Start-up Acquisitions, Venture Capital and Innovation: A Comparative Study of Google, Apple, Facebook, Amazon and Microsoft

Gugler, Klaus; Szücs, Florian; Wohak, Ulrich

DOI:
10.57938/2f8d7b88-c44e-4933-b926-03d1460bf880

Published: 01/07/2023

Document Version:
Publisher's PDF, also known as Version of record

Department of Economics
Working Paper No. 340

Start-up Acquisitions, Venture Capital and Innovation: A Comparative Study of Google, Apple, Facebook, Amazon and Microsoft

Klaus Gugler
Florian Szücs
Ulrich Wohak

July 2023

Start-up Acquisitions, Venture Capital and Innovation: A Comparative Study of Google, Apple, Facebook, Amazon and Microsoft

Klaus Gugler, Florian Szücs and Ulrich Wohak*

February 2023

Abstract We evaluate the impact of big-tech acquisitions on the incentives for investment and innovation. Using data on several hundred acquisitions by Google, Apple, Facebook, Amazon and Microsoft (GAFAM), we study the evolution of venture capital investment and patenting relative to control groups. The results show a clear negative impact on investment, while the effect on innovation depends on the acquirer and period. Both outcomes improve over time, as GAFAM firms become more similar in terms of their product and tech-portfolios, increasing competition. Yet, around 10% of acquisitions impact both metrics negatively.

Keywords: M&A, big-tech, innovation, investment

JEL Codes: D22; G34; K21; L41

*Klaus Gugler, Vienna University of Economics and Business, klaus.gugler@wu.ac.at; Florian Szücs, Vienna University of Economics and Business, florian.szuecs@wu.ac.at; Ulrich Wohak, Vienna University of Economics and Business, ulrich.wohak@wu.ac.at.

1. Introduction

Acquisitions of start-up firms by large incumbents in technology markets have increasingly drawn the attention of antitrust agencies in recent years. Five of the biggest tech companies - Google (now Alphabet), Amazon, Facebook (now Meta), Apple and Microsoft, collectively referred to as GAFAM - have jointly acquired more than 900 companies since their foundation. Very few of these acquisitions have been scrutinized by competition authorities, as most transactions remain below the size thresholds specified in the Hart-Scott-Rodino (HSR) Act.[1] Only a handful of these acquisitions have been remedied or blocked.

While the combined market valuation of these five platforms is over $5 trillion (more than a third of the value of the S&P100), they predominantly acquire small, nascent start-ups. Their acquisition strategy has been interpreted as eliminating potential competitors and, in turn, pre-empting future competition. This development has given rise to severe economic but also political and societal concerns about competition in high-tech markets and the accumulation of assets and power. An important aspect of the vivid M&A activity of GAFAM firms is its effects on the innovative potential of economies. On the one hand, there are concerns that these acquisitions could lead to a decline in innovation, particularly in digital markets. "Killer acquisitions" (Cunningham et al., 2021), the creation of "kill zones" (Kamepalli et al., 2020) or conventional raising rival's cost strategies (Salop and Scheffman, 1983) may follow these acquisitions and leverage market power to new markets, curbing future innovative activity. This is the overwhelming concern in a number of competition reports,[2] the predominant finding in the empirical literature[3] as well as parts of the theoretical literature.[4]

On the other hand, there is a strand of the theoretical literature that also sees positive aspects of big-tech acquisitions.[5] For example, Rasmusen (1988) conjectures that one of the reasons innovative firms are founded and funded in the first place is the prospect of being acquired by a large incumbent firm as the most important exit strategy ("entry for buyout"). Cabral

[1]For this reason, the Federal Trade Commission (FTC) issued orders to these five companies requiring them to provide information about prior acquisitions not reported to the antitrust agencies under the HSR in the course of the large congressional investigation. See Federal Trade Commission (2021).

[2]See "Investigation of Competition in Digital Markets: Majority Staff Report and Recommendations, Subcommittee on Antitrust, Commercial and Administrative Law of the Committee on Judiciary", "Final Report of the Stigler Committee on Digital Platforms (September 2019), "Unlocking Digital Competition: Report of the Digital Competition Expert Panel" (Furman report), "Ex-post Assessment of Merger Control Decisions in Digital Markets" (2019 LEAR report), "Australian Competition & Consumer Commission's Digital Platforms Inquiry" (July, 2019), "The French Competition Authority's Opinion on the Online Advertising Sector" (March, 2018), "A report to the European Commission, Competition Policy for the Digital Era", (Jacques Cremer, Yves-Alexandre de Montjoye and Heike Schweitzer, 2019).

[3]See e.g. Kamepalli et al. (2020) (negative effects on venture capital, creation of "kill zones"), Gautier and Lamesch (2021) (substantial portion of acquired products and services discontinued), Affeldt and Kesler (2021) (half of acquired apps in Google Play Store shut down), and Koski et al. (2020) (reduction of entry rates and venture capital funding in target's product market).

[4]See e.g. Motta and Peitz (2021) providing an overview over the mostly harmful competitive effects of such acquisitions.

[5]See also a recent empirical paper by Doan and Mariuzzo (2022) finding positive effects of leading firm acquisitions on patenting in the cloud computing market.

(2021) stresses incomplete markets for technology transfer, and acquisitions may be a means for incumbents to source complementary assets. Their high willingness to pay for such assets creates innovation incentives for entrants. Letina et al. (2021) find in their model that a complete prohibition of large tech acquisitions would result in lower innovation efforts.

The above literature, however, suffers from a number of drawbacks and cannot draw general conclusions on the effects of GAFAM mergers on innovation incentives. First, there is no comprehensive study of GAFAM acquisitions on innovation outcomes. For example, Kamepalli et al. (2020) analyze nine large software acquisitions by Facebook and Google. Doan and Mariuzzo (2022) focus on the cloud computing market, where only three out of the five GAFAMs are active (Amazon, Microsoft and Google). Conversely, in this paper we analyze the entirety of acquisitions on which data are available, by all five GAFAM firms.

Further, the outcome variables used in the extant empirical literature are not entirely convincing proxies for innovation. For example, venture capital (VC) funding, the number of products or the number of apps - while arguably related to innovation - have their deficiencies in measuring innovation input or output. VC funding may also include other, non-innovation outlays such as general purpose expenditures, while new products and apps are not necessarily innovative. This article analyzes venture capital funding as well, but complements it with an analysis of patenting behavior. While patent measures suffer from their own problems,[6] they are closely related to innovation. We argue that patents are important measures of innovation in digital markets, as evidenced by the sheer size of the patent portfolios of GAFAM firms as well as the prevalence of patenting activity among their targets. Moreover, we stress the combination of our two measures of investment and innovation, VC funding and patenting activity. Acquisitions as well as their effects are a heterogeneous form of investment. One measure of innovation does not appear to fully capture this heterogeneity.

Thus, while the literature finds a preponderance of potentially negative effects of large tech acquisitions on the competitive process and in particular on innovation incentives, strong conclusions are not yet possible. We contribute to this literature by analyzing the effects of all GAFAM acquisitions on i) venture capital funding in the relevant markets and ii) patenting in the relevant technology classes. We interlink three sources of data and assemble a database containing (1) all acquisitions of GAFAM firms (SDC Platinum) from 1990 - 2020, (2) all firms with venture capital funding including GAFAM targets (Crunchbase), and (3) all patents granted to all firms involved (PATSTAT).

As in any study on the effects of ownership changes, it is crucial to determine the proper counterfactual. Which innovation and investment activities would have been observed had the

[6]Not all innovations are patented, and not all patents are (valuable) innovations. Relatedly, patents suffer from skewness in value (some patents are very valuable, most are not). There may be strategic patenting, e.g. for foreclosure reasons (patent thickets), etc. See Lanjouw and Schankerman (2004) for ways to increase the accurateness of patents as measure for innovation.

merger not taken place? This is a particularly difficult problem in the case of tech mergers involving very young, small and nascent firms, where a wide range of potential counterfactuals can be imagined. Fortunately, our research question allows us to circumvent this problem in so far as we are not interested in the effect of tech mergers on the innovation of a particular firm (e.g. the target or the combined firm), but rather in their effect on a group of firms that constitutes the relevant innovation and venture capital markets. Thus, we evaluate the effect of mergers on patenting activity in the affected technology classes (delineated through 8-digit International Patent Classification (IPC) groups) by comparing them to a matched set of technology classes. To identify the relevant VC markets, we construct a semantic caliper around the M&A target, i.e. a group of firms with relatively similar business descriptions (Kamepalli et al., 2020). Thus, the treatment groups are not the merger targets or the combined firms, but rather a group of firms that we expect to be affected by a specific GAFAM acquisition, based on their patenting activity and business descriptions.

Our approach alleviates two problems. First, looking at the counterfactual for a treated group of companies is – arguably – less controversial than determining the counterfactual for a specific start-up firm. The prospective or counterfactual development of these groups of companies is less heterogeneous and more predictable than the counterfactual development of a particular firm. Second, and more importantly, we want to determine whether acquisitions benefited society, e.g. because they increased patenting and/or venture capital funding, or whether acquisitions destroyed value, e.g. because they were "killer acquisitions" or helped establish "kill zones", reducing overall innovation in the economy. To determine the overall impact, we need to estimate the effects of these acquisitions on all affected companies.[7]

Analyzing all acquisitions and two outcome variables related to investment and innovation allows us to achieve several objectives. First, we can draw a more nuanced picture of the innovation effects of GAFAM acquisitions than has so far been possible. Not all acquisitions are killer acquisitions nor do all acquisitions create kill zones. Some, however, do and we try to document patterns among them. If an acquisition reduces both overall VC funding and patents, the negative effects of this acquisition are likely to outweigh any positive effects. Second, there is a lot of heterogeneity across acquisitions. Some acquisitions are unambiguously undesirable, as they decrease both VC and patenting. Some, however, affect the two metrics asymmetrically, while others increase both. We characterize these types of acquisitions. Finally, the size of our sample allows us to estimate heterogeneous effects across GAFAM acquirers.

