blue and for the 1st decile are reported in red. Confidence bands are computed by bootstrap methods on the process generating factors. Bands represent one standard deviation (or 68%) of the distribution of bootstrapped IRFs.

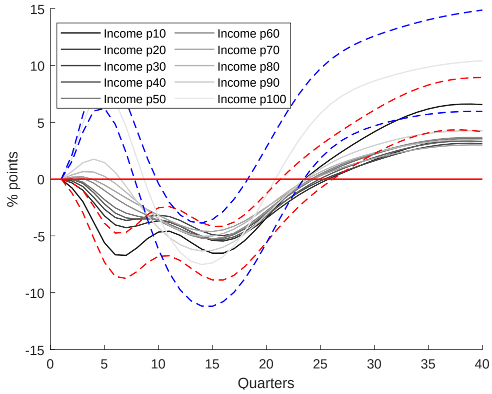
lower deciles and lighter lines for the upper ones. Confidence bands are reported only for the 1st and 10th deciles, so as to give an impression of the overall significance, without creating too much visual noise. As one can clearly see, the response is positive for all deciles both in the short and in the long run. As a consequence the entire distribution of markups is shifting upward. But this is not all, the shape of the distribution is changing, as the upper end of the distribution is more responsive to the shock than the lower deciles[50]. Firms in the top decile increase their markup by 15% in the long run, while firms in the lower deciles increase it by less than 2%. The result of Figure 3.2 implies that markups are increasing after a shock to the number of horizontal M&A Deals, and the distribution of markups is spreading and becoming more unequal.

[50]This result agrees with De Loecker, Eeckhout, and Unger (2020), who observe that the upper end of markups distribution explains a large portion of the variation of markups, and thus it is the one that is driving the recent increase in average markups.

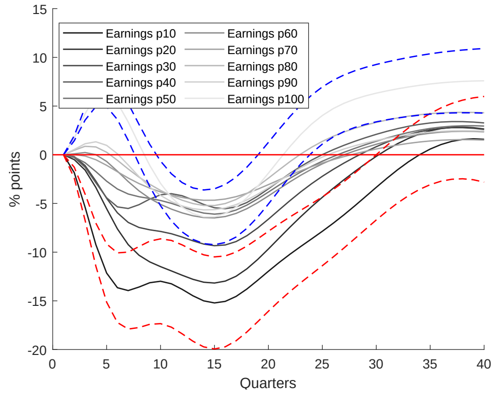
Once the shock to M&A has been described and characterized, one can study its effects on households inequality. Figure 3.3 shows the response of distributions of Income and Earnings, as well as the response of these variables by education groups. Starting from Figure 3.3a, the IRF of Income, one can clearly see that most of the distribution is loosing income after the shock, apart from the highest deciles which shows a significant gain in income both in the short and in the long run. One particular feature of this figure is the ordering of the IRF. Lower deciles, the poor, loose more income than the rich, and the ordering of deciles is almost perfectly preserved. This can be clearly interpreted as an increase in income inequity, driven mainly by losses suffered by the poor. The income distribution spreads substantially in the short run, in the first two to five years, but the effect of the shock can be seen also in the long run, when the income distribution shifts upward, driven likely by the increase in GDP. Still, the upper end of the distribution is gaining almost 10% of income, more than the lower end, generating permanent changes in inequality. Although merger efficiencies are generating higher output, they have no effect on inequality, which remains higher also in the long run. Figure 3.3c shows the response of income by education categories. Again, the disadvantaged are loosing more than the College educated in the short run. In the long run, however, there is no significant difference in the two responses.

Figure 3.3b shows response of the distribution of Labor Earnings. The overall pattern is similar to the response of income, since most of households income is composed by labor earnings. Earnings are substantially decreasing after the shock, and the poor are affected the most[51]. The richest gain both in the short and in the long run. Again, one can see a clear ordering of earning deciles responses, meaning the richer households are gaining more from the shock than poorer households, contributing to increasing inequality. A similar picture is drawn in Figure 3.3d, where one can see that the different response of college educated and non-college educated can explain some portion of the difference in

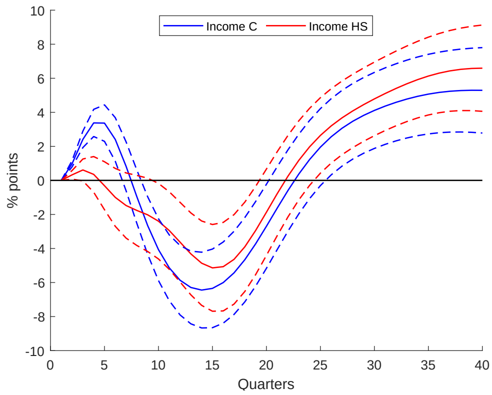
[51] the lowest decile looses 15% of earnings in the short run



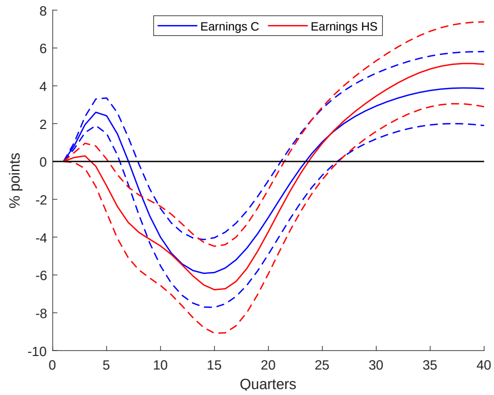
(a) Income by Deciles



(b) Labor Earnings by Deciles



(c) Income by Education



(d) Earnings by Education

Figure 3.3: Impulse Response Function of the distribution of Hoouseholds variables to a shock to M&A Deals. Confidence bands for the 10th decile are reported in blue and for the 1st decile are reported in red. Confidence bands are computed by bootstrap methods on the process generating factors. Bands represent one standard deviation (or 68%) of the distribution of bootstrapped IRFs.

earning deciles only in the short run.

Another way to study the effect of a market power shock is to inspect the responses of standard measures of inequality, such as Gini Indexes. This serves also as a robustness check on previous results, since Gini coefficients are computed on the same disaggregated data, but their response to a shock is not the same as Gini index computed on responses of deciles. Figure 3.4 shows IRF for Gini coefficients, and clearly it implies that the identified shock increases inequality. This is true for both income and earnings in the short run, where the sharp increase in Gini indexes for income and earnings mirrors the dynamics of their distributions. After a rapid increase in the first two years, the effect on income and earnings inequality is persistent, and remains significant in the long run. It is worth stressing again that IRFs of Gini indexes are perfectly consistent with IRFs of income and earnings deciles, providing further support to previous results. This feature does not arise by construction[52]. Their coordinated response, then, is a clear sign that a shock to market power does increase income inequality, and this change is a long lasting one.

