2020-01-31

Automation, research technology, and researchers’ trajectories: evidence from computer science and electrical engineering

J.L. Furman, F. Teodoridis. 2020. "Automation, Research Technology, and Researchers’ Trajectories: Evidence from Computer Science and Electrical Engineering." Organization Science, Volume 31, Issue 2, https://doi.org/10.1287/orsc.2019.1308

https://hdl.handle.net/2144/44298

Automation, Research Technology, and Researchers' Trajectories: Evidence from Computer Science and Electrical Engineering[1]

Jeffrey L. Furman
Boston University & NBER
595 Commonwealth Ave., Boston, MA 02215
furman@bu.edu

Florenta Teodoridis
University of Southern California
3670 Trousdale Pkwy, Los Angeles, CA 90089
florenta.teodoridis@marshall.usc.edu

Abstract

We examine how the introduction of a technology that automates research tasks influences the rate and type of researchers' knowledge production. To do this, we leverage the unanticipated arrival of an automating motion-sensing research technology that occurred as the consequence of the introduction and subsequent hacking of the Microsoft Kinect system. To estimate whether this technology induces changes in the type of knowledge produced, we employ novel measures based on machine learning (topic modeling) techniques as well as traditional measures based on bibliometric indicators. Our analysis demonstrates that the shock associated with the introduction of Kinect increased the production of ideas and induced researchers to pursue ideas more diverse than and distant from their original trajectories. We find that this holds for both researchers who had published in motion-sensing research prior to the Kinect shock (within-area researchers) and those who did not (outside-area researchers), with the effects being stronger among outside-area researchers.

Keywords:
Automation, Knowledge Production, Innovation, Research Technology, Rate and Direction of Innovation, Technological Change, Topic Modeling, Machine Learning, Idea Space, Research Trajectories, Knowledge Trajectories, Diversification, Breadth and Depth of Knowledge

[1] Acknowledgements: Authors listed in alphabetical order; both authors contributed equally. We thank Ajay Agrawal, Kevin Bryan, Avi Goldfarb, Danielle Li, Olav Sorenson, Scott Stern, and Martin Watzinger for substantive discussions of the analysis, and three anonymous referees and seminar participants at numerous conferences and seminars for constructive feedback. We are grateful for funding from the National Science Foundation's Program on the Science of Science & Innovation Policy (SciSIP), grants SES-1564368 & SES-1564381. We thank Aparna Anand, Avik Basu, Robyn Johnson, and Lindsay Raymond for excellent research assistance. All errors are our own.

I. Introduction

The automation of physical and mental tasks in the production of goods and services is among the most profound factors affecting the modern economy and the role of human capital in economic output (Mokyr, 2002). Recent work on this topic documents that the extent to which automating technologies substitute for or complement human labor depends upon the extent to which such technologies substitute for the specific tasks associated with production (Deming, 2017; Brynjolfsson et al., 2018; Acemoglu and Restropo, 2018; Felten et al., 2018). However, to date, most of the work on this topic has focused on manufacturing or service industries, while relatively less research examines the impact of such automating technologies on the production of new-to-the-world knowledge (Ding et al., 2010; Cockburn et al., 2017).

In this project, we contribute to investigating the impact of IT-based automating technology[2] on the production of knowledge. This is important because of the central role knowledge production plays in economic growth and the unique features of the knowledge production environment that may provide insights into the role of automation in other sectors. The impact of automating technologies on knowledge production may differ in important ways from the impact of such technologies on goods and services production because of the variety of tasks associated with knowledge work and the extent of autonomy that knowledge workers possess in allocating time across such tasks. These features may enable knowledge workers to respond to automating technologies in ways that differ from those of sectors that face greater rigidity in task re-allocation and, potentially, a greater chance for displacement of tasks previously performed by humans (e.g., Acemoglu and Autor, 2011). Furthermore, in knowledge production, tasks are knowledge-based in the sense that the completion of a research task requires a certain type of research expertise. Hence, we anticipate that researchers' response to automating technologies will be influenced by the overlap in knowledge that is being automated through the research technology (Murray et al., 2016; Zyontz and Thompson, 2017; Teodoridis, 2018). As a result, we distinguish between two types of researchers, those who had previously worked in the research area where the tasks affected by the

[2] Throughout the paper we use the terms 'automating technology', 'IT-based technology', 'research technology' and similar variants interchangeably to refer to IT-based devices (including both software and hardware) that are used in the process of research and that automate certain research tasks e.g., data collection tasks, data analysis tasks, etc.

automating research technology reside, whom we call "within-area researchers" and those who had not previously worked in the area, whom we describe as "outside-area researchers."

Estimating a relationship between the availability of automating research technology and the rate and type of knowledge output is difficult because numerous unobservable factors are likely associated with both research technology availability and researchers' project choices. To generate causal insights, we exploit an unexpected arrival of research technology as an instrument that is correlated with research task automation but not with factors affecting researchers' efforts, other than through its effect on automation. Building on Teodoridis (2018), we leverage the introduction of the Microsoft Kinect gaming system as an unexpected shock that suddenly and dramatically made available a technology that automates core motion-sensing research tasks. Kinect was launched in November 2010 as an add-on for Microsoft's Xbox 360 gaming console. Soon after its launch, hackers developed and released through the open-source community a driver that enabled devices other than the Xbox to interact with Kinect, thus making it possible for scholars in electrical engineering, computer science, and electronics to harness Kinect's motion-sensing output for use in research applications.

Kinect automates several key tasks required to track, collect and analyze real-time complex 3D motion data. Prior to Kinect, in order to complete these tasks, researchers had to rely on developing complex custom algorithms for collecting, analyzing and attempting to augment the output of 2D and 3D cameras. It is important to note that full automation of motion-sensing research tasks implies a system that can recognize and understand any 3D motion scene. While such a system is still to be developed, Kinect helped move a large step closer towards it (Han et al., 2013). In our email exchanges with Kinect researchers, one individual noted that, "*[With Kinect], researchers can get [motion] data without worrying about how it is obtained in the most of indoor scenes. Now, the research topics have been moving from how to get to shape and motion to how to use that information for applications such as robots and artificial intelligence.*"

We utilize the Kinect shock in difference-in-differences estimations to ascertain the impact of its introduction on measures of researcher-level project choice in the domains of computer science and electrical and electronics engineering research. To measure researchers' rate and type of knowledge output, we rely on bibliometric indicators. Counts of academic publications and citations are established (though imperfect) measures of research output. We utilize these to evaluate the impact of Kinect on researchers' rate of knowledge output (i.e., productivity). Estimating changes in the type of knowledge output presents a greater challenge, as it requires tracking research trajectories in an ever-evolving, multidimensional space of ideas. We focus on two characteristics of the type of knowledge output: changes in the diversity of research projects and changes in the trajectory of research.[3] We propose novel measures of units of knowledge in idea space based on topic modeling analysis, using unassisted machine learning techniques that capture the latent categorization of academic publications without relying on an ex-ante, pre-imposed, structure (such as author-defined keywords or institutionally defined research fields). Our analysis also considers more traditional measures of research diversity and trajectories based on observable characteristics of academic publications. Each measure involves advantages and disadvantages that we discuss in our analysis. Although each measure captures somewhat different attributes of research behavior, we believe that, taken together, they provide an informative window into the effects of automating research technology on researchers' type of knowledge output.

Contrary to concerns that automation displaces human labor and consistent with the idea that knowledge workers face substantial latitude for shifting across tasks, we find that the automation of key research tasks leads both within-area researchers and outside-area researchers to experience (a) an increase in research output, (b) an increase in the diversity of their research, and (c) a shift in their research trajectories. The outside-area researchers experience the largest impact, while within-area researchers experience a more modest effect. Interestingly, outside-area researchers achieve a boost both in their work on motion-sensing research and in the set of their projects that do not directly engage with motion-sensing. For example, an outside-area researcher in our sample who, prior to Kinect, focused on research involving sound waves, extended his research to the study of infant seizures by developing detection techniques combining audio and video inputs. Among the group of within-area researchers, those whose research was most focused on motion-sensing before Kinect experience the highest benefits. This suggests that the automating research technology allows within-area researchers to work more efficiently and to engage in

[3] We consider 'diversification' to be the breadth of researcher's publication portfolio at each point in time *t* and 'trajectory' as the distance in knowledge space between researcher's portfolio of projects at times *t-1* and *t*.

new types of projects. For example, after Kinect, a within-area researcher in our sample engaged in adapting computer vision detection and visualization algorithms to developing malaria diagnostic and tumor identification techniques, while another researcher pushed forward her motion-sensing agenda to include studies of virtual reality. Although these results are consistent with our hypotheses, they are not obvious *ex ante*. Automating technologies often increase returns for incumbents relative to entrants and do not necessarily enable organizational diversification.

Our analysis underscores the significant role played by research technology in knowledge production and sheds light on the processes by which the presence of technologies that automate certain research tasks influences knowledge workers' behavior. Understanding these processes is thus important for the study of the strategic management of science-based organizations (Murray and O'Mahony, 2007; Nelson, 2016) including research-oriented private firms competing for profits and research institutions competing for reputation and opportunities for knowledge creation. Furthermore, our study suggests that firms' decisions to engage in technology development and their technology commercialization strategy might influence knowledge flows and hence subsequent innovation efforts.

II. Automating Technology and Knowledge Production

2.1 Automating technology in knowledge production vs. other economic sectors

While relatively little scholarly work has examined the impact of automating technology on knowledge production, a substantial amount of work has been devoted to understanding the impact of such technological progress on the production of physical goods and, more recently, services. Since the earliest days of the Industrial Revolution, concerns about the impact of automating technology on labor and society have been a recurrent feature of public and academic debate (Mokyr et al., 2015). Over the past few decades, research in management and economics has examined the impact of technology on multiple outcomes, including firm-level outcomes such as productivity (Pinsonneault and Kraemer, 2002; Brynjolfsson and Hitt, 2003), competitive performance (Dos Santos and Peffer, 1995; Bloom et al., 2012), organizational change (Dean et al. 1992; Argyres, 1999; Volkoff et al., 2007), and firm boundaries (Forman and McElheran, 2012). More recently, scholars have begun to study the economic impact of machine learning on individuals, firms, and society (Agrawal et al., 2018; Choudhury et al., 2018).

A persistent question in this literature regards whether automating technologies constitute substitutes or complements for human labor. Most recently, the literature has shifted its focus from approaching this question with a view of whether technology and labor are substitutes overall and towards the recognition that the impact of automation on labor depends on the extent to which technology substitutes for specific tasks and affects the allocation of human effort across tasks within a personnel function (Deming, 2017; Brynjolfsson et al., 2018; Acemoglu and Restropo, 2018; Felten et al., 2018). The question is important for macro-level issues, such as the share of labor in the economy (Karabarbounis and Neiman, 2014), and for micro-level issues, such as the organization of work and the level of firm productivity (Bresnahan, et al., 2002; Bloom, et al., 2012).