Our main findings indicate that while big-tech acquisitions have a consistently negative impact on venture capital funding, their impact on innovation has become positive after 2010 (with the exception of Apple acquisitions). Moreover, the effects on VC funding become less negative

[7]For example, an acquisition might increase the number of patents of the combined firm, but lower innovation incentives for rival firms. If the net effect is negative, total welfare would likely decrease.

after 2010. Competition among GAFAM firms, as evidenced by the growing similarities in their patent and product portfolios, has increased over time. We find that this increase in competitive rivalry among GAFAM firms appeared to have led to less negative effects of tech acquisitions on VC investment and even positive effects on patent citations. For around 11% of deals, however, we find negative impacts on both innovation and VC investment. We observe two significant determinants: acquisitions with a publicly available transaction price as well as acquisitions in which both parties to the acquisition are active in the same technological field (and have patented there) are less likely to entail negative effects on both innovation and VC investment.

The rest of this paper is organized as follows. Section 2 discusses theory and potential channels for big-tech acquisition effects. Section 3 describes our dataset, while our empirical strategy is detailed in section 4. Results of the empirical analysis are discussed in section 5 and section 6 concludes.

2. Theoretical predictions

Any theory of the effects of big tech mergers must incorporate the characteristics of the firms and industries involved. The standard setup involves a dominant digital platform (the incumbent) acquiring a potential or nascent competitor (the start-up). The start-up is often small but quickly growing in adjacent markets. Crucially, if the start-up's project is successful it can become a substitute to the incumbent's product or service, i.e. it becomes an actual competitor. Characteristics of digital industries include network effects, multi-sidedness, free provision of services, and the importance of data. Motta and Peitz (2021) provide a good starting point for which effects can be expected in this setup. They stress the importance of the likely counterfactual, and show that whenever the start-up has the ability to pursue its project, the merger will be anti-competitive. Put differently, the acquisition can only be pro-competitive if the potential competitor is unable to pursue the project absent the merger and if the incumbent has an incentive to pursue, rather than shelve, the project.

This simple intuition has several implications on which effects to expect. First, if the replacement effect (Arrow et al., 1962) does not hold (such that duopoly profits are larger than monopoly profits), then projects that an entrant would not carry out, while an incumbent would, may exist. Motta and Peitz (2021) state that the replacement effect holds for several standard oligopoly models, but may not hold for general quality-enhancing innovations. Thus, if potential targets are active in adjacent markets and develop complementary products which may become substitutes in the future, the incumbent has an increased incentive to acquire. This results in the early elimination of potential competitive threats.[8]

[8]Cunningham et al. (2021) find that when incumbents acquire firms which have been developing competing drugs, they are more likely to abandon such projects if they possess more market power. Thus, the more competition there is among incumbents, the less likely killer acquisitions take place.

Under the assumption that monopoly profits are higher than the sum of duopoly profits (the replacement effect holds), the incumbent will always make the highest bid and win the competition. Thus, a high willingness to pay on the side of the incumbent may indicate that the incumbent wants to protect its incumbency rents.[9]

Motta and Peitz (2021) endogenize the resources of the start-up and allow for exclusionary conduct by the incumbent. Exclusionary practices (e.g. refusal to supply, degradation of interoperability, tying/bundling or imitation of the entrant's products) may be costly for the incumbent in the short run but may increase future profits.[10] Exclusionary conduct and acquisition may be complementary, since an exclusionary strategy reduces the acquisition price. Thus, the existence of a "kill zone" (according to which new firms/investors would stay away from the core market of large digital platforms) can be explained by the threat the incumbent may engage in exclusionary practices. Kamepalli et al. (2020) were the first to show that acquisitions by the incumbent may lower payoff prospects of new entrants and thus discourage them from investing. Choi and Jeon (2021) analyze bundling strategies and find that bundling can lead to a leveraging of monopoly power from the incumbent into the target market, creating a "kill zone" as more efficient competitors would decide not to enter the acquisition market.

Another theory of harm arising from big-tech acquisitions relates to the increased capacity for data collection. With a larger pool of consumer data, the matching accuracy between advertisers and consumers is enhanced, leading to an increase in the market share of the dominant player. This can result in competitors providing lower quality, causing a decline in their market share. Thus, while the merger may be consumer welfare increasing in the short run, it may be harmful to consumers (and innovation) in the long term, as firms may be induced to exit.

Authors have not only looked at the quantitative innovation incentive effects of mergers, but also whether the prospect of being acquired changes the type of innovation undertaken. Bryan and Hovenkamp (2020) conjecture that the leader will always acquire the start-up to prevent the laggard from catching up technologically. Start-ups will then bias their R&D investment towards improving the incumbent's technology rather than towards technology helping the laggard to catch up.

Cabral (2021), somewhat in contrast to Motta and Peitz (2021), highlights the importance of technology transfer through acquisitions in the presence of imperfect knowledge markets. Digital industries are characterised by high uncertainty about where the next competitive threat comes from, which lowers the pre-emption motive for acquisitions. Acquisitions allow the transfer of

[9] The acquisitions of Instagram and Whatsapp by Facebook and of Waze by Google to prevent competition in the social network apps market and with Google Maps, respectively, are often cited to follow this pattern. A policy of preventing incumbent firms from bidding must, however, trade off the pro-competitive effects against possible negative innovation effects if expected payoffs of the innovator are reduced, see Fumagalli et al. (2020).

[10] Exclusion may be all the more likely if there is no trade off between short and long run profits as possibly with some imitation strategies, see Shelegia and Motta (2021).

complementary technology to the incumbent. If this technology is worth more to the incumbent, the higher acquisition price generates innovation incentives for entrants in the first place (see also Letina et al. (2021)).

Of course, all theories of harm suggesting decreased innovation incentives must be traded off against any possible efficiency effects arising from the merger. These may include innovation complementarities, internalization of innovation externalities, positive network effects as well as other efficiencies. One general finding of the theoretical literature is that increased competitive pressure decreases the likelihood of anti-competitive effects of mergers. Thus, the likely effects of big tech mergers on innovative activity are ambiguous and a matter of empirical investigation.

3. Data

3.1. Mergers and acquisitions

We use Thomson Reuters SDC Platinum to obtain a full list of all mergers and acquisitions associated with Google, Apple, Facebook, Amazon and Microsoft. We manually clean the results and remove entries that do not correspond to proper M&A (e.g. repurchases) and in addition use financial news websites to corroborate all deals. We discard transactions the status of which is 'pending' or 'unknown' and only retain 'completed' deals. We obtain a list of 912 M&A, 301 by Google, 116 by Apple, 100 by Facebook, 124 by Amazon and 271 by Microsoft. While the acquisitions range from 1987 to 2020, only 7% occur before 1999 and more than half occur in the 2010-2019 period (see figure 1).

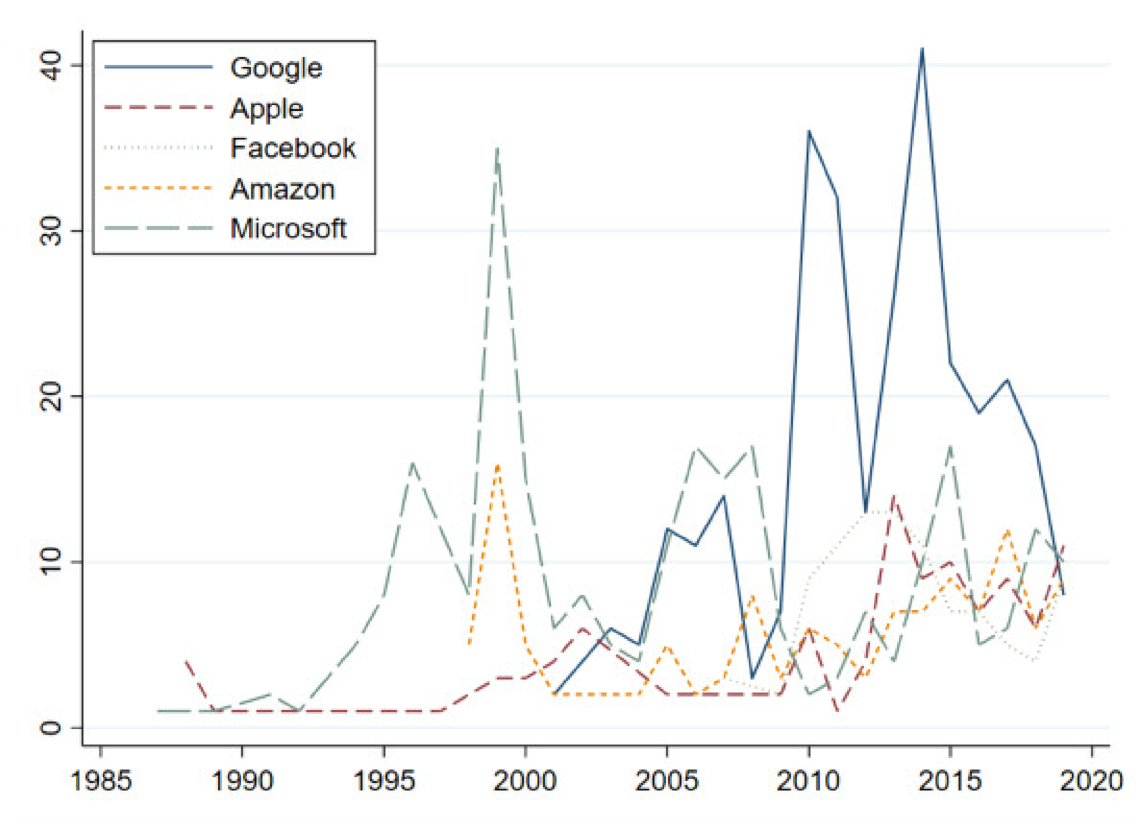
3.2. Venture capital investment

To investigate the effects of GAFAM acquisitions on VC investment, we collect data on business activities of start-ups and innovative firms as well as venture capital financing activity from Crunchbase,[11] a commercial database that collects information on firms with a particular focus on start-ups. We obtain data on the identity, business description, total number of employees, total venture capital investment as well as a list of industry classifications for a total of approximately 1.4 million firms.