Relevance of the identified shock

With regard to the magnitude of the increase in inequality, one can see from Figure 3.4 that after about one year the shock generates an increase of about 1.5 Gini points for income, while in the long run the shock increases income by 1 Gini point. This is the simulated effect of an exogenous occurrence of 40 Stealth Horizontal M&A. To put things in perspective, the data show a Gini index for household income of about 38.6 for 2001Q1, which implies that a shock of this size can increase Gini index for income by about 2.5% in the long run.

Another insightful way to assess the magnitude of the identified market power shock is Error Variance Decomposition[53]. From Figure 3.5 one can see that the market power shock is an important driver of income and earnings Gini index in the long run, accounting for roughly 20% of their common component variance. Moreover, is the most relevant component of variance of horizontal M&A Deals, and it is a significant driver of both GDP and profit share, especially in the long run[54].

A further way to quantify the effect of the market power shock is to construct an

[52] All these variables are part of the dataset and are driven by the same factors, but with variable specific factor loadings.

[53] This exercise decomposes the forecasting error variance for the common component of each variable in the portion that is explained by each shock. Red bars represent the portion explained by the identified shock to M&A, while green bars represent the portion explained by the shock driving the variable that is used as main control in the identification equation, that is the number of Stealth Non-Horizontal Mergers. Blue bars represent all the other shocks moving the economy, which are not of particular interest for this analysis.

[54] In particular the makret power shock can explain about 30% of profit share variation.

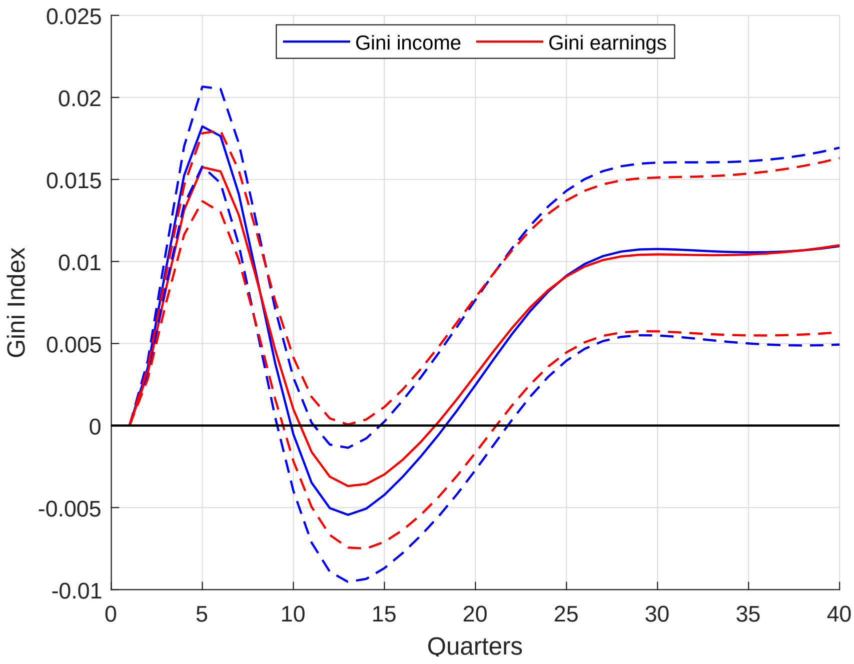


Figure 3.4: Impulse Response Function of Gini index for Income, Earnings and Consumption to a shock of M&A Deals. Confidence bands are computed by bootstrap methods on the process generating factors. Bands represent one standard deviation (or 68%) of the distribution of bootstrapped IRFs.

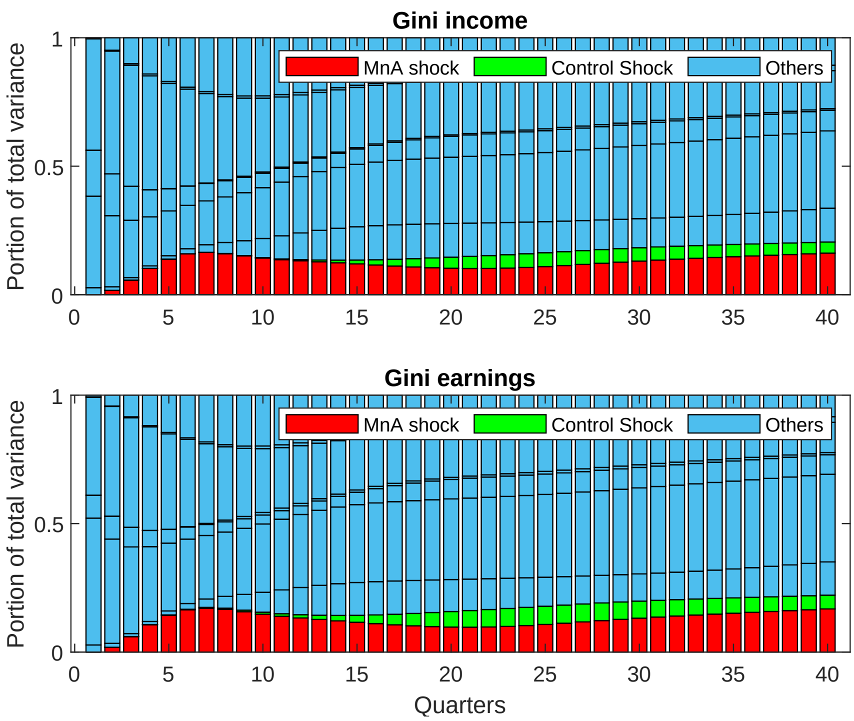


Figure 3.5: Error Variance Decomposition of the common component of several variables. The shock of interest is the M&A Shock, represented by red bars. The other shocks are the ones driving factors and thus the rest of the variables in the dataset