An important underlying assumption in tackling this question is that automation is a process that reduces the cost of performing certain tasks. It is this cost reduction that creates incentives for economic actors to substitute automating technology for human labor (Acemoglu and Autor, 2011; Nordhaus, 2007). Thus, when automating technology improves the productivity of a task, individuals and organizations shift human labor away from that task and, under certain conditions, into other tasks.[4]

In knowledge production, we expect similar forces at play but in a context where workers have substantial autonomy in choosing their tasks and a broad array of tasks to choose from. As Smith (1776) observed in his pin factory visits, manufacturing jobs can often involve substantial task specialization. In addition, manufacturing jobs characterized by task specialization are often embedded in relatively rigid hierarchies that provide relatively little worker autonomy. By contrast, the job of knowledge production, particularly for academic researchers, involves both a broad array of potential tasks and substantial autonomy (Aghion et al., 2008; Cohen et al., 2018). The tasks of a researcher are often extremely varied and can include project conception, choice of project partners and co-authors, grant writing, choice of technical and support staff and project execution on a variety of dimensions (e.g., preparation of research materials, conduct of specific analyses, project presentations, conference attendance, collaboration with

[4] In one illustrative and compelling example, Bessen (2015) describes the impact of ATM machines on bank tellers and bank branches between the 1980s and 2000s. He notes that, on one hand, ATM machines constituted a cheaper, strict substitute for the tasks provided by bank tellers and thus enabled the closing of some locations and a reduction in the number of tellers per branch. On the other hand, however, the diffusion of ATM machines induced net gains in bank teller employment by enabling bank branches to be opened in new locations and by enabling tellers to focus on tasks in which ATMs were poor substitutes, such as relationship banking.

industry, etc.). Further, research autonomy, including the freedom to choose research topics and to choose how to allocate time across a variety of research tasks, is the hallmark of scientific scholarship (Merton, 1938). While many service professions also involve a wide breadth of tasks and enable some autonomy (e.g., graphic design, civil engineering, legal practice, investment banking, financial consulting), others have a lesser degree of both (e.g., bank tellers, hair styling, fast food preparation, financial accounting). In other words, we consider knowledge production to be an area in which task variety and autonomy are particularly great and thus a sector worthy of attention, especially considering the impactful role of knowledge production in the economy.

2.2. Research technology automation and research expertise

We hypothesize that technology automation will differently affect researchers with substantial expertise in the research area into which the automating technology becomes available, i.e., within-area researchers, from those who do not have specialized human capital in that area, i.e., outside-area researchers. The case of statistical analysis packages provides an illustrative example. The availability of statistical software such as SAS, Stata, and R automates tasks that would otherwise require substantial human capital investments, including merging data, computing test statistics, inverting matrices, and composing graphics. Researchers could complete these tasks if they made sufficient investments in their human capital. Alternatively, researchers could collaborate with individuals who have made such human capital investments. The ready availability of off-the-shelf software, however, automates each of these tasks, and, via revealed preference, appears to enable researchers to complete their work more efficiently than would be possible in the absence of such software. In addition, the software packages may enable researchers to shift the time made available by the automation of statistical analysis to other research activities including undertaking new and different projects. While these effects likely hold for all researchers who conduct quantitative analysis, they likely differ in their impact and nature for experts in statistical analysis, e.g., econometricians, relative to other researchers. Econometricians might be displaced from their core research tasks and that could lead to exit or to additional time to be spent on other projects such as advancing the field of statistical analysis. For other researchers, the availability of the technology could enable them to employ the technology as a substitute for performing the task using specialized human capital.

As this example suggests, the automation of some core research tasks could affect the knowledge production function in several ways and may do so differently for researchers who have developed specialized human capital related to the now automated tasks and for those who have not. We formalize these thoughts in a set of hypotheses below.

2.2.a. Research technology automation and within-area researchers

We begin by focusing on those researchers with historical expertise in the domain affected by the automation. The automation of core research tasks can have multiple effects for these individuals. If a researcher's entire efforts are focused on the area in which the automation occurs, it is possible that the automation will substitute for her efforts and her productivity will decline or that she will exit the domain of research or the conduct of research entirely. This would be consistent with Autor et al. (1998), who highlight the prospect that technology substitutes for related human capital. If, however, the automated tasks constitute a consequential fraction of her tasks but not the full set of her tasks, she may reallocate her time to these other tasks. Thus, as the productivity of the automated tasks rise while the cost of the other tasks remains constant, we expect the productivity of such researchers to improve.[5] Considering the variety of tasks associated with research and the extensive autonomy of most researchers, we anticipate a positive net effect of automating technology for within-area researchers.

> *H1a: The availability of research technology that automates core research tasks boosts the rate of research output of within-area researchers*

An additional first-order effect of research task automation regards its influence on researcher's type of knowledge output. Aghion et al. (2008) suggest that research technology that affects the productivity of domain-specific knowledge may induce greater mobility across research trajectories. The idea is that the increase in productivity of the automated task allows these within-area researchers to spend the additional time on either experimenting with new topics or applying their expertise to projects in other domains. Thus, the availability of automating research technology may boost the opportunities set among within-area

[5] This requires the additional assumption that the standards for quality for publication in the domain of the automation do not change in response to the automation. We anticipate that shifting standards in the affected domain would be a second-order impact of research tool automation, though we hope that this issue will be examined more fully by further research.

researchers, facilitating tackling new research questions within their domain of expertise or branching out into other research trajectories.

> *H1b: The availability of research technology that automates core research tasks changes the composition of research project types for within-area researchers.*

2.2.b. Research technology automation and outside-area experts

Researchers that have not invested in the specialized human capital associated with the tasks automated by the research technology can be affected by the automation in ways that differ from those of the within-area researchers. A core insight here is that an automating research technology that reduces the cost of performing certain research tasks compensates for the fixed-costs of research expertise within the knowledge area of the technology (Cohen and Klepper, 1992, 1996). This, in turn, reduces returns to that within-area knowledge and, hence, opens opportunities for outside-area researchers. For example, Furman and Stern (2011) and Murray et al. (2016) suggest that open access to research tools can exert a democratizing influence on other fields and may compensate for the fixed-costs of within-area expertise. In addition, outside-area researchers' lack of specialized skills in the affected domain means that they are less likely to experience substitution in any of their research tasks. Thus, if these researchers adopt the automating technology, it will have a positive impact on their research productivity. Furthermore, the impact could be higher than that for within-area researchers who experience a boost in productivity through some degree of task substitution.

> *H2a: The availability of research technology that automates core research tasks boosts the rate of research output of outside-area researchers on all topics on which they work. The effect is greater than for within-area researchers.*

In addition to affecting productivity, we anticipate that the automating technology will influence outside-area researchers' type of knowledge output. Teodoridis (2018) shows that a democratizing change in the availability of research technology alters team composition to include specialists from knowledge areas outside the domain of the research technology. To the extent that changes in collaboration reflect changes in researchers' project choices, this suggests the possibility of a change in the project portfolio composition of outside-area researchers. Moreover, we expect that the automating research technology will influence the composition of research project types of outside-area researchers by creating opportunities for novel idea recombinations (Uzzi et al., 2013) and, thereby, new lines of inquiry. Thus, we expect that outside-area researchers can leverage the availability of automating research technology to explore ideas that incorporate the automated tasks either directly in the research domain of the technology or by broadening lines of inquiry in their areas of historical focus. As before, we expect the impact to be higher than that for within-area researchers who also experience a change in the opportunities they face but through some degree of task substitution.

> *H2b: The availability of technology that automates core research tasks changes the composition of research project types of outside-area researchers. The effect will be greater than for within-area researchers.*

III. The Kinect shock & the automation of motion-sensing research

The shock we examine in this paper is the launch of the Microsoft Kinect gaming system that suddenly and dramatically made available a research technology that automates certain motion-sensing research tasks. We exploit the unexpected availability of this technology as an instrument that is correlated with task automation in knowledge production but not with factors affecting researchers' efforts, other than through its effect on automation.

The setting appears unusual at first glance: It is the result of the unexpected impact of Microsoft's successful launch of a controller-free video game system designed to compete with rival products launched by Nintendo and Sony. In the two months following Kinect's launch on November 4, 2010, Microsoft sold more than 8 million units (>130,000 Kinect units/day), outpacing the iPhone and the iPad to become the Guinness World Records' all-time fastest-selling consumer electronic device (Bilton, 2011). The surprise that makes Kinect valuable for our research context is not, however, its commercial success but its wide-ranging and near-immediate impact on scholarship in some areas of computer science and engineering.

III.1. Microsoft's introduction of the Kinect system

On November 4, 2010, Microsoft introduced the Kinect system for its Xbox video game console with the aim of competing with handheld gesture-recognition remotes introduced previously by Nintendo (Wii) and Sony (PlayStation). With the Kinect, Microsoft attempted to leapfrog its video console rivals by creating the first hands-free controller for electronic devices, a game controller system that responded to the natural movements of the player.[6]

In addition to being a treat for video gamers, Kinect also constituted a feast for hackers, who descended upon the system with the aim of accessing the vast data obtained by Kinect's sensors and linking it to other devices, such as computers and robots. These efforts received a twofold infusion of interest on Kinect's launch day. The first came from Adafruit Industries, a manufacturer of do-it-yourself electronics kits operated by alumni of MIT's Media Lab, Limor Fried and Phillip Torrone, that offered a $1,000 prize for the first individual or organization to post an open-source Kinect driver to GitHub (Carmody, 2010a). The second spur of interest arose as a result of Microsoft's actively (and quixotic) effort to thwart the hackers. In a same-day response to Adafruit's prize offer, Microsoft released a statement to CNet: "*Microsoft does not condone the modification of its products… [and will] …work closely with law enforcement and product-safety groups to keep Kinect tamper-resistant*" (Terdiman, 2010). Adafruit responded immediately by doubling its Kinect driver bounty to $2,000, further intensifying the race for the driver.

The race was won on November 11, by a Spanish computer science undergraduate student, Héctor Martín, who did not own an Xbox but who had purchased a Kinect that morning when it went on sale in Europe (Giles, 2010). Within days of the driver's release, researchers and hobbyists had adapted Kinect for numerous uses, including the creation of 3D computer holograms and a modified iRobot Roomba that could respond to human hand and voice commands and could create visual maps of the rooms it had visited (Wortham, 2010).

During the week that hackers had raced to create an open-source driver to harvest Kinect's data, consumers purchased nearly a million Kinect units. In the wake of its success, Microsoft initially continued to resist working with the hacker community and even refused to acknowledge that its system had been hacked. Within ten days of the release of Martín's open-source driver, however, Microsoft, convinced of the value of embracing the experimentation of the hobbyist and scientific communities, pivoted entirely

[6] The Kinect system was composed of a color camera, a depth sensor, and a multi-array microphone. These physical features, along with artificial intelligence pattern recognition software, enabled Kinect to recognize in three dimensions and in real time the movements and facial expressions of multiple individuals and to acknowledge and distinguish their voice commands (Zhang, 2012). The Kinect system translated this motion-sensing information into actions enabling players to control gameplay.

and announced that it would not pursue legal remedies against those who adapted the Kinect system for other purposes (Carmody, 2010b).

III.2. Microsoft Kinect as an automating technology

Motion-sensing research involves a series of topics in the broader research area of computer vision, an interdisciplinary field in computer science and electrical engineering with the principal goal of enabling machines to “see” as well as (or better than) humans. Achieving this goal requires the ability to scan 3D environments and recognize where objects appear in both static and dynamic environments and under various lighting conditions. While often taken for granted by humans, vision is an exceptionally complex task for machines. Humans use their eyes to observe 3D scenes and their brains to process the types of objects and movement in the scenes. Machines rely on cameras as their *eyes* and on human developed software and hardware as their *brain*.