Crunchbase also collects information on financing rounds, covering funding stage (Angel, Seed, Private Equity, etc.), identity of both recipient and investor, the amount invested as well as the date of the financing round. Importantly, we also have access to the firms' business descriptions. As described in section 4.1, this allows us to compare targets and the early-stage VC investment in them to a control group of similar firms.

[11] https://www.crunchbase.com, last accessed on Jan 31, 2023

Figure 1: GAFAM M&A



We are able to match 678 of the 912 GAFAM M&As to the Crunchbase data. Among these transactions, only 68 (10%) occured before the year 2000 and the vast majority took place between 2000 and 2020. 151 (22%) occurred between 2000 and 2010, while 459 (68%) occurred after 2010. For each matched transaction, we obtain the year and amount of early-stage venture capital investment for all firms operating in the same area as the target.[12] We then aggregate firm-specific early-stage investment in a 5-year window around each merger to generate yearly sums for each target, allowing us to track VC investment for firms operating in the same market as GAFAM targets over time. We exclude transactions for which relevant pre-treatment data are missing, leaving us with a total of 577 GAFAM mergers. This results in a panel dataset of 10,706 market-year observations of early-stage VC investment.

3.3. Innovation

We use data from PATSTAT to track firms' innovation activities. PATSTAT is maintained by the European Patent Office and contains data on worldwide patent applications and indicators.

[12] We classify venture capital investment as early-stage investment if the funding stage was "Angel", "Pre-seed", "Seed", "Series A" or "Series B". For a full list and description of funding types see: `https://support.crunchbase.com/hc/en-us/articles/115010458467-Glossary-of-Funding-Types`.

We start by generating a balanced panel dataset at the IPC8-level. The IPC-grouping sorts patents into fine-grained technology classes of up to 8-digits. We retain all patents granted in the 1980-2020 (the last year for which patent data are available) period in a total of almost 75,000 IPC classes, yielding more than 3 million observations. Thus, the unit of observation is an 8-digit IPC class in a year and the resulting panel tracks the total amount of (worldwide) patents and the number of citations received in a 5-year window (Hall et al., 2005), as well as some technology-class specific covariates, such as the number of inventors active, their share of total patents and the growth rate of patents, used for matching below. In order to obtain a more meaningful measure of a patents' importance, we remove self-citations (i.e. citations by the patent-owning entity) when counting the number of times a patent has been cited.

Next, we match acquirers and targets to their respective patents. The large technology companies in the data account for more than 400,000 patents in the sample period (Google 86k, Apple 112k, Facebook 22k, Amazon 40k and Microsoft 142k) and have patents in a total of 10,139 IPC8 technology classes. Out of the 912 GAFAM M&A in the data, we are able to link 355 target firms to PATSTAT, i.e. 39% of targets have at least one, granted patent. The fact that this share is much higher than the population average indicates, that these are technology-focused acquisitions.[13] Most acquisitions linked to patents are undertaken by Google and Microsoft (118 and 120 respectively), followed by Amazon and Apple (45 and 46 respectively), while there are only 26 deals by Facebook. In sum, the target firms have been granted a total of almost 90,000 patents in 5,665 different IPC8 technology classes.

We combine the IPC8-panel with the information on which big-tech company acquired targets. Thus the resulting dataset contains almost 75,000 IPC classes in the 1980 - 2020 period and indicates in which IPC classes and years big-tech acquisitions were made. As the majority of IPC classes were not affected by GAFAM deals, they serve as a donor pool for matching a control group in section 4.2.

4. Empirical strategy

4.1. Constructing VC investment markets

We characterize the treatment group as firms that are sufficiently similar to actual GAFAM targets and hence are potential recipients of VC investment. Conventional economic market delineation is based on demand side (e.g. SSNIP) and/or supply-side substitution approaches using data on prices, quantities and other observable market outcomes. The challenge in applying this approach to nascent technology markets is the fact that such data are not available, as often the start-ups are not selling products yet. In their absence, we leverage text data on firms' business descriptions. Crunchbase has data on approximately 1.4 million firms of which

[13]For example, only around 5% of firms listed in Bureau van Dijk's ORBIS database have any patents.

1,396,188 have a business description. The median business description is 21 words long. In the subsequent analysis we only consider firms with a business description of at least 10 words to ensure that descriptions are meaningful. This leaves us with 1,197,670 business descriptions for further analysis.

We apply natural language processing (NLP) techniques to assess the similarity of firms' business activities to identify firms that are likely to operate in the same VC market. The main challenge in the analysis of text data is the removal of noise associated with polysemy (multiple meanings for one word, e.g. the word "sound") and synonymy (multiple words for one meaning, e.g. "work", "occupation", and "profession"). To capture this underlying semantic structure we employ Latent Semantic Analysis (LSA) (Deerwester et al., 1990) to the business description of firms listed on Crunchbase. LSA is a natural-language processing method that retrieves semantic similarities between words if they are used in similar contexts and hence allows us to assess the similarity of business description on the basis of their latent semantic content.

Before applying LSA, we exclude stopwords[14] and apply a lemmatizer, an NLP tool that performs a morphological analysis on each word and reduces it to its constituents.[15] Next, we construct a $D \times W$ document-term matrix A, where D is the set of documents and W is a set of distinct words. Each element $c_{i,j} \in A$ corresponds to the number of times word $w \in W$ occurs in document $d \in D$. We re-weight the entries in A by their term-frequency inverse-document frequency, such that words that appear frequently across documents are assigned a low weight, whereas words that appear frequently in a single document are assigned a high weight. This re-weighted matrix A is then decomposed using singular-value decomposition resulting in the best rank-C approximation $\boldsymbol{B_C} = \boldsymbol{U_C}\boldsymbol{\Sigma}_C\boldsymbol{M}_C^T$. Each document is represented by a $1 \times C$ vector v_i in the document-component matrix $\boldsymbol{U_C}\boldsymbol{\Sigma}_C$. Finally, we assess the similarity between documents using cosine-similarity between document vectors.

We define a caliper around each target based on the cosine similarity of business descriptions between a GAFAM-target and other firms (see appendix A for more details). For each target, firms with a cosine similarity greater than 0.3 are included in the caliper and constitute the set of firms affected by the merger. There is no clear rule for choosing this particular cutoff, but visual inspection of the similarity in business description suggests that 0.3 is a sensible choice. We provide robustness results with respect to the cutoff in appendix B. The median target-specific caliper is composed of 645 firms, for which 388 early-stage investments are recorded on average.

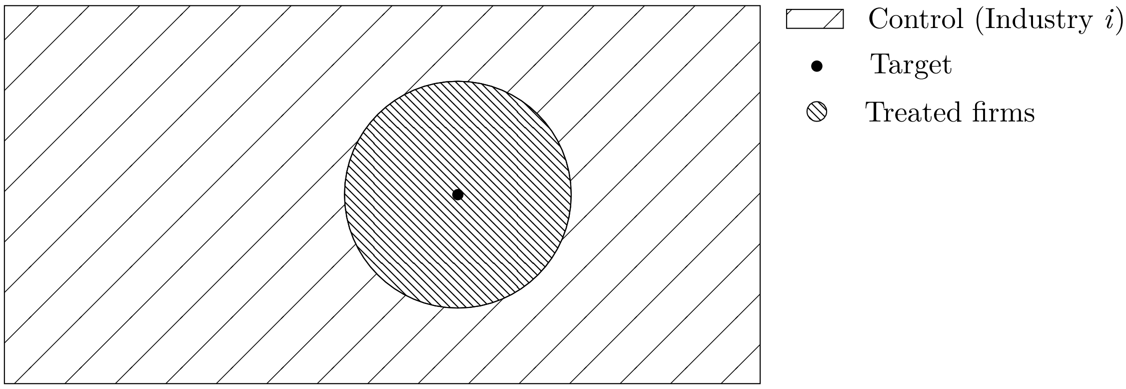
We construct the counterfactual investment trajectory for each caliper using industry classifications from Crunchbase. Firms are categorized in industries such as "Augmented Reality",

[14] Stopwords are words that are considered to carry little semantic meaning, such as "and", "or", "this", "that", etc. In the context of our analysis, we extend this list by additional words that are likely to add little value, such as "LLC", "www", "GmbH", etc.

[15] For example, the words "buy", "buys", "bought", "buying" are reduced by the lemmatizer to its lemma "buy".

"Cloud Computing" and "Machine Learning".[16] The average firm is assigned to more than one industry group. Therefore, we include in the counterfactual all firms that are active in at least one industry of the target. From this set of firms, we exclude all firms that belong to the target caliper. The median target-specific control group is composed of approximately 7,760 firms for which we, on average, have records on 13,157 early-stage investment rounds.

Figure 2: Illustration of treatment caliper and control groups



We illustrate this approach in figure 2. For each target (depicted as a black dot) we consider all firms that belong to at least one of the target's broader industry (e.g. augmented reality). From this set of firms we construct two disjoint groups. The treatment group is defined as all firms with a cosine similarity of 0.3 or more with regard to the target's business description (the hatched circle). All other selected firms constitute the control group, depicted as the hatched rectangle in figure 2. We next describe the matching process for innovation outcomes.

4.2. Matching technology classes

When looking at the different IPC classes in the dataset it becomes apparent, that (just like the underlying individual patents) patenting is very heterogeneous across classes. For example, while the median amount of patents (citations) in an IPC8 class and a year is 14 (49) on average, these numbers rise to 104 (1153) in IPC classes with GAFAM acquisitions. It seems plausible that tech-giants strategically select which targets and technologies they acquire (rather than random selection into treatment). We therefore need to account for this selection process in order to avoid bias.