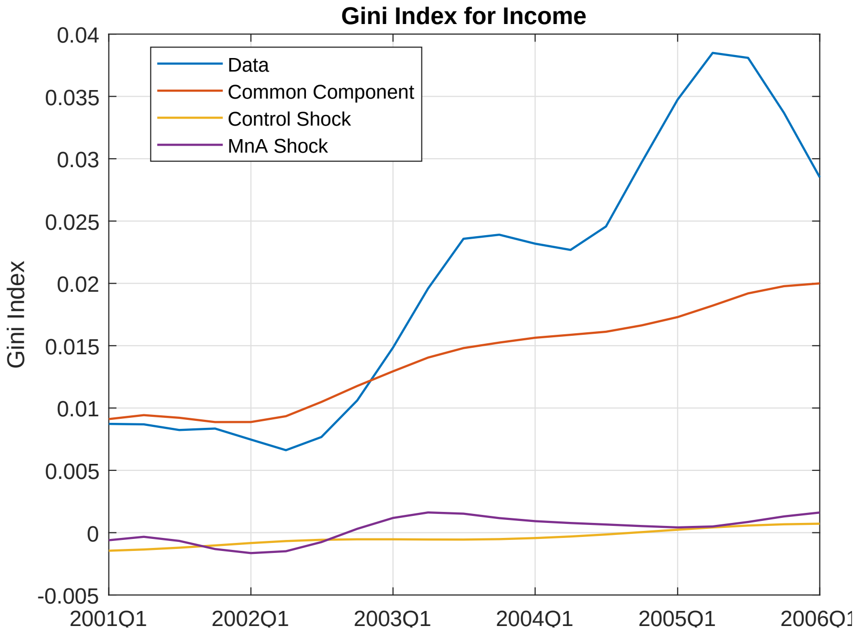


Figure 3.6: Historical Decomposition of Gini Index for household income. The variable is centered, so that the level of the vertical axes is not relevant. The scale is preserved, though, meaning that Gini Index increased by roughly 2 Gini points from 199Q1 to 2006Q4. Common Component represents the portion of the variable which is explained by the model, while M&A Shock represent the portion of the variable which is explained by the identified shock.

historical decomposition[55]. Figure 3.6 shows the historical decomposition of Gini Index for Income. The variable is centered, so that the level of the vertical axes is not relevant. The scale is preserved, though, meaning that Gini Index increased by roughly 2 Gini points from 2001Q1 to 2006Q4. The red line shows that the common component of the model, what the whole model can explain, accounts for most of the increase in Gini income. This ensures that the model is capturing most of inequality variation. The common component is the sum of various pieces, including the contribution of each shock. The purple line shows the contribution of the identified market power shock, and it is clear that the shock contributed to an increase of about 0.4 Gini points in the five subsequent years after 2001.

[55]This exercise consists in decomposing a variable in the parts that are explained by the various pieces of the model, and in particular the parts explained by each identified shock. This procedure is applied only on the common component of the variables, which is reported in red.

Industry level analysis

In order to further understand the channels through which the identified shock to market power propagates to the whole economy, one can assess the effects that it had on each industry. This analysis, given its very nature, can only comprise industry level variables, as it is not possible to relate a certain part of the population to a single industry. Therefore this section will focus on merger activity, markups level and concentration (as measured by the HHI) within each industry[56].

This analysis comprises 23 industries with enough firms and merger activity. For each industry the number of Stealth Horizontal M&A deals, the average markup and the HHI are added to the dataset[57]. The response of other variables does not change significantly, and the same goes for the main results presented until now. Figure 3.7 shows the impulse response function of M&A Deals, markups and concentration in each industry. In 20 industries out of 23 the number of horizontal M&A increases as a consequence of the shock, showing that the identified market power shock generates increases in merger activity across the whole economy[58]. On top of that, in 15 industries out of 23 the response of markups tracks the response of M&A Deals, showing that the two series move in the same direction following the shock. An example could be Mining, Quarrying, and Oil and Gas Extraction, a sector in which the shock generated an increase of about 3 mergers per quarter and an increase of about 0.1 in markups. For this industry, also the concentration level, as measured by the HHI, increases in response to the shock.

Six industries out of 23 show an inverse pattern, where M&A Activity and markups respond in opposite directions[59]. All these industries show an increase in the number of

[56]Industries are defined as two-digit NAICS levels, so as to have enough mergers in each quarter to generate variation and so as to have enough firms in each industry to compute reliable measures of markups.

[57]With regard to the identification strategy, these industry level variables play the role of additional controls driven by the factors.

[58]Notable exemptions are Non-Metal Manufacturing; Professional, Scientific, and Technical Services; and Other Services, except Public Administration.

[59]This is the case for Utilities, industry NAICS 44 and 45 (Retail Trade), Transportation and Infor-

M&A but a decrease in markups[60]. Interestingly, for these industries the HHI responds in a similar way to the markups, decreasing after the shock, even if the number of mergers is increasing. One possible interpretation is that in these industries the increased merger activity is more than compensated by entry of new competitors. This results in lower concentration and lower markups.

Further inquiry into the reasons behind different industry dynamics is beyond the scope of this work. This section is meant to provide evidence on the fact that macroeconomic trends deriving from the identified market power shock can still be found in a large portion of US industries. In particular, the shock generated an increase in the number of horizontal M&A deeals, concentrating these industries, and it raised industry level markups. As a consequence the identified shock made these industries more concentrated, less competitive and more profitable.

mation.

[60]In a similar way, Other Services saw an increase in markups and a decrease in M&A activity in the short run, but then in the long run markups decreased, following M&A actovity.

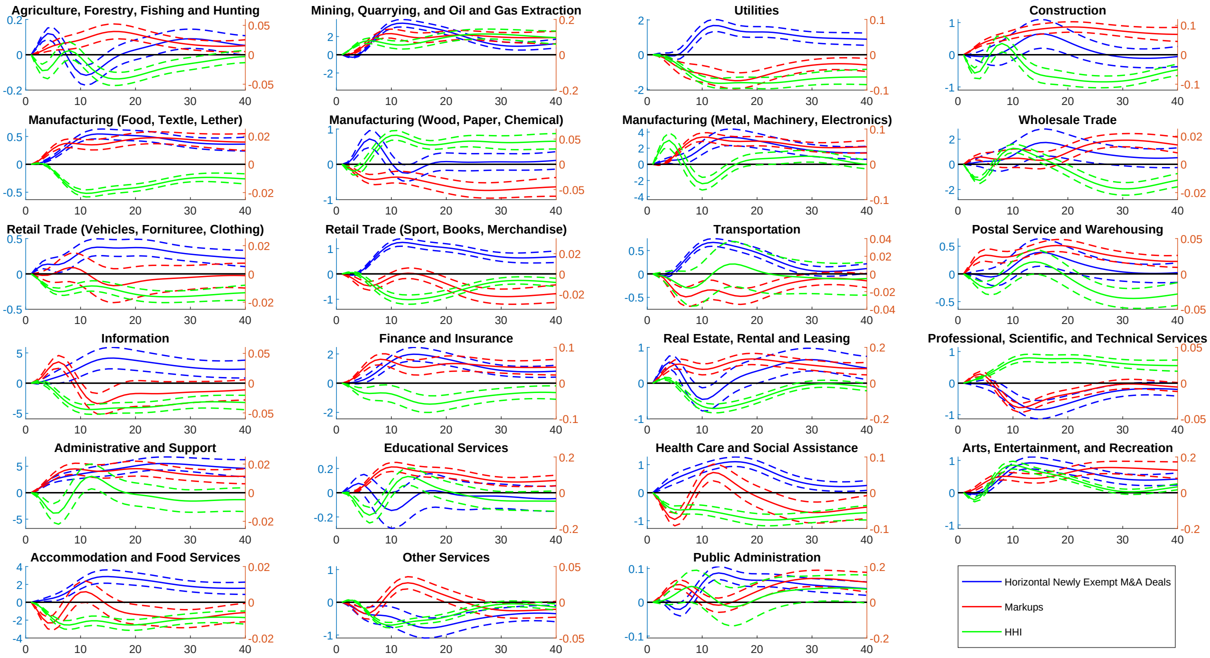


Figure 3.7: Industry level Impulse Response Function of Horizontal Newly Exempt M&A Deals, Markups and HHI to a shock of M&A Deals. Confidence bands are computed by bootstrap methods on the process generating factors. Bands represent one standard deviation (or 68%) of the distribution of bootstrapped IRFs.