Prior to the introduction of Kinect, there were two main approaches to motion-sensing in computer vision. In a first generation, researchers employed 2D cameras and developed software as the sole method of inferring depth and movement based on the images captured by those cameras. This process was complex and time-consuming, as it required sophisticated mathematical calculations, such as those required to infer depth based on known object sizes and movement from serial images. Such custom-developed software was not particularly fast to run, and it was prone to errors because it relied heavily on factors such as the resolution of the images, the frequency of frames, and the ingenuity of the software developer in coding efficient calculations. The second generation of this work involved advances in images captured from 2D cameras to 3D cameras, such as time-of-flight cameras. These early 3D cameras operated with low resolution, were subject to especially high sensitivity to lighting conditions, and were not particularly accurate in tracking movements in all three dimensions. While the availability of such cameras lessened the need for some of the coding development tasks required to process the image output, a broad set of such tasks remained even after these cameras were introduced, including the needs to clean image data, accurately trace movement, speed up the processing of the image output, infer missing pieces in various unfavorable lighting conditions, and impute missing information resulting from low resolution images.

While some algorithms for working with these cameras and their output could be shared, the vast majority needed to be customized for the environment in which individual researchers were working.

By offering a higher resolution 3D images and an embedded processing capability that more accurately traced movement under a variety of lighting conditions, the introduction of Kinect represented a significant technical advance that eliminated a substantial set of coding tasks that had previously fallen on motion-sensing researchers. The advance enabled by Microsoft's new tool was sufficiently great that it enabled a novice to extract and use Kinect's output without the need for specialist coding to process the data to render it useful. As a result, Kinect automated many tasks that previously required specialized humans involved in developing software algorithms to process images captured by cameras with a goal of uncovering the same insights that a human would when observing a 3D scene in movement. Kinect did not fully resolve all challenges associated with computer vision, e.g., Kinect performed poorly in bright sunlight, but automated a substantial set of tasks that had previously required specialized human coding.

The impact of Kinect was perceived by the computer vision research community to be wide-ranging. For example, a computer vision researcher told us that *"Kinect had a game-changing effect on the research possibilities. We work in robotics perception, i.e., how can robots perceive and act in the environments. Since our world is 3D and Kinect gives 3D information, the data becomes extremely powerful. This has enabled significant advances in applications such as object detection, human activity recognition and anticipation for robots, as well as robotic grasping and path planning."* In addition to its impact in computer vision research, Kinect raised the interest of researchers from other domains. Richards-Riessetto et al. (2012) describe the value of Kinect for work in archaeology, and Rafibakhsh et al. (2012) describe its value for construction engineering and management. In our own discussions with motion-sensing researchers, a researcher noted that his group provided a Kinect-based algorithm developed in his lab to "*ICT's Medical VR group which applied [it] to various motor rehabilitation applications. We've also researched the use of Kinect for human activity analysis (examining body language, fidgeting, etc.) to help therapists understand behavior of their patients […] Our lab's director and some of his students used the Kinect to create really interesting education game experiences. We have put together a number of other experiments, prototypes, and demonstrations involving the Kinect. It runs the gamut from wide area* *tracking to virtual human puppeteering to 3D scanning, and more. I can't list them all."* More broadly, after the launch of Kinect, motion-sensing appears to have found its way into an increasing variety of research projects with applications in a wide set of domains, from artificial intelligence and virtual reality to education, healthcare, music, cinematography, market research, and advertising. For example, faculty and graduate students at MIT's CSAIL laboratory have designed a motion-sensing system, called Emerald, which tracks individuals' movements within their homes, can alert medical personnel in the event of a medical catastrophe or fall, and can even be used to predict fall events.

In addition to reducing the need for specialized human capital investments, a secondary impact of Kinect is that it also reduced the monetary cost of capturing and leveraging data from 3D images. As we noted earlier, an important underlying aspect of automation that explains the incentives it generates for economic actors to substitute the technology for human labor is that automation is a process that reduces the cost of performing certain tasks with human labor (Acemoglu and Autor, 2011; Nordhaus, 2007). While Kinect indeed reduced the cost of executing certain research tasks with human labor, e.g., software coding, the monetary cost of Kinect was lower than that of the previous generation of 3D cameras: whereas the cost of time-of-flight cameras ranged from $10,000 to $20,000, Kinect cost $150 at launch. To sharply isolate the reduction in cost via automation, we would have needed for Kinect to be priced in the same range as the previous generation of motion-sensing technology (e.g., time-of-flight cameras). While this remains a limitation of our study, we believe the dominant effect is that of the reduction in cost through automation. The reason is that for researchers committed to computer vision research (within-area researchers), this change in monetary cost was likely not especially great relative to the overall cost of operating their laboratories. In other words, we argue that the Kinect impact would have been roughly the same even if Kinect was to be priced in the same range as the previous generation of motion-sensing technology. For outside-area researchers, however, the change in cost may have enabled experimentation that these researchers would not have considered if Kinect was priced in the same range as time-of-flight cameras. It may be, therefore, that changes in the monetary cost of motion-sensing research technology drive some of the effects we observe. We believe, however, that the more substantial change enabled by Kinect is not related to the price of the research technology but is more related to the fact that the advent of Kinect obviated the need for specialized human capital and enabled researchers in other areas to appreciate the potential value of motion-sensing techniques to sets of problems related to their areas of interest.

Overall, the launch of Microsoft Kinect appears to have changed the opportunity set for innovation in computer science and engineering. This development appears to be exogenous and to have been a surprise to the incumbent research community, the community of potential users that had been working outside traditional motion-sensing topics, and even to Microsoft itself.

IV. Data and Empirical Strategy

To examine the impact of Kinect on researchers' productivity and portfolio of project types, we draw on the population of publications, early-access publications, and conference proceeding papers included in the IEEE *Xplore* database, which covers nearly 200 computer science and electrical engineering journals and more than 1,800 conference proceedings, between 2001 and 2014.[7] We conduct our estimation on the subset of papers published in the four years before and four years after the launch of Kinect (2007-2014), and we use the remainder of the data to obtain better estimates of researchers' pre-Kinect research behavior and trends.

IV.1. Empirical Strategy

We employ a differences-in-differences analysis to compare research productivity and type of knowledge output before and after the launch of Kinect. Formally, we estimate for researcher *i* and year *t*:

$$DV_{it} = \beta(TreatedResearcher_i * AfterKinect_t) + Age_i + Age_i^2 + \delta_i + \gamma_t + \epsilon_{it} \quad (1)$$

where $TreatedResearcher_i$ is a dummy variable equal to 1 if research *i* is a treated unit, and 0 otherwise. We define treated researchers as individuals who were publishing in motion-sensing before the launch of Kinect or individuals who started to publish in motion-sensing only after the launch of Kinect. To identify motion-sensing publications, we search the full text and metadata of publications in the IEEE *Xplore* database using carefully identified keywords, through interviews with subject matter experts and cross-referenced with IEEE's taxonomy.[8] $AfterKinect_t$ is a dummy variable equal to 1 if the observation year is between 2011 and 2014, namely after Kinect's launch, and 0 otherwise. We also control for a quadratic

[7] The data were collected during late 2014. Hence, there is a truncation in the 2014 data coverage, which is evident in Figures 1-4 as a decrement in estimated effects in the year 2014. All results are robust to dropping the 2014 data.
[8] These terms are available upon request and are described in greater detail in Teodoridis (2018).

effect of individuals' (research) age, calculated as the number of years since the occurrence of the first publication in our large dataset, starting in 2001. The term δ_i represents individual fixed effects and controls for time-invariant individual attributes. The term γ_t captures year-specific fixed effects that account for changes in publication trends over time. As a consequence of including individual and time fixed effects, the terms $TreatedResearcher_i$ and $AfterKinect_t$ drop out of the estimating equation.

We exploit three categories of dependent variables of interest DV_{it}. First, we estimate changes in researchers' productivity by employing two measures of output: (a) $PubCount_{it}$ captures the number of academic publications of researcher *i* at time *t*, and (b) $CitationWeightedPubCount_{it}$ captures the number of academic publications of researcher *i* at time *t* weighted by the number of cumulative citations received until 2014. Second, we distinguish between two concepts capturing changes in researches' portfolio of project types: (a) the extent to which a researcher's *i* portfolio is concentrated or diverse at time t ($Diversification_{it}$) and (b) the extent to which a researcher's *i* portfolio involves topics that are closer or further from each other in ideas space at time *t-1* vs. t ($Trajectory_{it}$). We do this because a researcher's work could be concentrated in a small number of domains, but these could, theoretically, be quite distant from one another. For example, a researcher whose work includes topics in only labor economics and materials science would have a high research concentration, though her projects would be quite distant in ideas space. We provide details on the construction of these measures in the next section.

Our main coefficient of interest is β. We interpret a positive value of β as indicating a higher increase in productivity, diversification or trajectory shift, respectively, at the individual level for treated researchers after the launch of Kinect, when compared with the change of matched researchers from before to after the launch of Kinect. In other words, a positive value of β indicates a positive effect of the research technology on researchers' rate of knowledge output, or their propensity to diversify their research projects or to shift their research trajectories.

We construct a plausible counterfactual using coarsened exact matching (CEM; Iacus et al., 2011, 2012) based on individual researcher characteristics in the before period (2007-2010). Specifically, we match on (1) yearly productivity in each of the four years before Kinect's launch, (2) the number of co-authors in each of the four years before Kinect's launch, (3) a measure of diversification across knowledge topics between 2007 and 2010, and (4) distance in knowledge space to the motion-sensing domain of knowledge before the launch of Kinect. We measure yearly productivity as a count of publications weighted by citations, and diversification as 1 minus the Euclidean distance in the space of IEEE-defined research categories.[9] We define distance in knowledge space to motion-sensing based on the network of authorship with within-area researchers. Specifically, we label within-area researchers as being the closest to the motion-sensing domain of knowledge (distance one), followed by individuals who coauthored with within-area researchers on other, non-motion-sensing, projects (distance two), and followed by coauthors of coauthors of such researchers (distance three). All other researchers are categorized as being the furthest away in knowledge space from motion-sensing, to a total of a four-level distance categorization. We conduct two matching procedures, one for each of our two sets of researchers: within-area researchers and outside-area researchers. We include our measure of distance in knowledge space to motion-sensing only in the latter matching procedure since, by construction, all researchers with an assigned distance of one are labeled as treated in the former sample. We employ CEM with weights rather than one-on-one matching to use as much of the available data as possible.

We match on productivity in the before period to ensure that our results on changes in productivity, diversity and trajectory shift at the individual level are not confounded by researchers at the right tail of the productivity distribution. We match on the number of co-authors in the before period to ensure that our results are not driven by researchers' abilities or preferences for collaborating more intensely or more broadly. This matters for the analysis, since higher levels of collaboration could be correlated with more diverse output or with changes in research trajectory as each new collaborator increases the potential pool of expertise and perspectives. We also match on the level of diversification in the pre-Kinect period to ensure that the results are not driven by individuals with a taste for exploring new avenues that may manifest regardless of research technology availability. Finally, we match on distance to motion-sensing to ensure

[9] IEEE assigns each publication to one of the 51 main categories listed in its taxonomy. We calculate the Euclidean distance based on the percentage of keywords from each category that a researcher collects in her publication portfolio between 2007 and 2010. Formally, we calculate:

$$DiversificationIndex_i = 1 - \sqrt{\sum_{k=1}^{51} CategoryPercentage_{ik}^2} \quad (2)$$

where i is the individual researcher and $CategoryPercentage_{ik}^2$ represents the squared percentage of keywords assigned to researcher i's publications in each of the k main 51 categories of the IEEE taxonomy. Note that, by construction, the measure is less than or equal to 1 and never 0, and it increases with higher levels of keywords spread across IEEE categories.

that our results are not driven by proximity to the motion-sensing field that would influence researchers' project choices regardless of the availability of motion-sensing technology. Our results remain robust to using the full set of researchers in fixed-effects estimations, under the assumption that all researchers publishing in IEEE outlets are at risk of engaging with new technological developments in their research (Appendix A). Furthermore, raw data trends displayed in Appendix A indicate the presence of parallel pre-trends ensuring that our CEM procedure does not impose this structure to a phenomenon that follows a different pattern.