We do so by employing a propensity-score matching procedure, in which a binary indicator for whether an IPC8 class is affected by GAFAM deals or not is explained through a set of innovation-related indicators and fixed-effects. The predicted values of this model, the propensity scores, then represent the ex-ante probability that a technology class is affected by GAFAM

[16] A full list of industry classifications can be found here: `https://support.crunchbase.com/hc/en-us/articles/360043146954-What-Industries-are-included-in-Crunchbase`, last accessed on Jan 31, 2023.

acquisitions. We then pair treated and non-treated IPCs with similar propensity scores in order to obtain a balanced sample.

We use a logit model to regress an indicator for GAFAM acquisitions in an IPC8 class on potential determinants. An IPC8 class is treated if a GAFAM firm acquired a target with patents in that class. Thus, a single acquisition can affect multiple technology classes. The estimation results are reported in table 1. While patent count has a negative coefficient (likely due to its high correlation with citations, $\rho = 0.91$), the total number of received citations in an IPC8 class and a year increases the likelihood of a GAFAM acquisition (both patents and citations were logged for this estimation to avoid tiny coefficients). The log number of inventors has a positive impact on M&A, the log C1 (defined as the share of patents belonging to the largest patentor in a specific class), has a negative impact. High citation IPCs (defined as the 90th percentile of total citations) are more likely to experience M&A. The current and average patent growth rate (patent growth and trend) have a negative and a positive sign, respectively.

Table 1: Selection model for technology classes affected by GAFAM acquisitions

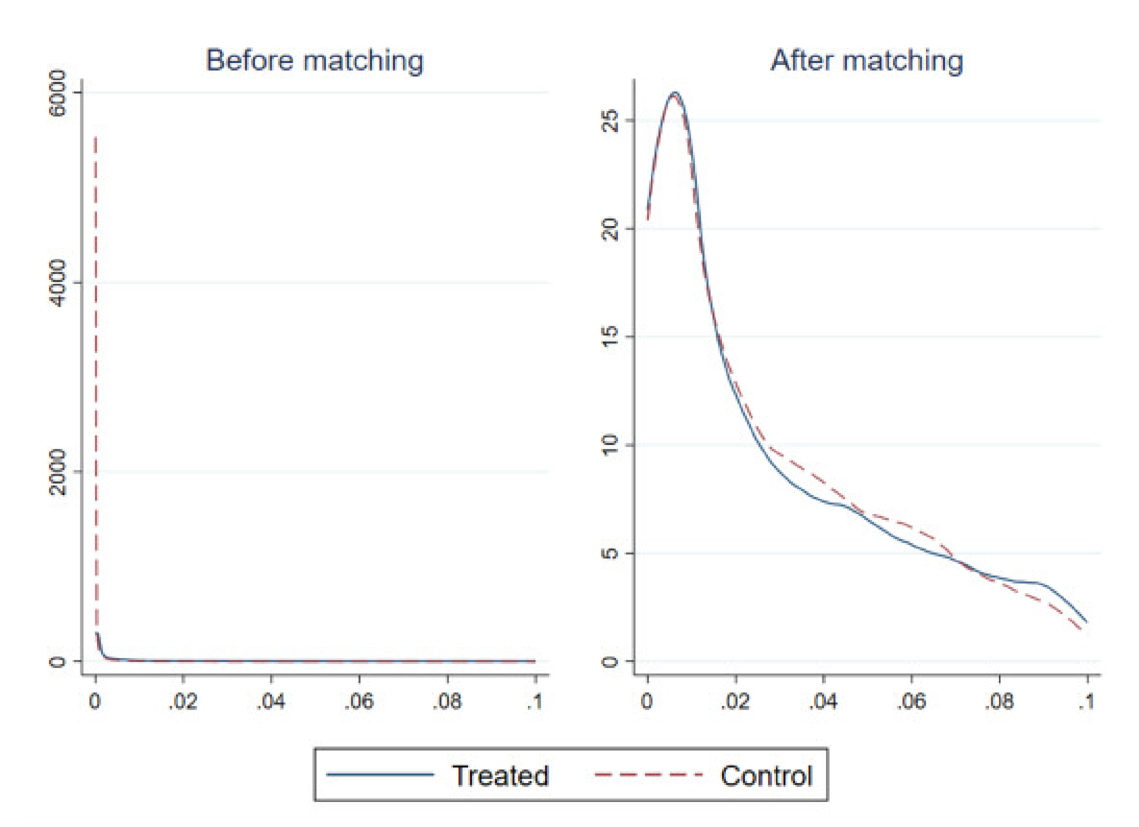
Patents	-0.101***	(0.034)
5-year citations	0.294***	(0.025)
Inventors	1.272***	(0.030)
C1	-0.371***	(0.031)
High citation IPC	0.215***	(0.053)
Trend	2.432***	(0.056)
Patent growth	-0.215***	(0.022)
Observations	1269189	
Pseudo R^2	0.476	

Notes: Standard errors in parentheses, * p<0.1, ** p<0.05, *** p<0.01. Fixed-effects for years and IPC4 groups are included.

From this model, we obtain the propensity scores and proceed to match a control group. For each IPC8 class with GAFAM M&A, we indicate all other IPC8 classes at the same point in time, that have not been affected by M&A and have not already been assigned as a control unit (matching without replacement). We further require the potential controls to be in the same quartile of patent count and growth across IPC classes, to help with balancing. Among those candidate classes, we choose the one with the lowest (absolute) difference in propensity score as a control observation. Iterating this procedure across all IPC classes with GAFAM M&A yields 3,829 treated/control IPC8 pairs and a total of approximately a quarter million observations.

Figure 3 illustrates the distribution of propensity scores before and after matching. While the distribution of propensity scores in the donor pool consists essentially of a mass point at (almost) zero (left side of figure 3), the two distributions look very similar after matching (right side of figure 3). Additional details on the matching procedure are collected in appendix C.

Figure 3: Distribution of the propensity score before and after matching



We have thus created control groups for the affected VC investment markets and for the affected patent technology classes. We will use these counterfactuals to estimate, how these markets would have evolved in absence of GAFAM M&A.

4.3. Estimation

The average treatment effect of big-tech acquisitions on VC investment and patenting behaviour is estimated in a DiD framework. The estimation equation is given by

$$\mu_{i,t} = \delta \text{post}_t + \gamma(\text{treated}_i \times \text{post}_t) + \iota_i + \tau_t + \varepsilon_{i,t},$$

where $\mu_{i,t}$ is the outcome (either log 5-year citations in an IPC8 class or log VC investment in a caliper) of unit i in year t, ι_i and τ_t are unit- and time-fixed-effects respectively and γ identifies the ATT. The error term $\varepsilon_{i,t}$ is robust to heteroskedasticity and allowed to cluster at the IPC8-class or VC-market level.

5. Results

5.1. Main results

Figure 4 provide a general overview over patent citations and VC investment for treatment and control groups. In the case of patents, the treatment group comprises those 8-digit IPC classes where at least one GAFAM firm has made an acquisition, while the "No GAFAM acquisition" group consists of matched control IPC classes without GAFAM acquisitions (see section 4.2). In the case of VC investment, the treatment group consists of semantic calipers based on business descriptions around the target of a GAFAM acquisition, while the control group is the remainder of the broader industry, in which the acquisition took place (see section 4.1).

The respectively left panels of figure 4 show that VC investment increases strongly, particularly between 2010 and 2020, while the evolution of patent citations is hump-shaped, with a peak around 2000. The time dynamics are similar for treated and untreated units in both cases, with treated technology classes having slightly more citations on average.

The right hand side panels of figure 4 plot the outcomes on a common timeline, relative to the event of a GAFAM M&A in the IPC class or VC market. In both cases, the pre-treatment outcomes for the control groups are slightly lower than those of the treatment group, but the dynamics are very similar.[17] After the treatment (in the $t > 0$ periods), we observe differential developments: while patent citations appear to increase, investment drops sharply in VC markets with GAFAM acquisitions.

Table 2 sheds light on the impact of GAFAM M&A on the outcomes. All regressions control for a full set of fixed effects (VC calipers/IPC classes as well as year fixed effects). The effect of GAFAM acquisitions on VC investment is negative and economically large: investment in treated markets decreases by around 30% after a GAFAM acquisition.[18] Conversely, the average impact on patent citations is insignificant.

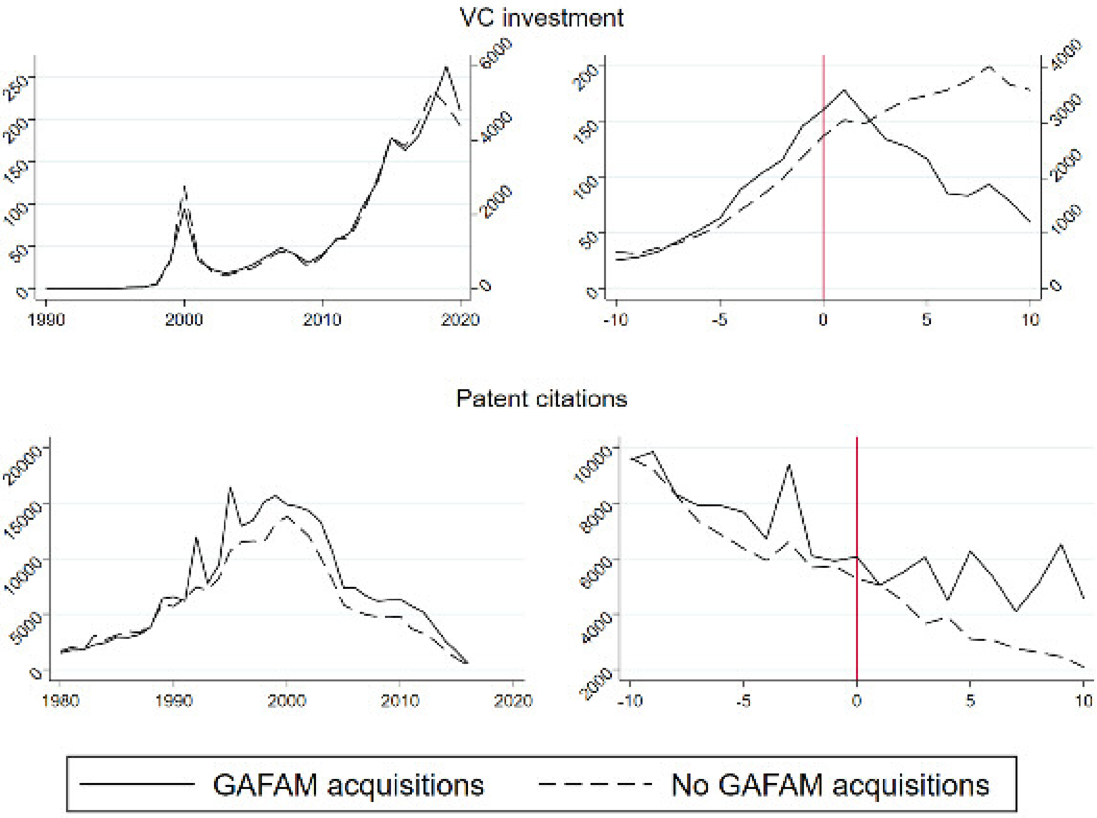
These results, however, mask important heterogeneities across time and firms. Columns (2) and (5) estimate ATTs separately for M&A before 2011 and those from 2011-2020.[19] GAFAM acquisitions before 2011 significantly reduce both patent citations in affected 8-digit IPC classes (by $exp^{(-0.151)} - 1 \approx -14\%$) as well as VC investment in treated calipers (by 37%). This contrasts with acquisitions after 2010. Acquisitions by GAFAM firms undertaken between 2011