3.4.2 Robustness

The first kind of robustness analysis that needs to be performed with a Cholesky identification strategy concerns the ordering of variables used in the recursive scheme[61]. It can be argued that the ordering chosen for the main result is the one imposing more stringent restrictions on the desired shocks, since a zero restriction are imposed on the instantaneous responses of all variables except H_t[62]. A radically different ordering would require M&A Deals to be ordered first, so that the VECM is run on the vector $FF_t = (NH_t, H_t, F_t)$[63]. This identification ordering is usually employed for "slow moving" variables, such as technology, since it implies that the identified shock is the only one affecting M&A instanta-

[61] It is well known that such identification produces orthogonalized shocks that are easy to interpret, but it is often required that results are robust to different orderings of relevant variables.

[62] As a consequence the identified shock affects H_t contemporaneously, but it cannot affect other variables.

[63] In this case the shock of interest is the one pertaining to the second column of the Cholesky matrix R.

neously. As a matter of fact, this identification scheme does not require that all variables, through the factors, act as controls in the equation for H_t, where the shock of interest is $u_{M,t}$.

$$H_t = h_m + \alpha_{1,m} NH_t + \beta_m H_{t-1} + \gamma_m NH_{t-1} + \delta_m F_{t-1} + ... + u_{1,t} + u_{M,t}$$

Figure **??** in the Appendix shows that the response of macroeconomic variables does not change much with respect to the previous identification strategy, especially in the long run. The response of markups does not change as well, as one can see from Figure **??** in the Appendix. The response of income and earnings distributions is not appreciably different, supporting the same patterns of the main results. Lastly, Figure **??** shows the IRF of Gini indexes, similarly to the ones reported in Figure 3.4. Patterns agree in the two figures, and noticeably the magnitude of the responses do not change, showing that conditioning on more variables does not affect the main result. The same goes for the historical decomposition, as one can see from Figure **??** in the Appendix.

One then could use different variables to measure M&A activity. Unfortunately the dataset provided by Thompson and Reuters lacks information regarding transaction values and firms assets for the majority of M&As. As a consequence a shock identified by using these measures is to be considered less reliable. Nevertheless in the Appendix one can find IRF for a shock identified using M&A Deal Value, rather than the number of Deals. As one can see from Figure **??** in the Appendix, the effect of a shock identified using all M&A deals has a similar effect on macroeconomic variables. The shock, however, is not calibrated, so that it is not possible to make considerations on its magnitude. Regardless, this shock has a similar effect on markups (Figure **??**) and on Gini indexes (Figure **??**).

A further sensitivity check can be performed by changing the way firm markups are computed. As explained in Appendix A, firm level markups are computed following the methodology of De Loecker and Warzynski (2012) and accommodating for the presence of fixed costs[64]. Figure **??** in the Appendix shows that response of macroeconomic variables does not change substantially, with the exception of TFP, which ramains always positive. Utilization adjusted TFP decreases in response to the shock, though. Markups estimated without fixed costs increase after the shock, as shown in Figure **??**. The same can be said for Income and Earnings Gini Indexes (Figure **??**). An even less refined measure of markups is given by the Lerner Index[65]. Results computed employing this measure do not differ substantially woth the main result, and are not reported here.

Rather than changing a controlling variable, one could change the variables that are used to identify the shock. One could simply use the number of Horizontal M&A Deals as identifying variable, and GDP as control. The resulting equation would be:

$$H_t = h_m + \alpha_{1,m} GDP_t + \alpha_{2,m} F_t + \beta_m H_{t-1} + \gamma_m GDP_{t-1} + \delta_m F_{t-1} + ... + u_{1,t} + ... + u_{M,t}$$

This strategy allows to identify a M&A wave which is not justified by favorable economic conditions, as GDP is assumed to not respond contemporaneously to the shock by construction. Figure **??** in the Appendix shows that this is the case, and that other variables respond similarly to the main result. The main difference is in the scale of the response, as an increase of 3200 Horizontal Deals is less relevant than an increase of 3200 Stealth Horizontal Deals[66]. Notwithstanding the scale of the IRF, the response of markups and Gini index are very similar to the main results, as one can see from Figure **??** and Figure **??**. This shows that not only Stealth Consolidations effects on inequality, but potentially any increase in horizontal mergers can. As a consequence the results of this paper extend also to larger M&A transactions.

Rather than a standard recursive identification procedure, one could try a more ag-

[64]De Loecker, Eeckhout and Unger (2020) show that markups computed without accounting for fixed costs are significantly higher.
[65]Which is nothing more than the ration between $(P - MC)$ and the price P, and it can be computed at the firm level.
[66]This is merely due to the fact that Stealth Deals are a subset of all Deals.

nostic one based on Antolín-Díaz and Rubio-Ramírez (2018) Narraive Sign Restrictions[67]. The response of macroeconomic variables was calibrated in the same way as the main result, so as to make them comparable. Such impulse responses of macroeconomic variables can be seen in Figure **??**. Similarly to the shock to M&A deals, the response of GDP is negative only in the short run, and it turns positive afterwards. Unemployment mirrors the path of GDP, increasing in the short run and decreasing afterwards, so as to accommodate the increase in output. Total factor productivity decreases in the short run, explaining the pattern of GDP.

The response of markup distribution is again similar to the one observed to the shock to M&A Deals (Figure **??**), with the high end of the distribution clearly increasing more than the lower end. Figure **??** shows the response of income and earnings distributions. The pattern of these variables shows a remarkable similarity with the main results, but confidence bands show less significance in the long run. Nonetheless, the response of inequality, as measured from Gini indexes, is positive both in the short and in the long run, as one can see from Figure 3.8, and with a magnitude that is similar to the main result. Furthermore the error variance decomposition shows that the shock identified with sign restrictions explain about 50% of the forecast error variance of income Gini index[68]. Overall, results from this alternative Sign Restriction scheme support the main ones, showing that a shock to market power identified with M&A Activity has a significant and long-lasting effect on income inequality.