IV.2. Measuring changes in the portfolio of project types

We attempt to distinguish between diversification and changes in research trajectories because the two concepts capture research behavior that reflects distinct features of a researcher's portfolio of project types. We define diversification as the breadth of researcher's portfolio of projects at one point in time *t*. We define trajectory as the distance in knowledge space between a researcher's portfolio of projects at times *t-1* and *t*. The measurement challenge is that estimating such changes requires delineating the boundaries of research trajectories. This involves a paradox, however, as the boundaries of research trajectories are part of the core unknown to be estimated. Unlike physical space, which consists of a well-known number of dimensions and distances between locations, ideas space consists of an unknown number of dimensions, the distances between which cannot be uniquely measured, and which evolve over time in ways that cannot be anticipated until they are realized (Doran, 2017).

Aghion, Dewatripont, and Stein (2008) model the development of ideas along research trajectories. In some cases, such as in mathematics, fields are relatively well defined, and stable field distinctions can form the basis for inquiry about location and movement in ideas space (Borjas and Doran, 2012, 2015; Agrawal et al., 2016). In most fields, however, it is difficult to measure such trajectories or to identify where they branch. To overcome these challenges, scholars often focus on measures of research breadth (Grupp, 1990; Rafols and Meyer, 2010) or other measures of topic overlap to reflect whether researchers are roughly in the same domain (Boudreau et al., 2017), or consider the development of ideas based on references to papers in a stream of research (e.g., Furman et al., 2012), or rely on a professionally-curated tools like the PubMed Related Article Algorithm (Azoulay et al., 2015; Myers, 2018).

These analyses, however, typically rely on observable indicators of innovation output, such as keywords, taxonomies, or citation maps. While helpful in providing some evidence of the evolution of research trajectories, these approaches face many of the challenges described above. For example, research that categorizes trajectories based on curated taxonomies has the advantage of consistency but faces the trade-off of either being stable and therefore comparable over time, though at the cost of inflexibility, or being dynamic and therefore evolving with the changing research landscape, though at the cost of consistency and classification standardization. Author-assigned keywords, or any other set of keywords not extracted from a defined vocabulary, fare better in capturing new knowledge trajectories but lack structure and may be more subject to gaming. This also limits the interpretability of hierarchical connections between keywords and changes in such relationships over time. Similarly, while the citation revolution, as a method for tracing knowledge linkages (Griliches, 1990), significantly helped advance our understanding of factors influencing the process of knowledge creation, it is subject to the same concerns, since measuring diversity in citation maps requires some form of categorization. Like author-assigned keywords, the selection of backwards and forwards citations is subject to social processes and strategic behaviors that complicate their interpretation for understanding changes in research trajectories.[10]

The research context we examine is not characterized by a relatively stable set of keywords and research topics. Indeed, the past two decades have seen the emergence of many new domains of research and associated new keywords enabled by ever-advancing computing power and methods within these fields. To measure research diversification and trajectory in these fields, we leverage advances in machine learning to develop measures that make use of more complete information in academic publications. We propose measures based on topic modeling algorithms, which we have adapted for inference in our context. Our empirical analysis also includes a set of more traditional measures of research diversification and trajectory based on observable characteristics of academic publications, including a measure of diversification based on the stable taxonomy maintained by the IEEE, a measure of research trajectory as a count of new authors, and a measure of research trajectory as a count of new publication outlets.

[10] Alternatives to using bibliometric measures to capture movement in ideas space do exist. For example, Krieger, Li, and Papanikolaou (2018) use the Tanimoto distance, which reflects the similarity of chemical structures, to measure mobility in ideas space in the pharmaceutical industry.

The main advantage and, hence, contribution of measures based on machine learning analysis,[11] is the ability to identify similarities between bodies of text without predefined assumptions about their structure. Note that the intended purpose of these algorithms is prediction, not inference. Their success rests with their ability to reveal the latent structure of a corpus of texts in order to predict with high accuracy where a new text would fit in the structure. We are interested in identifying the latent categorization of research publications in ways that (a) are less subject to the strategic behavior of researchers and (b) are sensitive to the fact that research fields evolve over time.

In the service of these objectives, we employ Hierarchical Dirichlet Process (HDP) (Teh et al., 2006)[12], which we adapt for our purposes. The algorithm falls into the topic modeling category of unassisted machine learning. HDP is a probabilistic model that employs a hierarchical Bayesian analysis of text (see, e.g., Hofmann, 1999; Teh et al., 2006; Buntine and Jakulin, 2004). The intuition is that of a generative process, in which the data are assumed to be characterized by a set of observed variables (words in the document or vocabulary) that develop from a set of hidden variables (the topic structure) (Teh et al., 2006). The algorithm generates collections of words (topics) that are found to appear together in the input text with a certain probability. In other words, the input text is 'assigned' to topics with a certain probability.

We conduct our analysis using the abstract of the academic publications in our dataset as input text into HDP.[13] We run the algorithm per year for the full set of publications available in our dataset. We modify the algorithm to output the set of words describing each topic[14] and to list the publication IDs of each abstract used to identify those topics. Each publication ID is assigned a score, which can be thought of as a probability of 'belonging' to a topic. All scores add up to 100% probability per publication ID.

The algorithm has advantages and limitations. First, HDP has the advantage of identifying the optimal number of topics per corpus of text analyzed. This differs from other topic modeling algorithms, such as

[11] Machine learning algorithms evolved from the study of pattern recognition in computer science but have increasingly found applications in a variety of fields, including genetics, medical imaging, computational biology and bioinformatics, image recognition, social network analysis, and economics and public policy (Athey and Imbens, 2015). Currently, there are a large number of algorithms customized for various tasks. While limitations remain, their complexity and accuracy are rapidly evolving.

[12] Hierarchical Dirichlet Process (HDP) is a more advanced version of the more well-known Latent Dirichlet Allocation (LDA).

[13] Despite substantial advances in computing power, each process is computationally intensive. Each run of HDP using our data requires days of computing time. As a result, we have run the current analysis allocating papers to topics using article abstracts but not full text. In addition, the large volume of data in our data-set prevents us from running the HDP or LDA algorithms on multiple years of data at once to e.g., compare the structure generated by these algorithms with the IEEE taxonomy that covers the entire corpus of publications across all years.

[14] We consulted with experts in computer science, electrical engineering, and electronics to ensure that the topics identified by the HDP algorithm reflect credible categorizations in this line of research.

Latent Dirichlet Allocation (LDA), that require the analyst to input the number of topics into which they would like the algorithm to group the text. We also run a sensitivity algorithm on multiple instances of LDA, consistent with state-of-the-art practices in computer science, to identify the number of optimal topics with the highest probability of accuracy.[15] Second, both algorithms treat the input text as a one-time group for which the latent categorization needs to be revealed. In other words, the algorithms cannot automatically track the evolution of topics over time by updating the set of keywords in each category over time. We address this limitation by calculating a cosine vector similarity between the yearly topics generated by the HDP algorithm. Specifically, we employ a Term Frequency – Inverse Document Frequency (TFIDF) cosine similarity where the frequency of words is weighted by the HDP-generated score that captures the relevance of each word for each topic. In addition, we use the HDP output in a regression with time fixed effects; hence our results are not hindered by the fact that the algorithms are executed on a per-year basis and thus reveal the latent categorization of topics for each year in our dataset.[16]

We calculate diversification as a yearly measure of spread across categories, as identified by the HDP algorithm. Specifically, we calculate an intensive measure of diversification equal to the sum of the number of topics where the focal researcher has their papers assigned by the HDP in the focal year. We also calculate an extensive measure of diversification equal to a count of unique topics where the focal researcher has their papers assigned by the HDP in the focal year. We also present results using more traditional measures of diversification, based on publication attributes. Specifically, we calculate a yearly reversed Euclidean distance in the space of fifty-one IEEE categories in computer science, electrical engineering, and electronics. To calculate this measure, we apply equation (2) to yearly publication data over the period of interest (2007-2014), four years before and four years after the launch of Kinect.

To generate measures that capture changes in research trajectory, we use the yearly topics generated by the HDP and the cosine similarity index between such topics. Specifically, we first calculate the distance between topics in consecutive years as one minus the similarity index between all such topic pairs. Next,

---

[15] All our results remain robust to using an LDA algorithm with 40, 60 and 90 topics. We chose the number of topics based on the optimal number of topics identified by the HDP algorithm, crossed checked against a sensitivity algorithm on multiple instances of LDA. We do not include these results because of space limitations but are happy to provide them upon request.

[16] Computer scientists are working on a variety of extensions of these algorithms. We chose to use algorithms that are considered robust among computer scientists rather than current experimental ones aimed at advancing the frontier in topic modeling.

for each researcher, we sum the distance between all pairs of topics in year *t-1* and *t* and then divide the sum to the number of unique topics covered in year *t-1*. The measure captures the average number of new topics for researcher *i* in year *t* relative to year *t-1*, weighted by the distance between the topics. One can think of this measure as capturing the yearly new areas of interest; the smaller the value, the less of a change in interests from one year to the next.

In addition, we employ two other measures based on traditional publication output. The first such measure counts the number of new co-authors that the focal author has in the observation year relative to previous years. To count the number of new co-authors, we take advantage of our full dataset going back to 2001. We do so since, by definition, the count of new co-authors requires a few years of reference data to get closer to reflecting the true number of new co-authors and to not be upward-biased due to left-side data truncation. This measure indicates changes in research trajectories to the extent that collaboration patterns reflect changes in the bases of expertise associated with a researcher's project choices. The second measure reflects the number of new publication outlets in which a researcher publishes each year, relative to each prior year, going back to 2001. This measure indicates changes in research trajectories to the extent that different journals address different audiences and cover different areas of ideas space.

Each measure has its own limitations and merits. The HDP-based measures are more flexible in capturing changes in knowledge space over time. However, the HDP categorization lacks stability and trackability over time. The IEEE diversification measure fairs well on these dimensions, but its disadvantage stems from the same attributes that we discussed above. In particular, the fixed taxonomy fails to capture changes in the categorization structure over time that would otherwise indicate changes in research trajectories. Similarly, the measures based on counts of new co-authors and new publication outlets, while perhaps easier to grasp than interpreting the HDP measure, are focused on outcomes that indirectly reflect the content or intellectual focus of academic research. Specifically, it is possible to change co-authors and publication outlets while continuing to work on the same knowledge trajectory, and it is also possible to continue working with old co-authors and publish in the same journals while shifting one's research focus.

As a result, we prefer our more novel measures, which we think are more likely to reflect changes in research diversification and trajectory because they are more tied to the content of researchers' published works. At the same time, we believe the other measures complement the HDP measures and enrich the insights than can be drawn from our analysis. Each measure captures different attributes of diversification and trajectory. Taken together, we argue, they paint an informative picture of changes in researchers' projects triggered by the availability of automating research technology.