[17]We test the plausibility of the common trends assumption by estimating period-specific ATTs for all periods except $t = -1$. Thus, we measure changes in outcomes relative to the last pre-merger period. Both regressions (patent citations and VC money raised) show no significant effects in the five years prior to GAFAM acquisitions. Conversely, ATTs turn significant in the first year after an M&A. This can be interpreted as evidence that i) treated and control IPCs/VC markets did not differ substantially before GAFAM M&A and that ii) GAFAM M&A were causal for the observed changes, as they occurred right after the deals.

[18]In unreported regressions we find that the reduction of VC investment occurs at the extensive margin. Thus, the average VC funding for start-ups in GAFAM acquisitions calipers given funding is not affected, but the likelihood of receiving funding is substantially reduced.

[19]We chose 2010 as the cutoff, as approximately half of the acquisitions in the data occured before 2010. We find similar results when subdividing into 5-year periods.

Figure 4: Impact of GAFAM acquisitions on patents and citations



Notes: The two upper panels show early-stage venture capital investment in absolute (left) and relative (right) time. The two lower panels report the same for total patent citations.

and 2020 increase patent citations in affected IPC classes by 36%. Conversely, these acquisitions have reduced VC investment in affected markets by 21%. Thus, while VC investment still responds negatively to GAFAM acquisitions, patent citations seem to have been spurred by them in the last decade.

In the third and sixth column of Table 2, we allow ATTs to vary with the identity of the acquirer. On the VC side, we obtain fairly uniform results: an acquisition by a GAFAM firm lowers follow-up VC investment in that area, irrespective of the identity of the firm (although the impact associated with Amazon M&A is insignificant). Conversely, the impact on patent citations is positive for some acquirers (Google, Facebook), but insignificant for Apple, Microsoft, and Amazon M&A.

In table 3 we investigate, if the observed changes across time can be attributed to specific firms. We therefore estimate treatment effects for individual GAFAM firms by time period.

Table 2: Venture capital investment, Patent citations and GAFAM M&A

	VC investment			Patent Citations		
	(1)	(2)	(3)	(4)	(5)	(6)
M&A	-0.307***			0.064		
	(0.052)			(0.047)		
M&A_2010		-0.463***			-0.151*	
		(0.096)			(0.084)	
M&A_2020		-0.236***			0.305***	
		(0.061)			(0.060)	
Google			-0.285***			0.135**
			(0.090)			(0.060)
Apple			-0.298**			-0.023
			(0.141)			(0.085)
Facebook			-0.330**			0.581***
			(0.133)			(0.133)
Amazon			-0.186			0.146
			(0.121)			(0.099)
Microsoft			-0.426***			-0.055
			(0.098)			(0.093)
Observations	10169	10169	10169	261721	261721	261721
R^2	0.894	0.894	0.894	0.672	0.673	0.672

Notes: Standard errors in parentheses (allowed to cluster at the VC market / IPC class level), * p<0.1, ** p<0.05, *** p<0.01.

Table 3: VC and patent citations by firm and period

		VC investment		Patent citations	
Google	2010	−0.584***	(0.179)	−0.284***	(0.062)
	2020	−0.172*	(0.104)	0.200***	(0.035)
Apple	2010	−0.823***	(0.315)	0.225*	(0.129)
	2020	−0.165	(0.154)	−0.111**	(0.048)
Facebook	2010	−0.181	(0.255)	−0.101	(0.172)
	2020	−0.353**	(0.150)	0.575***	(0.078)
Amazon	2010	0.090	(0.237)	0.115	(0.094)
	2020	−0.320**	(0.129)	0.114*	(0.060)
Microsoft	2010	−0.574***	(0.137)	−0.398***	(0.055)
	2020	−0.282**	(0.140)	0.739***	(0.069)
Observations		10169		261721	
R^2		0.894		0.674	

Notes: Standard errors in parentheses (allowed to cluster at the VC market / IPC class level), * p<0.1, ** p<0.05, *** p<0.01.

In general, the effects of GAFAM M&A on both metrics increase over time with two exceptions. Facebook and Amazon acquisitions worsen in their effects on VC investment in the latter decade, and Apple acquisitions begin to decrease patent citations in their calipers. Moreover, the effects of acquisitions on VC investment remains negative for all firms (except Apple), three out of the five firms displaying significantly negative effects in the latter decade. On the other hand, four out of the five estimated patent citation effects become significantly positive (the exception being Apple).

Summarizing the results across individual GAFAM firms and time, we conclude that innovation effects in treated 8-digit IPC classes have mostly improved over time: the sign of the overall ATT switches from negative to positive in the periods before and after 2010 (table 2, column (5)) and most of the firm-specific effects are significantly positive in the latter part of the sample period (table 3). The crowding-out effects of VC investment after GAFAM M&A remain significantly negative (overall, as well as for most individual firms), but are substantially lower than in the pre-2011 period.

5.2. Policy analysis

We split the policy analysis in two parts. The first part analyzes the question of which economic forces may explain the finding of negative but improving effects on patent citations and VC investment. The second part looks at the determinants of simultaneously negative effects on both metrics, and may shed light on the prevalence of "killer" or "kill zone" acquisitions.

5.2.1. The impact of potential competition

Both theory (Cabral, 2021) and empirical evidence (Cunningham et al., 2021) suggest that increased competition in the market reduces the risk of "killer acquisitions" or the emergence of "kill zones". More competition also decreases the profitability of exclusionary or discriminatory strategies, and reduces the likelihood of negative impacts on innovation and investment activity resulting from mergers. We focus on the competition *among* GAFAM firms and this is deliberate, as it represents a well-defined group of competitors and represents the largest competitive threats to one another.[20]

[20] The recent literature in the press highlights not only the monopolistic and harmful effects of GAFAM firms, but also their increasingly competitive nature, as evidenced by their investments in each other's markets. See e.g. The Economist, `https://www.economist.com/leaders/2021/02/27/the-rules-of-the-tech-game-are-changing`, according to which "the share of the five American giants' revenues that overlaps with the others has risen from 22% to 38% since 2015" or The Economist, `https://www.economist.com/business/2022/12/24/how-techs-defiance-of-economic-gravity-came-to-an-abrupt-end`, according to which "Amazon's cloud-computing arm has seen a sharp slowdown in growth, partly because Google is pouring billions into its own cloud service". Of course, competition also increased from outside the GAFAM group (see e.g. The Economist, `https://www.economist.com/leaders/2022/07/27/the-era-of-big-tech-exceptionalism-may-be-over`). There is also a race to invent the next "super app", where the owner of the "super app" would essentially control a large part the internet. This puts

We measure competition among GAFAM firms in two ways, potential current and potential future competition. The first measure of potential current competition utilizes the Hoberg and Phillips (2016) dataset of firm-by-firm pairwise similarity scores using their 10-K product filings. This dataset provides a time-varying empirical estimate of the matrix of product based cosine similarities of all publicly-listed U.S. firms. Intuitively, the measure is higher when two firms use similar words in their 10-K product descriptions. Importantly, it measures the degree to which specific firms are similar to their competitors and how this changes over time. Hoberg and Phillips (2016) show that their measure is able to capture competition among firms, in particular competition among potential rival firms that offer related products.Thus, for each GAFAM firm, we calculate the maximum product similarity vis-à-vis the respectively other GAFAM firms over time as a measure of potential current competition.[21]

Our second measure of competition calculates patent similarity scores among GAFAM firms. The more similar their patenting activity is across IPC classes, the more GAFAM firms are in competition to each other for future products. To quantify this, we calculate the yearly correlations of the shares of GAFAM patenting across IPC8 classes (Jaffe, 1989). Specifically, we i) calculate the yearly share of patents of each GAFAM firm in each IPC class, ii) correlate these shares and iii) calculate the maximum, yearly correlation of patent portfolios across GAFAM firms. This provides us with a time-varying measure of the intensity of potential future competition across big-tech firms as viewed from the respective firm.

Figure 5 illustrates that, on the one hand, both measures of competition have substantially increased over the sample period and, on the other, that the observed dynamics are rather similar across firms, with the possible exception of Apple.[22] Patent similarity increased from around 0.2 in the mid 90ies to around 0.8 in 2015, product similarity increased from around 0.05 to 0.15 in that time period.[23] In the regressions below, we use the yearly maximum similarity scores as viewed from the respective firm to the respectively other four GAFAM firms.