[67]Such identification scheme relies on imposing restrictions on the sign of impulse responses of certain variables, as well as the historical decomposition of those variables, and it is explained in Appendix C.

[68]The shock of interest, meaning the one on which sign restrictions are imposed, is shown in red in Figure **??**.

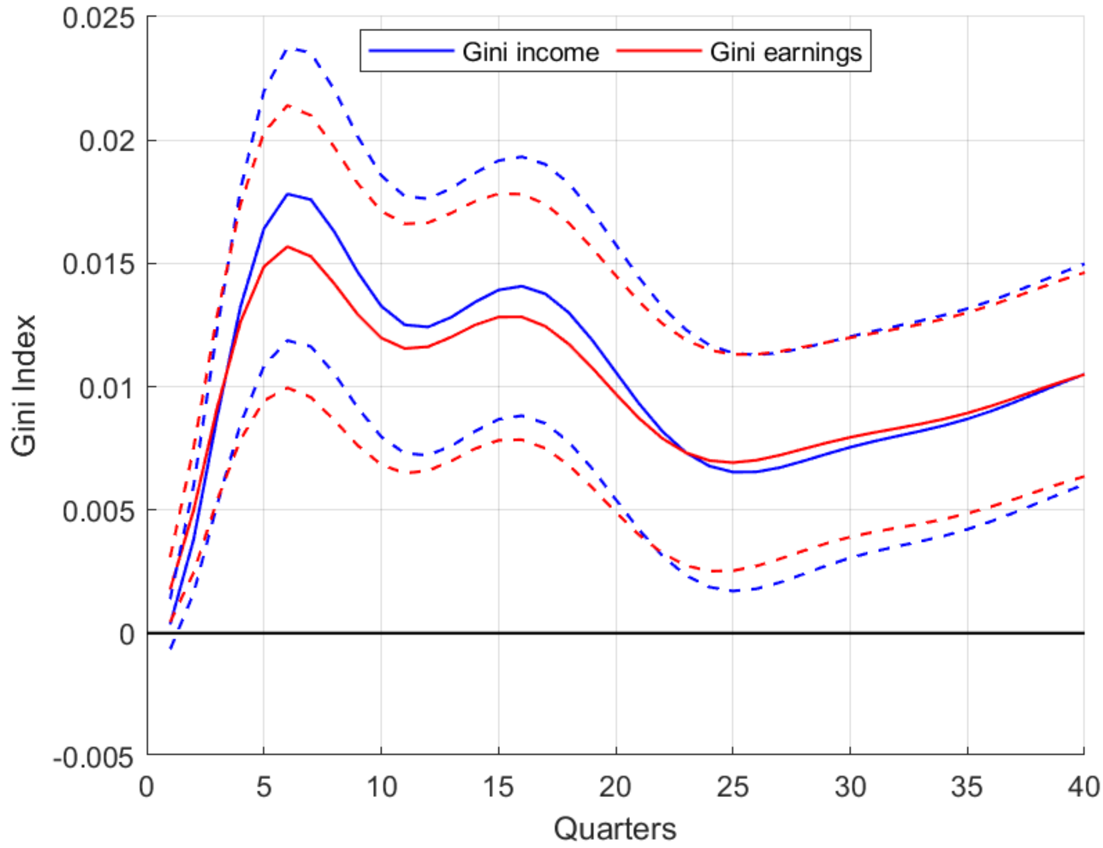


Figure 3.8: Impulse Responion of Gini indexes for Income and Earnings to a shock identified using Sign Restrictions. Confidence bands are computed by bootstrap methods on the process generating factors. Bands represent one standard deviation (or 68%) of the distribution of bootstrapped IRFs.

3.5 Conclusions

Inequality has recently reached the center of political and academical debate because of the dire consequences it can have on people's life and society as a whole. It increased at an alarming rate in past decades, changing our societies dramatically. But what are the determinants of income inequality? What can explain an increase in inequality? Which mechanisms play a role in changing income distribution? This paper studies one of such mechanisms: the effect of stealth consolidation and market power on income inequality. Stealth consolidation refers to anticompetitive mergers that go under the radar of antitrust authorities due to their unassuming size. Such mergers, however, can have significant effects on market power in segmented and local markets. Market power is the ability of firms to manipulate the market so as to increase their profits. This favors firm owners at the expenses of firm workers, and thus it has the potential to raise income inequality.

This work applies a Dynamic Factor Model to a large US dataset so as to model the whole economy and derive the effect of exogenous changes in market power. The dataset combines CEX survey data on heterogeneous households with Compustat data on heterogeneous publicly traded firms and Thompson Reuters data on M&As. It applies cointegration time series methods to such a large set of variables thanks to the Dynamic Factor Model. The shock to market power is identified by exploiting differences between horizontal stealth M&As and non-horizontal stealth M&As. This shock to market power decreases output and total factor productivity on impact. It has a positive effect on firm markups, especially for firms in the upper tail of markup distribution. The shock increases income and labor earnings inequality in the short run, and this is mainly due to an earnings loss for the poor.

In the long run the effect on output is positive, thanks to merger efficiencies that take several years to fully realize. The level of Market Power is changed permanently, thanks to the strong persistence of the shock. As a consequence the share of output that goes into profits increases in the long run. Notwithstanding an increase in output in the long run, also the effect on income and earnings inequality is permanent. The identified market power shock increases the income Gini index by 1 Gini points in the long run, or an increase of about 2.5% of income Gini. Error variance decomposition shows that the identified shock to market power accounts for 20% of the forecast error variance of Gini index in the long run. Moreover, an historical decomposition of Gini index shows that between 2001 and 2006 the identified market power shock accounted for an increase of about 0.4 Gini points.

Results of this work show how all agents in our complex societies are intertwined.

Starting from the 80s, both inequality and market power began to rise, and this increase continues to this day. These trends have already been paired together countless times, so as to answer the same question of this paper: does market power increase inequality? This work shows compelling evidence of the causal effect that stealth consolidation can have on income inequality, and it provides insights into mechanisms driving this causal link.