III.3. Sample construction and descriptive statistics

We collect data on every publication, early-access publication, and conference proceeding paper available through IEEE *Xplore* between 2001 and 2014. These data include 2,492,451 publications and 1,670,888 unique author names in the fields of computer science, electrical engineering, and electronics. Because of the importance of publications in conference proceedings in computer science and related fields, the IEEE possesses advantages relative to other libraries of publications, including the Web of Science and Scopus.

We focus our analysis on the four years leading up to and the four years following the launch of Kinect at the end of 2010, i.e., 2007-2010 and 2011-2014. We do so to ensure comparable timeframes and to allow for some publication data for controls and other measures that require a longer-run observation of publication trends, such as author age and changes in the number of yearly new co-authors and new publication outlets. The 2007-2014 dataset consists of 1,776,125 publications authored by 1,391,313 individuals as identified by IEEE. Within this subset, we further distinguish between researchers active both before and after Kinect's launch (430,779), only in the period before 2010 (442,395), and researchers who enter the sample after 2010 (518,139). We do so (1) to ensure that our main results are not driven by zeroes due to exits from or entry into our observation period and (2) to allow for an observable period before Kinect's launch from which we can identify trends and construct plausible counterfactuals. We focus our main analysis on the first dataset and we further eliminate outliers, namely individuals with fewer than three and more than 50 publications before Kinect's launch (2007-2010). We eliminate researchers with fewer than 3 publications since some of our measures rely on individuals' breadth of publications and low productivity mechanically translates into low diversification. Note that this set of researchers also includes authors who occasionally publish in outlets tracked by IEEE. We eliminate authors with more than 50 publications in the before period to account for potential disambiguation effects in the IEEE algorithm assigning unique author identifiers. Such author IDs would appear as very productive and potentially diversified individuals, risking an upward bias to our diversification-based estimations. We identify a total of 3,200 researchers with over 50 publications in the four-year period before Kinect's launch, less than 1% of our main sample.[17]

Our final sample includes 12,549 within-area researchers, 9,590 outside-area researchers, and 160,845 other researchers. We present descriptive statistics that show the balance in our CEM procedure for within-area and outside-area researchers in Appendix B, Tables B1.a and B1.b. We show average values for all our covariates used in the matching procedure, in both the full sample (Columns 1-3) and the matched sample (Columns 4-6). Across all definitions, treated researchers are more productive.[18] They have a higher number of co-authors and publish in more IEEE categories than the rest of the population. However, this fact results from a more skewed distribution in the full sample. The CEM procedure balances these observables for each of our two definitions of treated researchers. Our sample size is reduced by the matching procedure and by our eliminating outliers, as described above.

In Tables 1.a-1.c, we present average values of all our dependent variable measures across the two groups of treated researchers: two measures of publication rate, three measures of diversification and three measures of research trajectory. The descriptive statistics foreshadow our main findings: (1) an increase in quality publication output for both within- and outside-area researchers, (2) a positive impact on diversification that is more pronounced for outside-area researchers than for within-area researchers, (3) a positive impact on the trajectory of research that is more pronounced for outside-area researchers than for within-area researchers. Further, we observe differences across all our measures relative to baseline trends. Some of the effects come from mitigating the overall decrease in productivity, diversification or trajectory shift, while others come from an accentuated baseline trend of increasing productivity, diversification or trajectory shift respectively. Specifically, the increase in productivity measured by the count of academic

[17] All core results are robust to altering these cutoff choices and to including the full set of data available to us. Specifically, the results remain robust (1) to considering the full dataset, 2001-2014; (2) to considering other cutoff points for the minimum and maximum number of publications; and (3) to eliminating cutoffs for the minimum and maximum number of publications and utilizing the full set of 2007-2014 authors.

[18] Our dataset is a balanced panel where the productivity in non-publishing years is taken into account as years with zero publications.

publications is absolute but also offsets an overall trend of a decrease in publication output. Our measure of productivity based on publication output weighted by citation count offsets an overall downward sloping trend that can be explained not only by the average decrease in academic publication, but also by the skewed nature of citation accumulation over time. Similarly, our Euclidean measure of diversification, which captures the information contained in IEEE's taxonomy relative to changes in ideas space, shows an overall trend of decreasing diversification, with the research technology dampening that effect. The same trends can be observed in our measure of diversification based on counts of new publication outlets. This is important for at least two reasons. First, it underscores our point regarding the difficulty of crafting all-encompassing measures of diversification and research trajectory. Second, it highlights the differences in the various aspects of the knowledge creation process captured by each measure. For example, the overall trend of decrease in diversification as captured by the Euclidean measure appears to contrast with the overall trend of increase in diversification captured by our topic modeling measures. This apparent inconsistency may, however, result from an increase in the fragmentation of research trajectories, an effect aligned with the "knowledge burden" effect (Jones 2009, 2010) that cannot be captured by a measure that relies on a fixed taxonomy. More broadly, the differences suggest a need for approaching studies on changes in research diversification and trajectories using empirical strategies that triangulate across multiple measures.

V. Empirical Analysis

V.1. Did the Kinect shock induce changes in researchers' productivity?

We begin our analysis by estimating the impact of Kinect on the researchers' productivity. To do this, we estimate equation (1) using the annual count of academic publications at the individual level as the dependent variable as well as an annual count weighted by citations accumulated by 2014. We should note that we observe this effect in descriptive data, as the average number of publications for within-area researchers increases from 1.18 to 1.25, while in the matched control group the average number of publications changes from 1.21 to 1.15. In the group of outside-area researchers, the average number of publications changes from 1.07 to 1.68, while in the control group it changes from 1.09 to 1.02.

We show the results of our main difference-in-differences estimation in Table 2.a. In Column 3, we estimate the effect on the productivity of within-area researchers. We find evidence of a 9% increase in publication count for within-area researchers when compared to matched controls, or one additional paper for every ten, in line with hypothesis H1a. Column 2 shows results for outside-area researchers, and Column 3 shows results for the same group of outside-area researchers when excluding their motion-sensing publications. In both cases, we find evidence of an increase in the productivity of such researchers, when compared with the group of control researchers, as identified through our CEM procedures, in line with hypotheses H2a. Treated outside-area researchers increase their number of publications 65% more than the control group, an effect that corresponds to an increase of three additional papers for every two.[19] In Column 3, we show evidence that the increase in publication output is not driven by publications in motion-sensing. This is important because it suggests that the effect of research technology does not lead mechanically to increases in productivity via direct engagement with the technology, but rather the access to the research technology impacts the productivity of outside-area researchers on all topics on which they work, in line with our hypothesis H2a. Specifically, outside-area researchers experience a disproportionate increase of 36%, approximately one additional paper for every three publications, in their publications on topics outside of motion-sensing. Furthermore, the magnitude of the effects is higher for outside-area researchers than for within-area researchers,[20] in line with hypothesis H2a. All these trends persist when focusing on a measure of publication output weighted by citations (Table 2.b), suggesting that the increase in productivity does not arise as a result of a jump in studies of lesser quality.

Next, we examine the timing of this effect in order to ensure that our estimates of the boost in productivity are not driven by secular trends towards increased publication or changes in productivity that occur later during our after period, and which could therefore be attributed to events other than the launch of Kinect. To do this, we replace $AfterKinect_t$ with a set of year-specific dummy variables. We plot the estimated value of the interaction between these year dummies with our treated dummy in Figures 1.a and 1.b. All estimates consider 2010 as our baseline year (Kinect was launched on Nov 4, 2010). Each point on the graphs represents the estimated difference between the number of publications of treated vs. control

[19] The publication boost is consistent with those of other work that examines shocks to the availability of research tools. For example, Furman and Stern (2011) find that making life science research materials available through public resource collections, biological resource centers (BRCs), induces between 50% and 125% increase in research referencing these materials. Similarly, Murray et al. (2011) find that open access to certain types of research mice yields a 22% to 43% boost in research citing the use of such mice.

[20] In addition to observing these differences across estimations, we conduct a CEM procedure on our main sample that includes both within-area and outside-area researchers. We add both interaction terms in the same estimations to better compare the magnitude of the effects and observe that indeed the impact on within-area researchers was smaller in magnitude than the impact on outside-area researchers (Appendix C, Table C1).

researchers, for each year, relative to the same difference in 2010. In each graph, the difference is small and close to zero before the launch of Kinect, as expected given our CEM procedure, and increases immediately thereafter, consistent with the conclusion that the availability of motion-sensing technology triggered an increase in researchers' productivity. In line with the discussion of coefficient magnitudes in Tables 3.a and 3.b, the effect is higher for outside-area researchers than for within-area researchers.

V.2. Did the Kinect shock induce research diversification?

Our first evidence comes from descriptive statistics on entry into motion-sensing research. Our estimation strategy precludes including authors who published only in the post-Kinect period, i.e., we do not observe a before period in order to identify an appropriate counterfactual. In Table 3, we report the number of new unique authors (based on IEEE database author identifiers) that appear in the dataset over the eight-year course of our sample. We distinguish between researchers entering our sample with at least one motion-sensing publication in their first year (Column 1) and researchers entering with other types of publications (Column 2). Column 3 reports the ratio of new entries in motion-sensing research. These data suggest diversification via entry into motion-sensing research. Prior to the Kinect launch, approximately 1% of entry into the IEEE dataset occurs via publications in motion-sensing, but this percentage increases by 31% immediately following the Kinect launch and continues to increase, up to 2.5 times in 2014, relative to the period before Kinect.

To investigate this question in a more formal way, we return to our main difference-in-differences estimation from equation (1). First, we focus on the impact on within-area researchers in Tables 4.a and 4.b. Table 4.a presents estimates using a dummy variable for treated individuals, while Table 4.b shows estimates using a continuous measure of involvement in motion-sensing before Kinect. Specifically, we replace $TreatedResearcher_i$ in equation (1) with a variable capturing the level of specialization in motion-sensing before Kinect, calculated as the sum of motion-sensing publications over the period before the launch of Kinect (2007- 2010) and divided by the total publication output of researcher *i* over the same period. In both tables, we present estimates using our three different diversification measures, the intensive and extensive HDP topic counts and the Euclidean diversification measure based on the IEEE taxonomy.

Table 4.a, Columns 1 and 2 show evidence of an 8% increase in the number of HDP topics covered yearly and a 5% increase in the number of unique HDP topics covered yearly, relative to the control group, consistent with hypothesis H1b. This amounts to a modest increase of 0.33 topics (and 0.16 respectively) out of approximately 50 topics covering the complete set of research in computer science and engineering. These findings are consistent with those obtained using the traditional diversification measure (which indicates a 7% increase in the Euclidean-based diversification) but have the added benefit of ease of interpretation of magnitude relative to the number of additional topics. Table 4.b. shows that the increase in diversification is more pronounced for those who had previously focused their research in motion-sensing, as opposed to researchers who also published in other domains. This suggests that rather than frustrating opportunities for within-area researchers, the automating technology facilitates exploration of ideas in their domains of expertise.

As before, we test the timing of these effects by replacing the post-shock dummy, $AfterKinect_t$, with a series of dummy variables reflecting each year of the analysis. We plot the estimated difference in yearly level of diversification between our treated and control researchers in Figure 2.a. All values are computed relative to 2010. In each case, we confirm the absence of pre-trends as constructed through our CEM procedure and find evidence of an increase in average researcher diversification that begins in the year after the Kinect launch and persists across time.