GAFAM firms in competition to each other (see e.g. The Economist, `https://www.economist.com/briefing/2022/01/22/what-americas-largest-technology-firms-are-investing-in`.) Another example would be the recent launch of the artificial intelligence app ChatGPT by Microsoft, potentially putting Microsoft's Bing search engine in fiercer competition to Google's core business.

[21] We take the maximum on the assumption that one close competitor among GAFAM firms is sufficient to generate competitive pressure. Results are similar when taking average similarity measures among the respectively other GAFAM firms.

[22] This contrasts to the results of Ederer and Pellegrino (2023) finding decreasing product market centrality of GAFAM firms. Ederer and Pellegrino (2023) use the method of Pellegrino (2019) to calculate GAFAM product market centrality measures from the Hoberg and Phillips (2016) cosine similarity data. Two factors may explain the diverging results. We use the raw data from Hoberg and Phillips (2016), while Ederer and Pellegrino (2023) transform the data to arrive at a measure of product market centrality. Product market centrality essentially calculates – under assumptions on the mode of competition – cross-price elasticities utilizing the extent of overlap of common characteristics. More importantly, we calculate maximum similarities *among* GAFAM firms, while Ederer and Pellegrino (2023) calculate product market centrality of GAFAM firms to *all* other (listed U.S.) firms.

[23] While the Jaffe index of patent similarity can be viewed in a cardinal way, the product similarity score cannot be interpreted as such. Therefore, we abstain from judging which index increased by more.

Figure 5: Patent portfolio correlations over time

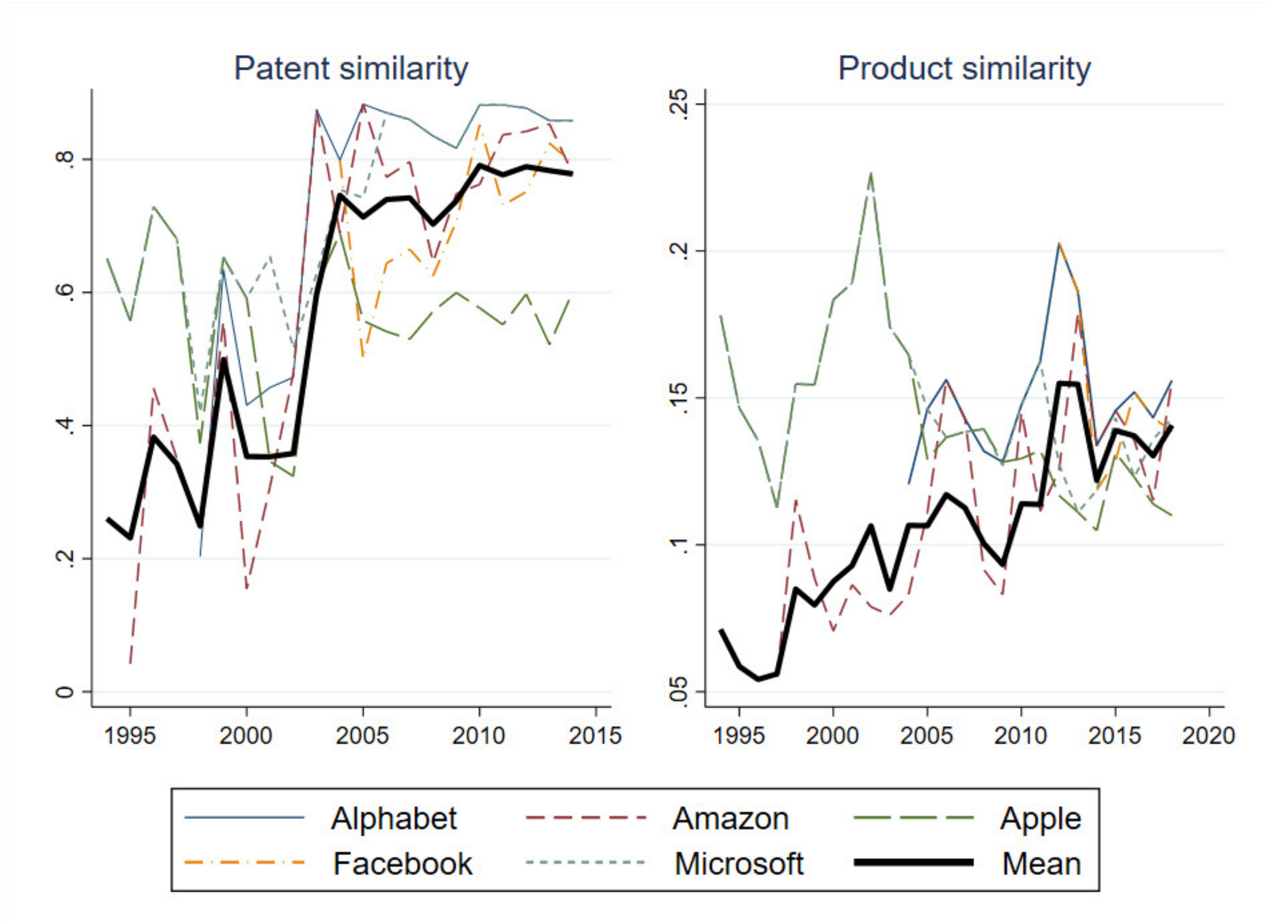


Table 4 presents two sets of regressions for the two outcome variables VC investment and patent citations. In columns (1) and (3) we interact the post M&A dummy with our patent and product similarity measures (patsim and prodsim) for the entire time period, in columns (2) and (4) we split the post M&A dummy into before 2011 and after 2010, and interact with the similarity measures, respectively.

The results in columns (1) and (3) show, that the negative impact on VC investment is mitigated by increasing product market competition among GAFAM firms, while the negative impact on innovation is mitigated by the increasing similarity of GAFAM patent portfolios.

Turning to columns (2) and (4) we note that ATEs do not change strongly over time, once the two dimensions of similarity are controlled for: investment decreases and innovation is insignificantly affected, both before and after 2011. While all similarity-interactions have positive coefficients, they remain insignificant in the period up to 2010. In this period, competition among GAFAM firms was still low due to their relative dissimilarity. After 2010 we see that investment increases with product similarity (potential current competition), while innovation increases with both product and patent similarity (potential current and future competition).

Table 4: VC investment, patent citations and similarity

	VC investment		Patent Citations	
M&A	-0.822***		-0.179***	
	(0.246)		(0.055)	
M&A× patsim	0.088		0.261***	
	(0.166)		(0.045)	
M&A× prodsim	2.883*		0.390	
	(1.540)		(0.253)	
M&A 2010		-1.013**		-0.101
		(0.508)		(0.121)
M&A 2020		-1.063***		-0.037
		(0.325)		(0.087)
M&A 2010× patsim		0.373		0.038
		(0.507)		(0.075)
M&A 2020× patsim		0.247		0.275***
		(0.187)		(0.078)
M&A 2010× prodsim		1.530		-0.769
		(2.108)		(0.544)
M&A 2020× prodsim		4.632**		1.025**
		(2.170)		(0.520)
Observations	8639	8639	261721	261721
R^2	0.894	0.894	0.672	0.673

Notes: Standard errors in parentheses (allowed to cluster at the VC market / IPC class level), * p<0.1, ** p<0.05, *** p<0.01.

Summarizing, the increasing similarity of GAFAM firms' product and patent portfolios has led to stiffer competition and improved the impact of acquisitions on investment and innovation outcomes.

5.2.2. The prevalence of detrimental acquisitions

The policy discussion surrounding big-tech acquisitions frequently focuses on the likely impact on future innovation and the differentiation between desirable and undesirable transactions. Our analysis, which is based on a substantial sample of big-tech mergers and acquisitions, and the corresponding estimates of their impact on patenting and venture capital investment, enables us to examine the factors that contribute to negative impacts on both metrics simultaneously.

To this end, we estimate the impact of every M&A in our data on patents and VC investment individually, i.e. instead of estimating an average ATT, we estimate a separate ATT for each deal by comparing treated units (IPCs or calipers) with their control group. We successfully estimate deal-specific ATTs on both outcomes for a total of 176 M&A. Based on the sign of the estimated effects, we distinguish four clusters which we report in table 5. From the rows of table 5 it can be seen that the majority of deals affect i) patenting in a positive fashion and ii) VC investment in a negative fashion. While only 16 acquisitions affect both metrics negatively, these deals are unequivocally undesirable (from a societal point of view) with regards to the performance measures considered here.[24]

Table 5: Classification of acquisition effects full sample

	Patents +	Patents −
Investment +	58	11
Investment −	91	16

In a next step, we attempt to identify observable determinants for undesirable acquisitions. We thus create an indicator variable for the 16 deals with negative effects on both performance measures and relate them to a set of observable characteristics. Table 6 contains the results.

Table 6: Determinants of negative patent and VC effects

	LPM		Probit		Logit	
Less than 6 years old	-0.008	(0.047)	-0.102	(0.312)	-0.106	(0.602)
Acquisition before 2010	-0.037	(0.057)	-0.462	(0.467)	-0.826	(0.872)
Number of funding rounds	0.019	(0.022)	0.174	(0.144)	0.293	(0.269)
Common IPC target	-0.154**	(0.062)	-0.861**	(0.342)	-1.473**	(0.613)
Transaction price public	-0.072	(0.047)	-0.695*	(0.371)	-1.269*	(0.726)
Patent similarity	-0.016	(0.099)	-0.064	(0.642)	0.056	(1.249)
Product similarity	-0.615	(0.870)	-6.721	(6.921)	-12.702	(13.295)
Observations	163		163		163	
R^2	0.071					

Notes: Standard errors in parentheses, * p<0.1, ** p<0.05, *** p<0.01.

The regression contains variables indicating young targets (5 years old or less; to distinguish start-ups from more established firms), deals before 2010 (to investigate time dynamics), the number of funding rounds a firm was involved in before its acquisition, whether the acquirer and target have patents in a common 8-digit IPC class, whether the acquisition price was made public or not and the measures for patent and product similarity described above.