The empirical results of this work can be the basis of several new strands of research. The acquisition of highly innovative start-ups by the tech giants has attracted the attention of antitrust authorities around the world. Since these transaction involve small parties, many tend to fall below reporting threshold, and can concur to stealth consolidation. One possible extension of this paper can focus on innovation effort by entrants, incumbents and acquired firms, by exploiting the existing threshold. This could shed some light on these high tech acquisitions are anticompetitive and detrimental to innovation. Moreover, the recent work of Cunningham, Ederer, and Ma (2019) describes the phenomenon of killer acquisitions. The authors document that in the pharmaceutical sector incumbents acquire potential entrants with the sole purpose of discontinuing competing products. These acquisitions tend to cluster just below reporting threshold. Studies exploiting these thresholds could provide further evidence on the practice of acquisition aimed at eliminating rivals. Lastly, the empirical findings of this work justify the study of a theoretical model featuring both heterogeneous agent and an antitrust authority, so as to study the effects that merger policy can have on resource allocation and redistribution.

References

Anderson, E., Inoue, A., Rossi, B., 2016. Heterogeneous Consumers and Fiscal Policy Shocks. Journal of Money, Credit and Banking 48, 1877–1888. https://doi.org/10.1111/jmcb.12366

Antolín-Díaz, J., Rubio-Ramírez, J.F., 2018. Narrative Sign Restrictions for SVARs. American Economic Review 108, 2802–2829. https://doi.org/10.1257/aer.20161852

Bai, J., Ng, S., 2002. Determining the Number of Factors in Approximate Factor Models. Econometrica 70, 191–221. https://doi.org/10.1111/1468-0262.00273

Barigozzi, M., Conti, A.M., Luciani, M., 2014. Do Euro Area Countries Respond Asymmetrically to the Common Monetary Policy? Oxford Bulletin of Economics and Statistics 76, 693–714.
https://doi.org/10.1111/obes.12038

Barigozzi, M., Lippi, M., Luciani, M., 2016. Non-Stationary Dynamic Factor Models for Large Datasets. arXiv:1602.02398 [stat].

Barigozzi, M., Luciani, M., 2017. Common factors, trends, and cycles in large datasets. arXiv:1709.01445 [stat].

Barkai, S., 2019. Declining Labor and Capital Shares. SSRN Journal. https://doi.org/10.2139/ssrn.3

Berger, D., Herkenhoff, K., Mongey, S., 2019. Labor Market Power (No. w25719). National Bureau of Economic Research, Cambridge, MA. https://doi.org/10.3386/w25719

Bhandari, A., Evans, D., Golosov, M., Sargent, T.J., 2018. Inequality, Business Cycles, and Monetary-Fiscal Policy (Working Paper No. 24710). National Bureau of Economic Research.
https://doi.org/10.3386/w24710

Blonigen, B.A., Pierce, J.R., 2016. Evidence for the Effects of Mergers on Market Power and Efficiency 36.

Blundell, R., Etheridge, B., 2010. Consumption, income and earnings inequality in Britain. Review of Economic Dynamics 13, 76–102. https://doi.org/10.1016/j.red.2009.10.004

Boar, C., Midrigan, V., 2019. Markups and Inequality (No. w25952). National Bureau of Economic Research, Cambridge, MA. https://doi.org/10.3386/w25952

Breen, R., Garcia-Penalosa, C., 2005. Income Inequality and Macroeconomic Volatility: An Empirical Investigation. Rev Development Economics 9, 380–398. https://doi.org/10.1111/j.146 9361.2005.00283.x

Colciago, A., Mechelli, R., 2019. Competition and Inequality: Aiyagari meets Bertrand and Cournot (SSRN Scholarly Paper No. ID 3312919). Social Science Research Network, Rochester, NY.

Coroneo, L., Giannone, D., Modugno, M., 2016. Unspanned Macroeconomic Factors in the Yield Curve. Journal of Business Economic Statistics 34, 472–485. https://doi.org/10.1080/0735001

Cunningham, C., Ederer, F., Ma, S., 2019. Killer Acquisitions 106.

De Giorgi, G., Gambetti, L., 2017. Business cycle fluctuations and the distribution of consumption. Review of Economic Dynamics 23, 19–41. https://doi.org/10.1016/j.red.2016.07.005

De Loecker, J., Eeckhout, J., Unger, G., 2020. The Rise of Market Power and the Macroeconomic Implications*. The Quarterly Journal of Economics 135, 561–644. https://doi.org/10.1093/qje/qjz041

De Loecker, J., Warzynski, F., 2012. Markups and Firm-Level Export Status. American Economic Review 102, 2437–2471. https://doi.org/10.1257/aer.102.6.2437

Dıez, F.J., Fan, J., Villegas-Sanchez, C., 2019. Global Declining Competition 41.

Edmond, C., Midrigan, V., Xu, D.Y., 2018. How Costly Are Markups? (No. w24800). National Bureau of Economic Research, Cambridge, MA. https://doi.org/10.3386/w24800

Eggertsson, G.B., Robbins, J.A., Wold, E.G., 2018. Kaldor and Piketty's Facts: The Rise of Monopoly Power in the United States (Working Paper No. 24287). National Bureau of Economic Research.
https://doi.org/10.3386/w24287

Elhauge, E., 2015. Horizontal Shareholding. Harv. L. Rev. 129, 1267.

Fernald, J., 2014. A Quarterly, Utilization-Adjusted Series on Total Factor Productivity (No. 2012–19). Federal Reserve Bank of San Francisco.

Forni, M., Gambetti, L., 2010. The dynamic effects of monetary policy: A structural factor model approach. Journal of Monetary Economics 57, 203–216. https://doi.org/10.1016/j.jmoneco

Forni, M., Gambetti, L., 2014. Sufficient information in structural VARs. Journal of Monetary Economics 66, 124–136. https://doi.org/10.1016/j.jmoneco.2014.04.005

Forni, M., Gambetti, L., Sala, L., 2014. No News in Business Cycles. Econ J 124, 1168–1191.
https://doi.org/10.1111/ecoj.12111

Forni, M., Giannone, D., Lippi, M., Reichlin, L., 2009. Opening the Black Box: Structural Factor Models with Large Cross Sections. Econom. Theory 25, 1319–1347.
https://doi.org/10.1017/S026646660809052X

Fry, R., Pagan, A., 2011. Sign Restrictions in Structural Vector Autoregressions: A Critical Review. Journal of Economic Literature 49, 938–960. https://doi.org/10.1257/jel.49.4.938

Gali, J., 1999. Technology, Employment, and the Business Cycle: DO Technology Shocks Explain Aggregate Fluctuations? The American Economic Review 89, 23.

Galí, J., Rabanal, P., 2004. Technology Shocks and Aggregate Fluctuations: How Well Does the Real Business Cycle Model Fit Postwar U.S. Data? NBER Macroeconomics Annual 19, 225–288.
https://doi.org/10.1086/ma.19.3585339

Galı, J., Gertler, M., Lopez-Salido, J.D., 2017. Markups, Gaps, and the Wefare Costs of Business Fluctuations. THE REVIEW OF ECONOMICS AND STATISTICS 16.