Next, we turn our attention to the effect for outside-area researchers and show results in Tables 5.a and 5.b. As before, Table 5.a presents estimates using our three different diversification measures, the intensive and extensive HDP topic counts and the Euclidean diversification measure based on the IEEE taxonomy, considering the motion-sensing publications produced by these researchers after the launch of Kinect. Table 5.b eliminates the motion-sensing publications and focuses on the impact on diversification for all other projects of these researchers.

Table 5.a, Columns 1 and 2 show evidence of a 65% increase in the number of HDP topics covered yearly and a 43% increase in the number of unique HDP topics covered yearly, relative to the control group, in line with hypothesis H2b. This amounts to an increase of 2.37 topics (and 1.18 respectively) out of approximately 50 topics. These findings are consistent with those obtained using the traditional

diversification measure, which indicates a 32% increase in Euclidean-based diversification, but with the added benefit of ease of interpretation of magnitude relative to number of additional topics. Additionally, we observe that the effect on outside-area researchers is larger than that on within-area researchers, in line with hypothesis H2b. As before, we include more formal estimates of the difference in magnitude in Appendix C, Table C2.

Table 5.b. demonstrates that this boost in diversification persists when we exclude motion-sensing publications. Specifically, the automation of motion-sensing technology facilitates a 36% increase in the number of HDP topics covered yearly by all other, non-motion-sensing publications produced by these researchers, relative to the matched control group. Similarly, Column 2 indicates a 21% increase in the number of unique HDP topics covered yearly by all other, non-motion-sensing publications produced by these researchers, relative to the control group. The change in diversification amounts to an increase of 1.31 topics (and 0.58 respectively).

As before, we test the timing of these effects by replacing $AfterKinect_t$ with a set of dummy variables for each sample year. We plot the estimated difference in yearly level of diversification between our treated and control researchers in Figures 3.a and 3.b. All values are computed relative to 2010. In each case, we confirm the absence of pre-trends as constructed through our CEM procedure and find evidence of an increase in average researcher diversification following the launch of Kinect that begins immediately after the launch and persists across time.

V.3. Did the Kinect shock induce changes in researches' trajectories?

To turn to the question of whether the Kinect shock induced changes in researchers' trajectory of inquiry, we apply the same estimation strategy as in the previous section but replace the diversification variables with three alternative measures designed to reflect differences in the distance in knowledge space between researchers' portfolio of projects at times *t-1* and *t*. In Tables 6.a and 6.b, we present the results of changes in trajectory for within-area researchers. Table 6.a presents estimates using a dummy variable for treated individuals, while Table 6.b shows estimates using a continuous measure of involvement in motion-sensing before Kinect. Column 1 of Table 6.a estimates that within-area researchers experience a 7% increase in new topics, equivalent to a modest shift of 0.24 topics. Column 2 of the same table estimates a 5% increase in the number of new coauthors (1 in 10 new coauthors), while Column 3 estimates an 8% increase in the number of new publication outlets (1 in 15 new publication outlets). The advantage of the HDP measure is that it allows us to estimate the number of additional topics the researcher undertakes; an increase in the number of authors and publication outlets does not necessarily imply an increase in the number of topics since researchers can work with new coauthors and publish in different journals without changing their topics of interest, or, alternatively, by abandoning old topics and undertaking new ones. The effects are most pronounced for researchers who have a greater degree of involvement in motion-sensing research before Kinect (Table 6.b.). We test the timing of these effects and display the yearly estimated coefficients in Figure 2.b. We interpret the results as implying that Kinect induced within-area researchers to expand their research into areas in which they had not worked before, in line with our hypothesis H1b.

Tables 7.a and 7.b suggest that the same holds – though to a larger degree[21] – for outside-area researchers, in line with hypothesis H2b. As in prior analyses, Table 7.a presents estimates of the three trajectory measures, including motion-sensing publications produced by these researchers after the launch of Kinect. Table 7.b eliminates the motion-sensing publications and focuses on the impact on trajectory for all other projects of these researchers. Column 1 of Table 8.a indicates a 66% increase in new topics, equivalent to a shift in trajectory of 2 new topics. Column 2 of the same table estimates a 59% increase in the number of new coauthors (1.25 new coauthors), while Column 3 estimates a 57% increase in the number of new publication outlets (1 in 2 additional publication outlets). The effects persist when eliminating the set of motion-sensing publications from the portfolio of these researchers, suggesting that the effect is not mechanically driven by incorporating motion-sensing into the portfolio of projects (Table 7.b), in line with H2b. Specifically, outside-area researchers increase the number of yearly new topics by 22% (0.45 new topics out of approximately 50), the number of new-coauthors by 30% (more than 1 in 2 new coauthors) and the number of new publication outlets by 61% (1 in 2 new outlets). As before, we test the timing of these effects and display the yearly estimated coefficients in Figures 4.a and 4.b.

[21] As before, we include more formal estimates of the difference in magnitude in Appendix C, Table C3.

V.4. Complementary descriptive analyses based on Scopus data

The results of our regression analyses suggest that the availability of this motion-sensing technology induces greater research diversity and a shift in the trajectory of researchers' project portfolios. Ideally, we would also like to gain insight into the type of knowledge these individuals create as part of the identified effects. However, the difficulty of capturing changes in the type of knowledge produced also poses challenges to our ability to speak directly to the measured increase in diversification and shift in trajectory. To shed some light on this, in addition to the examples mentioned earlier, we provide some descriptive examples we collect through a different bibliographical database, Scopus, and its analyze function which allows user to select a group of papers and analyze their attributes in terms of e.g., spread across domains of knowledge, counts of authors affiliated with various institutions, spread across different publication outlets, etc. We observe two broad avenues through which researchers enhance the trajectory of their inquiry following Kinect: (1) an increase in the number of areas where they publish and (2) an increase in the percentage of their publications across knowledge areas in their portfolio. We include these data and additional information in Appendix D.

Additionally, we utilize the Scopus analyze feature to provide a glimpse into the mechanism through which the benefits of the automating technology manifest, by attempting to identify the type of institutions that benefit most from the availability of Kinect as automating motion-sensing research technology.[22] On the one hand, the benefit of Kinect as a technology that automates research tasks might manifest more for research institutions that are or aim to be leaders in research. The assumption is that these institutions are time constrained, rather than financially constrained, and the automation would help propel their productivity forward. On the other hand, the benefit of Kinect as technology that reduces the costs of performing certain research tasks might manifest more for institutions that are financially constrained. It follows that a research technology that is automating by significantly reducing the cost of performing certain research tasks should benefit both types of institutions. It is important to remember that we cannot separate these effects from those resulting from the low monetary cost of Kinect. In addition to reducing the cost of executing certain research tasks via automation, Kinect's low monetary cost (relative to the

[22] Data limitations prevent us from providing comprehensive evidence at the individual researcher level.

previous generation of motion-sensing technology) may have enabled experimentation that might not have been considered if Kinect was priced in the same range as the previous generation technology. It may be, therefore, that changes in the monetary cost of motion-sensing research technology drive some of the effects we observe. However, if these effects were dominant, then we would expect to see financially constrained institutions disproportionally benefiting from the Kinect phenomenon. In fact, when collecting author affiliations for the top 25% and top 10% most cited motion-sensing publications during the period before and after Kinect, we observe that the set of institutions experiencing the highest increase in the number of top cited publications is comprised of both highly ranked universities and private institutions with little financial constraints and lower ranked institutions with presumably higher financial constraints. This finding is aligned with the idea that the automating research technology reduces the cost of performing certain research tasks by substituting for human capital and thus benefiting both leading institutions racing to maintain their frontier position as well as more financially constraining institutions. We include these data and additional information in Appendix E.

VI. Discussion and Conclusion

Automating technologies are transforming the nature of work and competition across industries. Scholarship in management and economics has investigated the impact of such technologies on the manufacturing and service sectors and has documented the extent to which and the conditions under which these technologies substitute for or complement human labor. Although research on innovation emphasizes the importance of research tools, less work investigates the role of such automating technologies in the production of knowledge. In this paper, we contribute to addressing this gap by examining the impact of an automating technology on the rate and type of knowledge production.

Ideas production differs in key ways from the production of goods and services. We believe that a core contribution of the paper involves training a lens on the way in which automating technologies affect the organization of knowledge production. In contrast to some traditional manufacturing positions (e.g., assembly line work) and service positions (e.g., fast food preparation), knowledge work is characterized by a particularly wide range of tasks (e.g., problem selection, grant writing, data analysis, paper composition, and seminar presentation) and a substantial amount of autonomy for knowledge workers to select their tasks. This task variety and authority over their own work may enable knowledge workers to adapt to the introduction of automating technologies in ways that vary from those of workers – and even managers – in manufacturing and service sectors. Our results suggest that this is the case. Researchers with prior involvement in motion-sensing are not negatively impacted by the availability of motion-sensing technology that automates core research tasks. Moreover, such researchers, on average, accelerate the pace of their work in response to automation, and spread their knowledge wider.

A second area to which we believe this paper contributes is to the study of research technology as an input into knowledge production. In particular, we document that the availability of this motion-sensing research technology enables gains in production and movement in ideas space. Further, we demonstrate that the ability to leverage research technology varies across researcher types. While the Kinect shock supports some benefits among within-area researchers, its greatest impact is among researchers who have not previously engaged in motion-sensing research. Acemoglu (2002, 2012) and Bryan and Lemus (2017) worry that economic incentives lead to an underprovision of research diversity that is especially pronounced among for-profit firms and that can be alleviated by increasing research diversity among academics. This paper suggests that automating research technology can lessen this problem by enabling wider exploration of research ideas.

A third contribution we seek to make in this paper involves the adaptation of machine learning tools to develop generalizable measures of type of knowledge production. In our main specifications, we employ Hierarchical Dirichlet Process (HDP), which we train on paper abstracts to parse fields of research into multiple categories. The approach has several advantages over bibliometric measures for capturing the diversity and trajectory of research output. Furthermore, unlike LDA, one of the most-often used topic modeling tools, HDP identifies the optimal number of topics within a body of text. While the technique cannot automatically track the evolution of topics over time, we address this limitation by augmenting it with a clustering technique, Term Frequency – Inverse Document Frequency (TFIDF) cosine similarity, to calculate a cosine vector similarity between the yearly topics generated by the HDP algorithm. We believe that this approach adds to similar early efforts of employing machine learning techniques in developing measures of type of knowledge output (Kaplan and Vakili, 2015), and, importantly, represents an advance in measuring *changes* in the type of knowledge output.

Lastly, a fourth contribution we hope to make is to expand the discussion of automation in literature in strategic management and organization studies.[23] While a substantial and insightful literature on the impact of digital technologies exists, much of this appears in journals focused on information systems and economics and we hope that this work contributes to expanding attention to these topics in core strategy and management research.

One of the advantages of our context is that the peculiar surprise associated with the Kinect technology supports a closer to causal analysis of its impact on the community of researchers in computer science and electrical engineering research. There are numerous analogs outside these areas in which specific research tasks have been automated, including the automation of statistical analysis via tools such as SAS, Stata, and R and the automation associated with combinatorial chemistry techniques. By focusing on a relatively narrow research area and an unanticipated increase in the availability of a key automating technology in that area, we are able to get closer to obtaining causal identification, albeit at the cost of some external validity. Our approach is also limited in disentangling the effects of automation from that of publicity of the technology; the launch of Kinect and subsequent hype were highly public. Moreover, Kinect is an automating technology that reduces the cost of executing the tasks it automates with minimal adjustments or co-invention costs. This differs from other technologies such as mainframes, servers or cloud computing (Jin and McElheran, 2018) where the automation effects might manifest in more complex ways. In addition, Kinect was offered a lower monetary cost than previous generations of motion-sensing technology. It is possible that the magnitude of the effects we observe are in part explained by these circumstances.