[24] Interestingly, this corresponds to around 9% of the deals. Cunningham et al. (2021) state that conservative estimates would indicate that 5.3%–7.4% of acquisitions in their pharma sample are killer acquisitions. Thus, if one likens acquisitions that decrease both patent citations and VC investment to "killer" or "kill zone" acquisitions, the percentage of acquisitions that are detrimental to society may be fairly equal across industries.

We observe two significant determinants: acquisitions with a publicly available transaction price as well as acquisitions in which both parties to the deal are active in the same technological field (and have patented there) are less likely to entail negative effects on both innovation and VC investment. The former finding can be viewed as being consistent with Wollmann (2019) and Wollmann (2020) showing that pre-merger notifications are essential to antitrust enforcement. In our sample there is a close relation between deal values being not publicly available and the deal being not notified. As mentioned in the introduction, very few of these acquisitions have been scrutinized by competition authorities, as most transactions remain below the size thresholds specified in the HSR act. Our findings indicate that all GAFAM (or more generally all acquisitions by large platform firms) should be obliged to be notified to the anti-trust authorities.[25]

The latter effect is consistent with the notion that if acquisitions reinforce technological competences because the acquisition takes place in a technologically related field, "killer" motives may not be the driving forces behind the acquisition. While the existing empirical studies are ambiguous about the effects of technological relatedness on innovation of the combined firm,[26] there are no studies looking at the effects of technological relatedness on innovation in the concerned patent class or VC investment market.

6. Conclusion

Big-tech acquisitions of start-up companies are a matter of significant concern for antitrust agencies and society at large, as they are believed to undermine competition and stifle innovation. Our research analyzes all acquisitions made by GAFAM (Google (now Alphabet), Amazon, Facebook (now Meta), Apple, and Microsoft) for which the necessary data could be obtained. Our research approach overcomes the challenge of determining an appropriate counterfactual for a particular acquisition by examining groups of firms in the affected innovation and venture capital markets, rather than the combined firm post-acquisition. We evaluate the impact of these mergers on patent citations in the relevant technology classes (defined through 8-digit IPC groups) by comparing them to a matched set of technology classes. To identify the relevant venture capital markets, we construct a semantic caliper based on the M&A target's business description and compare it to the broader industry.

[25] The finding may also be viewed as being consistent with Fumagalli et al. (2020) if a non-public price might be used to obscure a high-price per employee acquisition. Fumagalli et al. (2020) theoretically show that acquisitions in which a high price (per employee or assets) was paid are more likely to lead to the shelving of products post acquisition, since a high takeover price signals that the acquisition may not be indispensable for the success of the start-up and therefore there is a high chance that it raises anti-competitive concerns.

[26] For example, while Ornaghi (2009) and Haucap et al. (2019) find unambiguously negative effects of technological relatedness between acquirer and target on innovation of the combined firm, Cloodt et al. (2006) finds a curvilinear impact, and Hussinger (2010) finds that technological relatedness is an important determinant of the decision to acquire small or medium sized companies, and that a related technology portfolio conveys an information advantage to acquirers.

Our main findings indicate that while big-tech acquisitions have a consistently negative impact on venture capital funding, their impact on innovation has become positive after 2010. Moreover, the effects on VC funding also become less negative over time. Competition among GAFAM firms, as evidenced by the growing similarities in their patent and product portfolios, has increased. In regressions interacting the post M&A dummy with product and patent similarity measures we find that this increase in competitive rivalry among GAFAM firms appeared to have led to less negative effects of tech acquisitions on VC investment and even positive effects on patent citations. For around 9% of deals, we find negative impacts on both innovation and VC investment. We observe two significant determinants: acquisitions with a publicly available transaction price as well as acquisitions in which both parties to the acquisition are active in the same technological field (and have patented there) are less likely to entail negative effects on both innovation and VC investment.

In light of the significant increase in technological competition among platform companies, it appears that the negative effects of large tech acquisitions have been mitigated. Schumpeterian forces appear to also work in the fields of large platforms with massive economies of scale and network effects. However, this does not mean that anti-trust and merger policies are unnecessary. The continued prevalence of negative effects in venture capital investments suggests that "kill zone" acquisitions may still occur. As such, it is recommended that all acquisitions made by major platform companies be reported to anti-trust authorities and that transaction values be made transparent. Furthermore, when evaluating such acquisitions, anti-trust agencies should consider the potential impact on innovation and (VC) investment markets, as these are often early indicators of future competition.

Our study is the first to analyze a broad sample of big-tech acquisitions and examine the factors that determine their effects. Although our findings contribute to the current understanding of big-tech acquisitions, there are still many unanswered questions that need to be addressed. Further research is needed to develop theories of harm for big-tech acquisitions and provide commensurate empirical evidence, making this a key area for future research.

References

Affeldt, Pauline and Reinhold Kesler, "Big Tech acquisitions—Towards empirical evidence," *Journal of European Competition Law & Practice*, 2021, *12* (6), 471–478.

Arrow, Kenneth et al., "Economic Welfare and the Allocation of Resources for Invention," *NBER Chapters*, 1962, pp. 609–626.

Bryan, Kevin A and Erik Hovenkamp, "Antitrust limits on startup acquisitions," *Review of Industrial Organization*, 2020, *56*, 615–636.

Cabral, Luis, "Merger policy in digital industries," *Information Economics and Policy*, 2021, *54*, 100866.

Choi, Jay Pil and Doh-Shin Jeon, "A leverage theory of tying in two-sided markets with nonnegative price constraints," *American Economic Journal: Microeconomics*, 2021, *13* (1), 283–337.

Cloodt, Myriam, John Hagedoorn, and Hans Van Kranenburg, "Mergers and acquisitions: Their effect on the innovative performance of companies in high-tech industries," *Research policy*, 2006, *35* (5), 642–654.

Cunningham, Colleen, Florian Ederer, and Song Ma, "Killer acquisitions," *Journal of Political Economy*, 2021, *129* (3), 649–702.

Deerwester, Scott, Susan T. Dumais, George W. Furnas, Thomas K. Landauer, and Richard Harshman, "Indexing by latent semantic analysis," *Journal of the American Society for Information Science*, 1990, *41* (6), 391–407.

Doan, Thanh and Franco Mariuzzo, "Digital Platform Mergers and Innovation: Evidence From Cloud Computing Market," 2022.

Ederer, Florian and Bruno Pellegrino, "The Great Startup Sellout and the Rise of Oligopoly," in "AEA Papers & Proceedings," Vol. 113 2023.

Federal Trade Commission, "Non-HSR Reported Acquisitions by Select Technology Platforms, 2010-2019: An FTC Study," 2021.

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

Gautier, Axel and Joe Lamesch, "Mergers in the digital economy," *Information Economics and Policy*, 2021, *54*, 100890.

Hall, Bronwyn H, Adam Jaffe, and Manuel Trajtenberg, "Market value and patent citations," *RAND Journal of economics*, 2005, pp. 16–38.

Haucap, Justus, Alexander Rasch, and Joel Stiebale, "How mergers affect innovation: Theory and evidence," *International Journal of Industrial Organization*, 2019, *63*, 283–325.

Hoberg, Gerard and Gordon Phillips, "Text-based network industries and endogenous product differentiation," *Journal of Political Economy*, 2016, *124* (5), 1423–1465.

Hussinger, Katrin, "On the importance of technological relatedness: SMEs versus large acquisition targets," *Technovation*, 2010, *30* (1), 57–64.

Jaffe, Adam B, "Characterizing the "technological position" of firms, with application to quantifying technological opportunity and research spillovers," *Research Policy*, 1989, *18* (2), 87–97.

Kamepalli, Sai Krishna, Raghuram Rajan, and Luigi Zingales, "Kill zone," Technical Report, National Bureau of Economic Research 2020.

Koski, Heli, Otto Kässi, and Fabian Braesemann, "Killers on the road of emerging start-ups–implications for market entry and venture capital financing," Technical Report, ETLA Working Papers 2020.

Lanjouw, Jean O and Mark Schankerman, "Patent quality and research productivity: Measuring innovation with multiple indicators," *The Economic Journal*, 2004, *114* (495), 441–465.

Letina, Igor, Armin Schmutzler, and Regina Seibel, "Killer acquisitions and beyond: policy effects on innovation strategies," *University of Zurich, Department of Economics, Working Paper*, 2021, (358).

Martin, Dian I and M Berry, "Mathematical foundations behind latent semantic analysis," *Handbook of Latent Semantic Analysis*, 2007, pp. 35–54.

Motta, Massimo and Martin Peitz, "Big tech mergers," *Information Economics and Policy*, 2021, *54*, 100868.

Ornaghi, Carmine, "Mergers and innovation in big pharma," *International journal of industrial organization*, 2009, *27* (1), 70–79.

Pellegrino, Bruno, "Product differentiation and oligopoly: a network approach," *WRDS Research Paper*, 2019.

Rasmusen, Eric, "Entry for Buyout," *The Journal of Industrial Economics*, 1988, *36* (3), 281–299.

Řehůřek, Radim and Petr Sojka, "Software Framework for Topic Modelling with Large Corpora," in "Proceedings of the LREC 2010 Workshop on New Challenges for NLP Frameworks" ELRA Valletta, Malta May 2010, pp. 45–50. `http://is.muni.cz/publication/884893/en`.

Rubin, Donald B, "Using propensity scores to help design observational studies: application to the tobacco litigation," *Health Services and Outcomes Research Methodology*, 2001, *2* (3), 169–188.

Salop, Steven C and David T Scheffman, "Raising rivals' costs," *The American Economic Review*, 1983, *73* (2), 267–271.

Shelegia, Sandro and Massimo Motta, “The" kill zone": Copying, acquisition and start-ups’ direction of innovation,” 2021.

Wollmann, Thomas G, “Stealth consolidation: Evidence from an amendment to the Hart-Scott-Rodino Act,” *American Economic Review: Insights*, 2019, *1* (1), 77–94.

_ , “How to get away with merger: Stealth consolidation and its effects on us healthcare,” Technical Report, National Bureau of Economic Research 2020.