Giannone, D., Reichlin, L., Sala, L., 2005. Monetary Policy in Real Time 40.

Gutiérrez, G., Jones, C., Philippon, T., 2019. Entry Costs and the Macroeconomy (No. w25609). National Bureau of Economic Research, Cambridge, MA. https://doi.org/10.3386/w25609

Gutiérrez, G., Philippon, T., 2017. Declining Competition and Investment in the U.S. (No. w23583). National Bureau of Economic Research, Cambridge, MA. https://doi.org/10.3386/w2358

Guvenen, F., Ozkan, S., Song, J., 2014. The Nature of Countercyclical Income Risk.

journal of political economy 40.

Hallin, M., Liška, R., 2007. Determining the Number of Factors in the General Dynamic Factor Model. Journal of the American Statistical Association 102, 603–617. https://doi.org/10.1198/016214506000001275

Heathcote, J., Perri, F., Violante, G.L., 2010. Unequal we stand: An empirical analysis of economic inequality in the United States, 1967–2006. Review of Economic Dynamics 13, 15–51.
https://doi.org/10.1016/j.red.2009.10.010

Iyigun, M.F., Owen, A.L., 2004. Income Inequality, Financial Development, and Macroeconomic Fluctuations. The Economic Journal 114, 352–376. https://doi.org/10.1111/j.1468-0297.2004.00212.x

Jäntti, M., Jenkins, S.P., 2010. The impact of macroeconomic conditions on income inequality. J Econ Inequal 8, 221–240. https://doi.org/10.1007/s10888-009-9113-8

Jappelli, T., Pistaferri, L., 2010. Does consumption inequality track income inequality in Italy? Review of Economic Dynamics 13, 133–153. https://doi.org/10.1016/j.red.2009.11.001

Jordà, Ò., 2005. Estimation and Inference of Impulse Responses by Local Projections. American Economic Review 95, 161–182. https://doi.org/10.1257/0002828053828518

Kaplan, G., Moll, B., Violante, G.L., 2018. Monetary Policy According to HANK. American Economic Review 108, 697–743. https://doi.org/10.1257/aer.20160042

Khan, L., Vaheesan, S., 2017. Market Power and Inequality: The Antitrust Counter-revolution and Its Discontents. Policy Review 11, 61.

Krueger, D., Perri, F., Pistaferri, L., Violante, G.L., 2010. Cross-sectional facts for macroeconomists. Review of Economic Dynamics 13, 1–14. https://doi.org/10.1016/j.red.2009.12.001

Lütkepohl, H., 2006. Structural vector autoregressive analysis for cointegrated variables. Allgemeines Statistisches Arch 90, 75–88. https://doi.org/10.1007/s10182-006-0222-4

Manning, A., 2011. Imperfect Competition in the Labor Market, in: Handbook of Labor Economics. Elsevier, pp. 973–1041. https://doi.org/10.1016/S0169-7218(11)02409-9

Mumtaz, H., Theophilopoulou, A., 2017. The impact of monetary policy on inequality in the UK. An empirical analysis. European Economic Review 98, 410–423. https://doi.org/10.1016/j.euroecorev.2017.07.008

Nekarda, C., Ramey, V., 2013. The Cyclical Behavior of the Price-Cost Markup (No. w19099). National Bureau of Economic Research, Cambridge, MA. https://doi.org/10.3386/w19099

Olivier Coibion, Yuriy Gorodnichenko, Lorenz Kueng, John Silvia, 2017. Innocent Bystanders? Monetary policy and inequality. Journal of Monetary Economics 88, 70–89. https://doi.org/10.1016/j.jmoneco.2017.05.005

Parker, S.C., 2000. Opening a can of worms: the pitfalls of time-series regression analyses of income inequality. Applied Economics 32, 221–230. https://doi.org/10.1080/000368400322903

Phillips, P.C.B., 1998. Impulse response and forecast error variance asymptotics in nonstationary VARs. Journal of Econometrics 83, 21–56. https://doi.org/10.1016/S0304-4076(97)00064-X

Piketty, T., 2014. Capital in the twenty-first century. Belknap Press of Harvard Univ. Press, Cambridge, Mass.

Romer, C.D., Romer, D.H., 2010. The Macroeconomic Effects of Tax Changes: Estimates Based on a New Measure of Fiscal Shocks. American Economic Review 100, 763–801. https://doi.org/10.1257/aer.100.3.763

Sims, C.A., Stock, J.H., Watson, M.W., 1990. Inference in Linear Time Series Models with Some Unit Roots. https://doi.org/10.2307/2938337

Song, J., Price, D.J., Guvenen, F., Bloom, N., von Wachter, T., 2019. Firming Up Inequality*. The Quarterly Journal of Economics 134, 1–50. https://doi.org/10.1093/qje/qjy025

Stock, J., Watson, M., 2005. Implications of Dynamic Factor Models for VAR Analysis (No. w11467). National Bureau of Economic Research, Cambridge, MA. https://doi.org/10.3386/w1146

Uhlig, H., 2005. What are the effects of monetary policy on output? Results from an agnostic identification procedure. Journal of Monetary Economics 52, 381–419. https://doi.org/10.1016/

Wollmann, T.G., 2019. Stealth Consolidation: Evidence from an Amendment to the Hart-Scott-Rodino Act. American Economic Review: Insights 1, 77–94. https://doi.org/10.1257/aeri.20

A: Firm Level Markup Computation

In the framework proposed by De Loecker and Warzynski (2012) markups are estimated at the firm level using the financial data and the cost minimization problem of the firm, without imposing any assumption on the demand and type of competition. In particular, the researcher has to madel the production function of the firm:

$$Q_{it} = Q_{it}(\Omega_{it}, V_{it}, F_{it}, K_{it})$$

Where Q are sales (SALE in Compustat), V is a vector of variable inputs (COGS in Compustat), F represents fixed costs (SG&A in Compustat) and K stands for capital (PPEGT in Compustat). All variables are deflated using appropriate deflators. The index i represents firms and t stands for time. Then, given the minimization problem faced by the firm:

$$\mathcal{L}(V_{it}, F_{it}, K_{it}, \lambda_{it}) = P^V_{it} V_{it} + r_{it} K_{it} + F_{it} - \lambda_{it}(Q(.) - \bar{Q}_{it})$$

One can note that the lagrangian multiplier λ_{it} actually represents the marginal cost faced by the firm, and thus it is possible to derive an expression for the markup:

$$\mu_{it} = \frac{P_{it}}{\lambda_{it}} = \theta^v_{it} \frac{P_{it} Q_{it}}{P^V_{it} V_{it}}$$