Overall, our analysis underscores the important role of IT-based research technology for innovation. Understanding this relationship should be of central interest for research-oriented organizations focused on being ahead of competitors in knowledge production. Our study suggests that both access to research technology and the type of knowledge workers employed by these organizations matter. In addition to productivity benefits, automating research technologies may induce benefits of greater research diversity among innovation-focused firms. Whereas absorptive capacity has focused on the complementarity between internal R&D efforts and external knowledge, this paper highlights the fact that expertise can be

[23] We are grateful to an anonymous reviewer for proposing this point.

embedded in a research technology and suggests that externally-developed research technologies may have effects on productivity and idea mobility within firms. Furthermore, the magnitude of impact of research technology in knowledge production draws attention to the role of market power in technology development and retailing, as well as the complex implications of technology development and pricing strategies. For example, offers of discounted technologies could lead to indirect returns in the form of accelerated rates of innovations. Thus, while our analysis has been conducted at the level of individual researchers, we hope the paper invites continued work on the impact of automating technologies on innovating firms and sectors and lights a path towards such investigations.

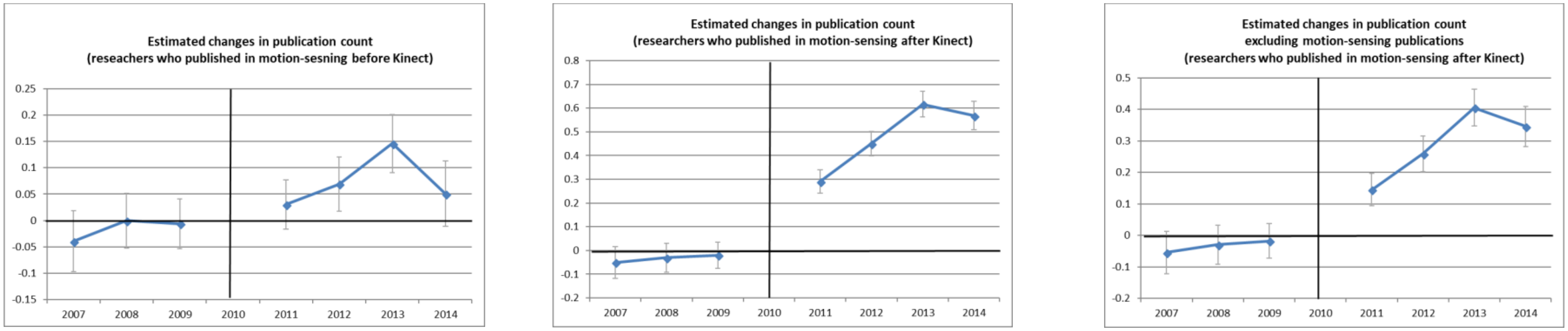
REFERENCES

1. Acemoglu, D. (2002) “Directed Technical Change,” *Review of Economic Studies*, 69(4), 781-809.
2. Acemoglu, D. (2012) “Diversity and Technological Progress” in *The Rate and Direction of Inventive Activity Revisited*, Scott Stern and Josh Lerner, ed. Chicago: IL: University of Chicago Press, 319-356.
3. Acemoglu, D. and Autor, D. (2011) "Skills, Tasks and Technologies: Implications for Employment and Earnings," in David Card and Orley Ashenfelter (ed.), *Handbook of Labor Economics*, Volume 4, 1043-1171.
4. Acemoglu, D. and Restrepo, P. (2018) ”Modeling Automation,” *American Economic Review Papers and Proceedings*, 108, 48-53.
5. Adafruit (2011) “Johhny Lee was behind the Kinect Bounty Contest -with a little help from us :),” Adafruit blog, https://blog.adafruit.com/2011/02/21/johhny-lee-was-behind-the-kinect-bounty-contest-with-a-little-help-from-us/, posted February 21, 2011 at 7:41 pm, accessed, July 24, 2017.
6. Aghion, P., Dewatripont, M. Stein, C. J. (2008) “Academic freedom, private-sector focus, and the process of innovation,” *Rand Journal of Economics*, 39(3), 617–635.
7. Agrawal, A., Goldfarb. A., Teodoridis, F. (2016). “Understanding the Changing Structure of Scientific Inquiry,” *American Economic Journal: Applied Economics* 8(1): 100-128.
8. Agrawal, A., Gans, J., Goldfarb, A. (2018) *The Economics of Artificial Intelligence: An Agenda*, NBER Conference, Toronto, September 13-14, 2017.
9. Argyres, N. S. (1999) “The Impact of Information Technology on Coordination: Evidence from the B-2 ‘Stealth’ Bomber,” *Organization Science* 10(2): 162-180.
10. Athey, S. and Imbens, G. (2015) “Lectures on Machine Learning,” NBER Summer Institute.
11. Autor, D. H., Katz, L. F., Krueger, A. B. (1998), “Computing inequality: have computers changed the labor market?. *The Quarterly Journal of Economics*, *113*(4), 1169-1213.
12. Azoulay, P., Furman, J. L., Krieger, J. L., Murray, F. (2015) “Retractions”, *Review of Economics and Statistics*, 97(5): 1118-1136.
13. Bessen, J. 2015. “How Computer Automation Affects Occupations: Technology, Jobs, and Skills”, Boston University Law & Economics working paper # 15-49.
14. Bilton, N. (2011) “Microsoft Sells 10 Million Kinects, Breaking Record,” New York Times Blogs, https://bits.blogs. nytimes.com/2011/03/09/microsoft-sells-10-million-kinects-breaking-record/, posted March 9, 2011 6:08pm; accessed 24 July 2017.
15. Borjas, G., Doran, K. (2012) “The Collapse of the Soviet Union and the Productivity of American Mathematicians,” *Quarterly Journal of Economics*, 127(3), 1143-1203.
16. Borjas, G., Doran, K. (2015) “Cognitive Mobility: Native Responses to Supply Shocks in the Space of Ideas,” *Journal of Labor Economics*, 33(S1), S109-S145.
17. Boudreau, K. J., Guinan, E. C., Lakhani, K. R., Riedl, C. (2017) “Looking Across and Looking Beyond the Knowledge Frontier: Intellectual Distance, Novelty, and Resource Allocation in Science,” *Management Science*, 62(10), 2765-2783.
18. Bloom, N., Sadun, R., Van Reenen, J. (2012) “Americans do IT better: US multinationals and the productivity miracle,” *American Economic Review*, 102(1), 167-201.

19. Bresnahan, T. F., Brynjolfsson, E., Hitt, L. M. (2002). Information technology, workplace organization, and the demand for skilled labor: Firm-level evidence. *The Quarterly Journal of Economics*, *117*(1), 339-37
20. Bryan, K. and Lemus, J. (2017) "The Direction of Innovation", working paper.
21. Brynjolfsson, E. and Hitt, L. (2003) "Computing Productivity: Firm-Level Evidence," *Review of Economics and Statistics,* 85(4), 793-808.
22. Brynjolfsson, E, Mitchell, T., Rock, D. (2018) "Why can Machines Learn and What does it Mean for Occupations and the Economy," *American Economic Review*, 108, 43-47.
23. Buntine, W. and Jakulin, A. (2004) "Applying discrete PCA in data analysis*,"* In Proceedings of the 20th conference on Uncertainty in artificial intelligence, 59-66.
24. Carmody, T., (2010a) "Adafruit Offers $1000 Bounty for Open-Source Kinect Drivers," *Wired Magazine*, https://www.wired.com/2010/11/adafruit-offers-1000-bounty-for-open-source-kinect-drivers/, posted November 04, 2010, 1.28pm; accessed 24 July 2017.
25. Carmody, T., (2010b) "How Microsoft Learned to Stop Worrying and Love Open Kinect," Wired Magazine, https://www.wired.com/2010/11/open-kinect/, posted November 22, 2010, 2.25pm; accessed July 24,2017.
26. Choudhury, P., Starr, E., Agarwal, R. (2018) "Machine Learning and Human Capital: Experimental Evidence on Productivity Complementarities," working paper
27. Cockburn, I., Henderson, R., Stern, S. (2017) "The Impact of Artificial Intelligence on Innovation: An Exploratory Analysis," forthcoming in "The Economics of Artificial Intelligence: An Agenda", University of Chicago Press.
28. Cohen, M. W. and Klepper, S. (1992) "The Tradeoff between Firm Size and Diversity in the Pursuit of Technological Progress," *Small Business Economics*, 4(1), 1-14.
29. Cohen, M. W. and Klepper, S. (1996) "A Reprise of Size and R&D," *The Economic Journal*, 106(437), 925-951.
30. Cohen, W M, Sauermann, H., Stephan, P. E. (2018) "Academics' motives, opportunity costs, and commercial activities across fields," NBER Working Paper #24769.
31. Deming, D. J. (2017) "The Growing Importance of Social Skills in the Labor Market," Quarterly Journal of Economics, 132 (4), 1593-1640.
32. Dean, J. W., Se Joon Yoon, and Gerald I. Susman (1992) "Advanced Manufacturing Technology and Organization Structure: Empowerment or Subordination?" *Organization Science*, 3(2), 203-229.
33. Ding, W. W., Levin, S. G., Stephan, P. E., Winkler, A. E. (2010) "The Impact of Information Technology on Scientists' Productivity, Quality and Collaboration Patterns," *Management Science* 56(9): 1439 - 1461.
34. Doran, K. (2017) "The Direction of Science: Natural Experiments," presentation at the June 2017 DRUID Conference, New York, NY.
35. Dos Santos, B. and Peffers, K. (1995) "Rewards to Investors in Innovative Information Technology Applications: First Movers and Early Followers in ATMs," *Organization Science* 6(3): 241-259.
36. Felten, E., Raj, M., Seamans, R. (2018) "A Method to Link Advances in Artificial Intelligence to Occupational Abilities", *AEA Papers & Proceedings*, 108: 54-57.
37. Forman, C. and McElheran, K. (2012) "Information Technology and the Boundary of the Firm: Evidence from Plant-Level Data," *Harvard Business School Working Paper.*
38. Furman, J. L., and Scott, S. (2011) "Climbing Atop the Shoulders of Giants: The Impact of Institutions on Cumulative Knowledge Production," *American Economic Review*, 101(5): 1933-1963.
39. Furman, J. L., Murray, F., and Stern, S. (2012) "Growing Stem Cells: The Impact of Federal Funding Policy on the U.S. Scientific Frontier," *Journal of Policy Analysis & Management*, 31(3), 661-705.
40. Giles, Jim (2010) "Inside the race to hack Microsoft's Kinect," *New Scientist*, 208(2789), 1.
41. Griliches, Z. (1990) "Patent Statistics as Economic Indicators: A Survey," *Journal of Economic Literature*, 28(4): 1661–707.
42. Grupp, H. (1990) "The Concept of Entropy in Scientometrics and Innovation Research," *Scientometrics*, 18(3-4), 219-239.
43. Han, J., Shao, L., Xu, D. and Shotton, J. (2013) "Enhanced Computer Vision with Microsoft Kinect Sensor: A Review," *IEEE Transactions on Cybernetics*, 43(5), 1318-1334.
44. Hofmann, T. (1999) "Probabilistic latent semantic analysis," in Proceedings of the Fifteenth conference on Uncertainty in artificial intelligence, 289-296.
45. Iacus, S. M. and Porro, K. (2011) "Multivariate matching methods that are monotonic imbalance bounding", *Journal of American Statistics Association*, 106(493): 345-361.
46. Iacus, S.M. and Porro, K. (2012) "Causal inference without balance checking: Coarsened exact matching", *Political Analysis* 20(1):1-24.