A. Appendix: Cosine similarity via LSA

We use the information contained in companies' business description to identify companies that engage in similar business activities and are thus likely to operate in the same VC market. We apply a method from the natural language processing literature: Latent Semantic Analysis (LSA)[27]. To that end, we represent all business descriptions ("documents") in a document-term matrix $\boldsymbol{A}$ where each row corresponds to a document $d \in D$ and each column corresponds to a word $w \in W$, where D and W correspond to the set of all documents ("corpus") and the complete vocabulary of the corpus, respectively. Each element in $\boldsymbol{A}$ corresponds to the number of times word w occurs in document d:

$$\underset{D\times V}{\boldsymbol{A}} = \begin{pmatrix} c_{1,1} & \cdots & c_{1,W} \\ \vdots & \ddots & \vdots \\ c_{D,1} & \cdots & c_{D,W} \end{pmatrix}$$

We re-weight the individual elements in the document-term matrix $\boldsymbol{A}$ by their term-frequency inverse-document frequency, as is standard in the NLP literature. This weighting assigns a low weight to words that appear frequently across documents and a high weight to words that appear frequently within only a few documents. In particular, each element in $\boldsymbol{A}$ is re-weighted by $tfidf(c_{i,j}) = (1 + log(c_{i,j}) * \left(log(\frac{1+D}{1+d_w}) + 1\right)$ where d_w is the number of documents containing word w.

$$\underset{D\times V}{\boldsymbol{B}} = \begin{pmatrix} tfidf(c_{1,1}) & \dots & tfidf(c_{1,W}) \\ \vdots & \ddots & \vdots \\ tfidf(c_{D,1}) & \dots & tfidf(c_{D,W}) \end{pmatrix}$$

This re-weighted matrix $\boldsymbol{B}$ is then decomposed using singular value decomposition (SVD). SVD transforms B of rank r into three matrices (Martin and Berry, 2007)

$$\underset{D\times W}{B} = \boldsymbol{U\Sigma M^T}$$

where $\boldsymbol{U}$ is a $D \times r$ orthogonal matrix, $\boldsymbol{\Sigma}$ is a $r \times r$ diagonal matrix and $\boldsymbol{M}^T$ is an $r \times W$ orthogonal matrix. The final step in LSA is the elimination of $(r - C)$ rows and columns in $\boldsymbol{\Sigma}$ corresponding to the smallest eigenvalues of $\boldsymbol{\Sigma}$, where C is a scalar chosen by the researcher. There is no optimal scalar C, but it is recommended to choose $100 \leq C \leq 1000$ (Martin and Berry, 2007). In our application, $C = 500$. This truncation process results in the best rank-C approximation $\boldsymbol{B_C} = \boldsymbol{U_C}\boldsymbol{\Sigma}_C\boldsymbol{M}_C^T$ to the input matrix $\boldsymbol{B}$. Each document d_i is represented as a $1 \times C$ vector v_i in the document-component matrix matrix $\boldsymbol{U}_C\boldsymbol{\Sigma}_C$. We construct a measure

[27] We implement LSA in Python using gensim (Řehůřek and Sojka, 2010)

of similarity between documents by calculating the cosine similarity between any two pairs of vectors v_i, v_j:

$$S(v_i, v_j) = \frac{\sum_{c=1}^{C}(v_i \cdot v_j)}{\sqrt{\sum_{c=1}^{C} v_i}\sqrt{\sum_{c=1}^{C} v_j}}$$

where $S(v_i, v_j) \in [-1, 1]$ increases in the similarity of documents.

B. Appendix: Robustness cosine similarity

In this section we conduct a robustness exercise with respect to the cosine similarity cut-offs for semantic calipers around targets (see Section 4.1) for details. Increasing the similarity score can be graphically represented by *decreasing* the radius of the circle in Figure 2.

Table 7: Early-stage VC investment for different cosine similarity cutoffs

	s=0.3	s=0.4	s=0.5
M&A_2010	-0.463***	-0.633***	-0.699***
	(0.096)	(0.136)	(0.220)
M&A_2020	-0.236***	-0.281***	-0.373***
	(0.061)	(0.083)	(0.104)
Observations	10169	7511	3984
R^2	0.894	0.902	0.916

Notes: Standard errors in parentheses, * p<0.1, ** p<0.05, *** p<0.01.

The first column of Table 7 repeats the ATT estimates of the main text in Table 2. The second and third column repeat the exercise for cosine similarity cut-offs at 0.4 and 0.5. The estimated coefficients for acquisitions prior to 2011 and after 2010 suggest that results are robust to the choice of cut-off value.

C. Appendix: Additional matching metrics and figures

Table 8 reports summary statistics. The first two columns contain the means of key metrics in IPC8 classes with GAFAM M&A and in the population, the third column contains the p-value of a t-test for equal means. As can be seen, classes with M&A tend to have more patents, more citations and more inventors active in them. Columns (4) to (6) report summary statistics after the matching procedure and show that all biases were strongly reduced and are mostly insignificant after matching.

Table 8: Biases before and after matching

	Before matching			After matching		
	Treated	Population	p-value	Treated	Control	p-value
Patents	584.42	91.17	0.00	366.42	369.93	0.83
5-year citations	2112.18	185.84	0.00	1337.25	1406.84	0.35
Inventors	8.45	6.72	0.00	8.23	8.01	0.00
C1	-4.34	-3.97	0.00	-4.40	-4.38	0.22
High citation IPC	0.60	0.09	0.00	0.52	0.54	0.10
Trend	0.38	0.43	0.00	0.33	0.34	0.22
Patent growth	0.15	0.37	0.00	0.11	0.11	0.95

Notes: The number of inventors and the share of patents belonging to the largest patentor in a class (C1) are in logs. 'High citation IPC' indicates IPCs in the 90th percentile of the distribution of citations. Patent growth is measured from one period to the next, 'trend' is the class-specific average over time.

Figure 6 illustrates the standardized biases[28] in covariates before and after matching, respectively. While standardized biases are substantial prior to matching, they mostly vanish in the matched sample. An exception to this is the number of inventors, which retains a standardized bias of 0.2 and is also significantly different between groups (see columns 4 - 6 of table 8). However, all other covariates and - most importantly - the dependent variable, 5-year citations, are balanced as can be seen from the t-tests in table 8. The remaining bias in the number of inventors is not too large (Rubin (2001) recommends post-matching biases to be <0.25) and is not likely to cause econometric concern.

[28]Rubin (2001) proposes that the standardized bias (defined as $\frac{\bar{X}^1-\bar{X}^0}{\frac{1}{2}(\sigma^1+\sigma^0)}$, that is the difference in means of treated and control groups divided by the mean of their standard deviations) of covariates should be below 0.25 after matching.

Figure 6: Biases in covariates before and after matching

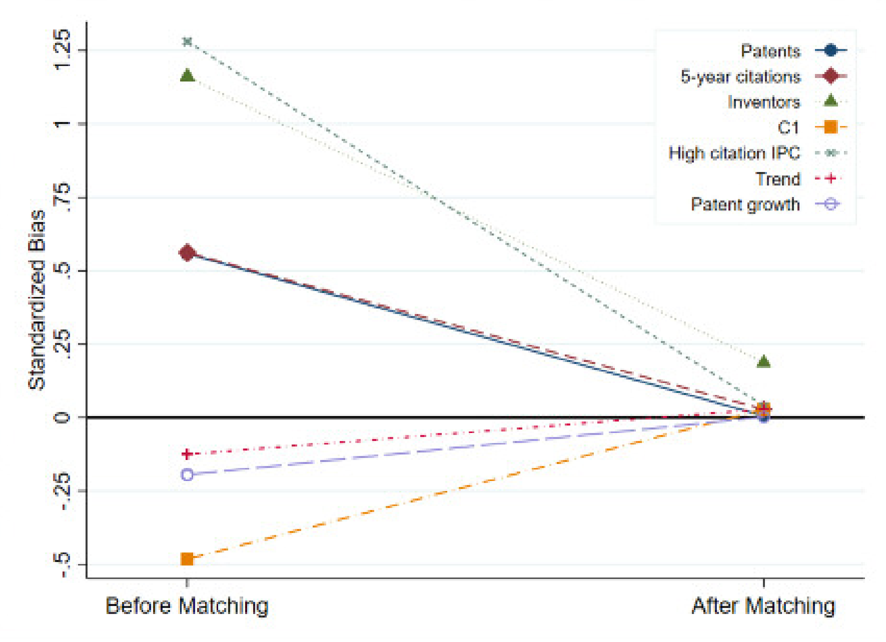


Figure 7: Impact of GAFAM acquisitions on patent citations

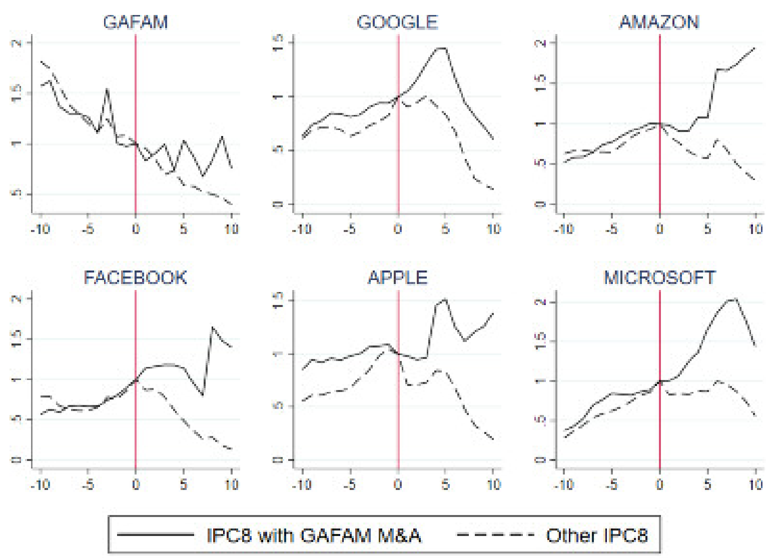


Figure 8: Impact of GAFAM acquisitions on VC investment