Where θ^v_{it} is the elasticity between output Q and variabe input V. This elasticity can be computed at the sector level (in this work 2-digit NAICS) by running sector specific panel regressions with variables in logs:

$$q_{it} = \theta^v_s v_{it} + \theta^k_s k_{it} + \theta^f_s f_{it} + \omega_{it} + \epsilon_{it}$$

Where ω_{it} represents an unobserved productivity shock, and it can be estimated by running a non-parametric regression

$$q_{it} = \phi(v_{it}, k_{it}, f_{it}) + \epsilon_{it}$$

And then just defining $\omega_{it} = \phi_{it} - (\theta_s^v v_{it} + \theta_s^k k_{it} + \theta_s^f f_{it})$. Then one can model the productivity process as an AR(1):

$$\omega_{it}(\theta_s^v) = \alpha\omega_{it-1}(\theta_s^v) + \xi_{it}$$

Lastly one can impose that variable input responds to current productivity shocks, but lagged variable input does not. Together with the condition that capital and fixed costs do not respond to current shocks, this gives moment conditions to identify the desired elasticity:

$$\xi_{it}(\theta_s^v)\begin{bmatrix} v_{it-1} \\ k_{it} \\ f_{it} \end{bmatrix} = 0$$

Once the sector level elasticity θ_s^v is computed, one can obtain firm specific markups for every period of time.

B: Test for the number of factors and shocks

Three tests are run to determine the number of factors and shocks in the dataset. In order to determine how many static factor r to use, Table 3.1 shows results of Bai and Ng (2002) test. Information Criteria is known to select less factors than Panel Criteria, and the selected number of factors ranges from 5 to 8. This work uses $r = 8$, which is one factor more than the number found by Barigozzi, Lippi, and Luciani (2016). In order to determine the number of structural shocks q, Hallin and Liška (2007) test the number of diverging eigenvalues of the spectral density $\Sigma(\theta)$ of the dataset in differences Δx. The test relies on an expansing subset of the dataset, starting from a random subset of $\frac{2}{3}N$ variables and adding variables in random order until the number N is reached. This procedure can be repeated several times with different variable draws, so as to have a better understanding of the results. Table 3.2 reports how many times each value for q is selected by the test. From this result it is clear that the test selects a number of structural shocks $q = 3$, in accordance with most of the literature on DFMs. Barigozzi, Lippi, and Luciani (2016) apply Hallin and Liška (2007) test on the long run spectral decomposition $\Sigma(0)$, and prove that the test selects the correct number of unit roots τ as $N, T \to \infty$. Table 3.3 shows this test applied to the dataset used in this work. The answer of the test is not as clear-cut as the one for q. For this work it is chosen $\tau = 1$ as the number of unit roots. It is worth noting that if one uses a number $\tau = 2$, the results of this work do not change significantly.

Criteria: Loss Function	IC	PC
p1	8	9
p2	7	8

Table 3.1: Number of static factors r selected by Bai and Ng (2002) test. Both standard Information Criteria and Panel Creteria are reported.

Loss Function: q	p1	p2	p3	p4
1	1.00	1.00	1.00	1.00
3	98.00	98.00	98.00	99.00
4	1.00	1.00	1.00	0.00

Table 3.2: Number of structural shocks q selected by Hallin and Liška (2007) test on $\Sigma(\theta)$. The test is repeated 100 times, and for each loss function it is reported the number of times a particular q was chosen.

Loss Function:	p1	p2	p3	p4
τ				
1	53.00	56.00	56.00	56.00
2	46.00	43.00	43.00	43.00
4	1.00	1.00	1.00	1.00

Table 3.3: Number of unit roots τ selected by Barigozzi, Lippi, and Luciani (2016) test on $\Sigma(0)$. The test is repeated 100 times, and for each loss function it is reported the number of times a particular τ was chosen.

Defining M as the closest integer to $\frac{1}{2}T^{1/2}$, loss functions are:

$$p1 = ((M/T)^{0.5} + M^{-2} + N^{-1}) * \log(\min([(T/M)^{0.5}; M^2; N]))$$
$$p2 = (\min([(T/M)^{0.5}; M^2; N]))^{-1/2}$$
$$p3 = (\min([(T/M)^{0.5}; M^2; N]))^{-1} * \log(\min([(T/M)^{0.5}; M^2; N]))$$
$$p4 = (\min([(T/M)^{0.25}; M^2; N]))^{-1} * \log(\min([(T/M)^{0.25}; M^2; N]))$$

C: Narrative Sign Restrictions

Given the desire to identify a shock to market power, one can characterize such shock by the effect that it has on certain key variables. Similarly to the main identification strategy, one can impose that the shock raises the number of horizontal M&A Deals. On top of that, one would want that a positive shock to market power decreases output on impact, since firms will find optimal to restrict supply in favor of profits. Lastly, in order to check that firms are gaining while output is waning, the final restriction is a positive response of stock prices. All these restrictions are imposed for the first five periods. Restrictions for identification of the rotation matrix R can be derived by imposing conditions on the IRF of certain variables. Rather than imposing some contemporaneous impulse responses to be 0, a quite strong assumption, one can be more agnostic and impose restrictions on the sign of such IRF. This approach was pioneered by Uhlig (2005) in his seminal work on sign restrictions. Given that the Amendment to the Hart-Scott-Rodino Act had such a strong effect on M&A Activity, one would like to factor this into the identification strategy. This can be done by imposing further restrictions on the historical decomposition of certain key variables. In this case, it is natural to impose that the identified shock is the main driver of horizontal M&A Deals at the date of the Amendment, the first quarter of 2001. This approach of using external information to identify the shock is akeen to narrative identification, and it was descibed and named Narrative Sign Restrictions by Antolín-Díaz and Rubio-Ramírez (2018).

The procedure in practice is quite simple, one draws a random rotation matrix R and checks whether it satisfies the desired restrictions. If this is the case, the rotation is stored, otherwise it is discarded. The extraction is repeated thousands of times, since the process is very easy to automate. Given that sign restrictions do not provide exact identification, but only set identification, an infinite set of R matrices satisfy the desired restrictions. One could look at all successful draws of R, but this work follows the methodology of Fry and Pagan (2011), who identify the rotation R^{FP} whose impulse responses are closer to the median of all successful impulse responses. Given the computationally intensive procedures involved in sign identification, it is convenient to focus on dynamic factors, rather than static factors. In practice this amounts to a double rank reduction through the matrix $\hat{K}$ that brings the rotation matrix dimension down from $r = 7$ to $q = 3$ (See Appendix B for derivation of these numbers). As a consequence the identified IRF are computed as:

$$IRF^{FECM} = \hat{\Lambda} \left[\hat{A}^{VECM}(L) \right]^{-1} \hat{K} R^{FP}$$