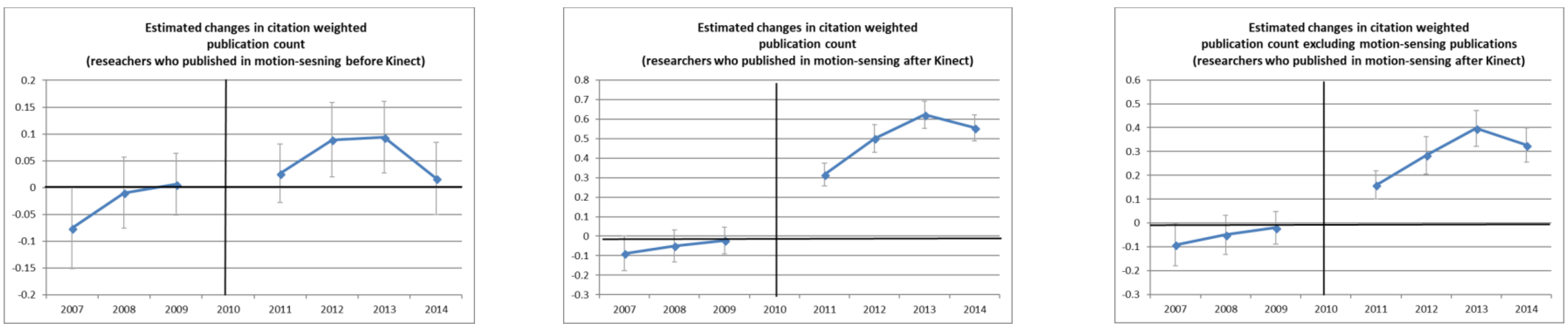
47. Jin, W., and McElheran, K. (2018) "Economies Before Scale: Survival and Performance of Young Plants in the Age of Cloud Computing", Rotman School of Management working paper no. 3112901.
48. Jones, B. (2009) "The Burden of Knowledge and the Death of the Renaissance Man: Is Innovation Getting Harder?" *Review of Economic Studies*, 76(1): 253-281.
49. Jones, B. (2010) "As Science Evolves, How Can Science Policy?" *NBER Innovation Policy and the Economy*, 11: 103-131.
50. Kaplan, S., and Vakili, K. (2015) "The double-edged sword of recombination in breakthrough innovation", *Strategic Management Journal*, 36(10): 1435-1457.
51. Karabarbounis, L. and Neiman, B.. 2014. "The Global Decline of the Labor Share." *Quarterly Journal of Economics*, 129(1), 61-103.
52. Krieger, J., Li, D., Papanikolou, D. (2018) "Developing Novel Drugs," *Harvard Business School Working Paper, 18-056.*
53. Lee, J. C. (2011) "Windows Drivers for Kinect, Finally!", http://procrastineering.blogspot.com/2011/02/windows-drivers-for-kinect.html, posted February, 21, 2011, accessed July 24, 2017.
54. Merton, R. K. (1938) "Science and the Social Order," *Philosophy of Science*, *5(3), 321-337.*
55. Mokyr, J. (2002) "The Gifts of Athena: Historical Origins of the Knowledge Economy," Princeton University Press: Princeton, NJ.
56. Mokyr, J., Vickers C., and Ziebarth N. L. (2015) "The History of Technological Anxiety and the Future of Economic Growth: Is This Time Different?" *Journal of Economic Perspectives*, 29(3), 31-50.
57. Murray, F., Aghion, P., Dewatripont, M., Kolev, J., & Stern, S. (2016) "Of Mice and Academics: Examining the Effect of Openness on Innovation," *American Economic Journal: Economic Policy*, 8(1), 212-52.
58. Murray, F. and O'Mahony, S. (2007) "Exploring the Foundations of Cumulative Innovation: Implications for Organization Science", *Organization Science*, 6(18): 1006-1021.
59. Myers, K. (2018) "The Elasticity of Science," working paper.
60. Nelson, A. J. (2016) "How to Share 'A Really Good Secret': Managing Sharing/Secrecy Tensions Around Scientific Knowledge Disclosure," *Organization Science* 27(2):265-285.
61. Nordhaus, W. D. 2007. "Two centuries of productivity growth in computing." *The Journal of Economic History*, 67 (1): 128-159.
62. Pinsonneault, A. and Kenneth L. K. (2002) "Exploring the Role of Information Technology in Organizational Downsizing: A Tale of Two American Cities," *Organization Science* 13(2):191-208.
63. Rafibakhsh, Nim, Jie Gong, Mohsin K. Siddiqui, Chris Gordon, and H. Felix Lee (2012) "Analysis of Xbox Kinect sensor data for use on construction sites: Depth accuracy and sensor interference assessment," Construction Research Congress 2012: Construction Challenges in a Flat World, 848-857.
64. Rafols, I., and Meyer, M. (2010) "Diversity and network coherence as indicators of interdisciplinarity: Case studies in bionanoscience," *Scientometrics*, 82(2), 263–287.
65. Richards-Rissetto, H., Remondino, F., Agugiaro, G., von Schwerin, J., Robertsson, J. and Girardi, G., (2012) "Kinect and 3D GIS in archaeology," IEEE, 18th International Conference on Virtual Systems and Multimedia, 331-337.
66. Smith, A. 1776. "An Inquiry into the Nature and Causes of the Wealth of Nations", W. Strahan and T. Cadell, London.
67. Teodoridis, F. (2018) "Understanding Team Knowledge Production: The Interrelated Roles of Technology and Expertise," *Management Science*, 64 (8): 3469-3970.
68. Teh, Y. W., Jordan, M. I., Beal, M. J., Blei, D. M. 2006. "Hierarchical Dirichlet Processes", *Journal of the American Statistical Association* 101:1566–1581.
69. Terdiman, D. (2010) "Bounty Offered for Open-Source Kinect Driver," CNET News, https://www.cnet.com/news/bounty-offered-for-open-source-kinect-driver/, posted November 4, 2010, 12.50pm, updated 4.03pm and 5.07pm; accessed 24 July 2017.
70. Uzzi, B., Mukherjee, S., Stringer, M., & Jones, B. (2013) "Atypical combinations and scientific impact" *Science*, 342(6157), 468-472.
71. Volkoff, O., Strong, D., Elmes, M. (2007) "Technological Embeddedness and Organizational Change," *Organization Science* 18(5):832-848.
72. Wortham, J. (2010) "Off the Shelf and into the garage," New York Times, November 22, page B1.
73. Zhang, Z. (2012) "Microsoft Kinect Sensor and Its Effect," *IEEE Multimedia*, 19(2), 4–10.
74. Thompson, N. and Zyontz. S. (2017) "Who tries (and who succeeds) in staying at the forefront of science: Evidence from the DNA-Editing Technology CRISPR, MIT-Sloan working paper.

Figure 1.a: Estimated changes in yearly publication counts



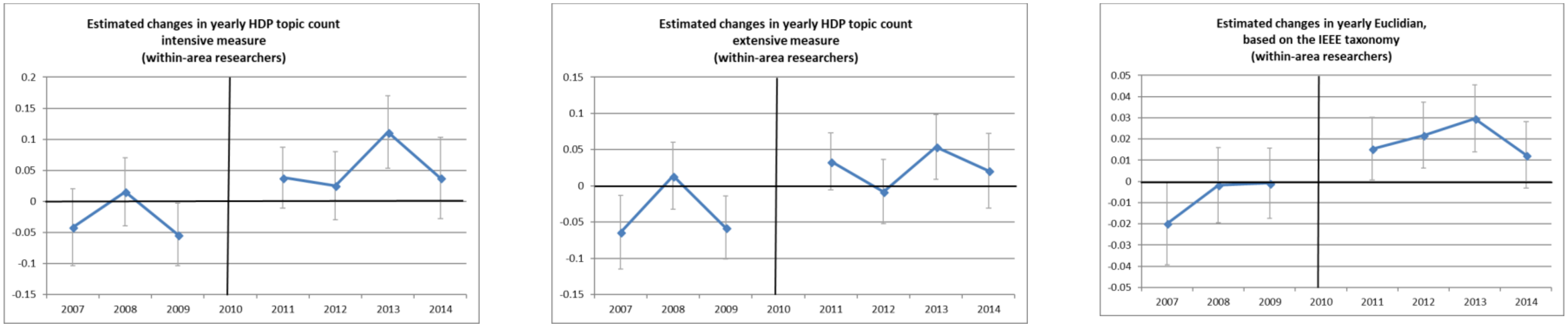
Notes: We base this figure on our 2007-2014 dataset. Each point on the graph represents the coefficient value on the covariate *TreatedResearcher x Year* and thus describes the relative difference in publication counts between treated and control authors in that year. The bars surrounding each point represent the 95 percent confidence interval. All values are relative to the base year of 2010.

Figure 1.b: Estimated changes in yearly publication counts weighted by citations



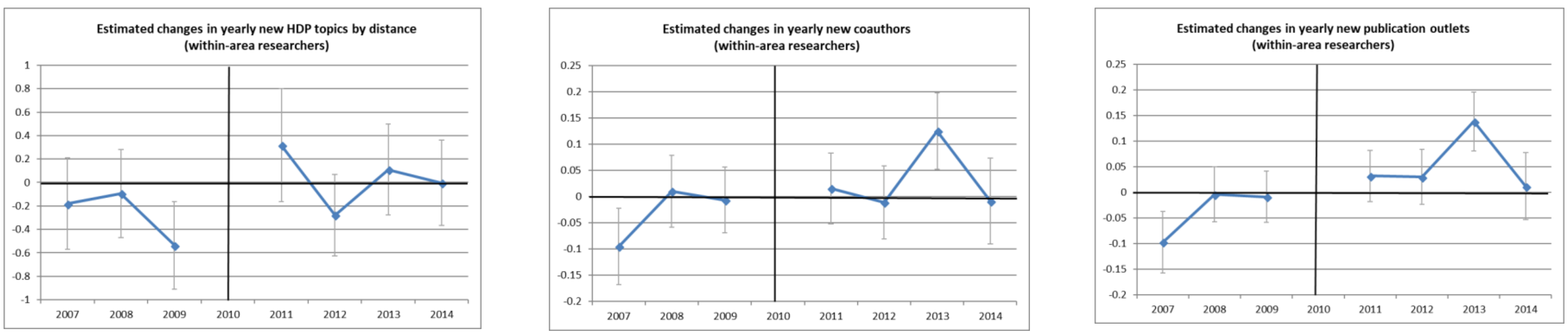
Notes: We base this figure on our 2007-2014 dataset. Each point on the graph represents the coefficient value on the covariate *TreatedResearcher x Year* and thus describes the relative difference in citation weighted publication counts between treated and control authors in that year. The bars surrounding each point represent the 95 percent confidence interval. All values are relative to the base year of 2010.

Figure 2.a: Estimated changes in the level of diversification of within-area researchers



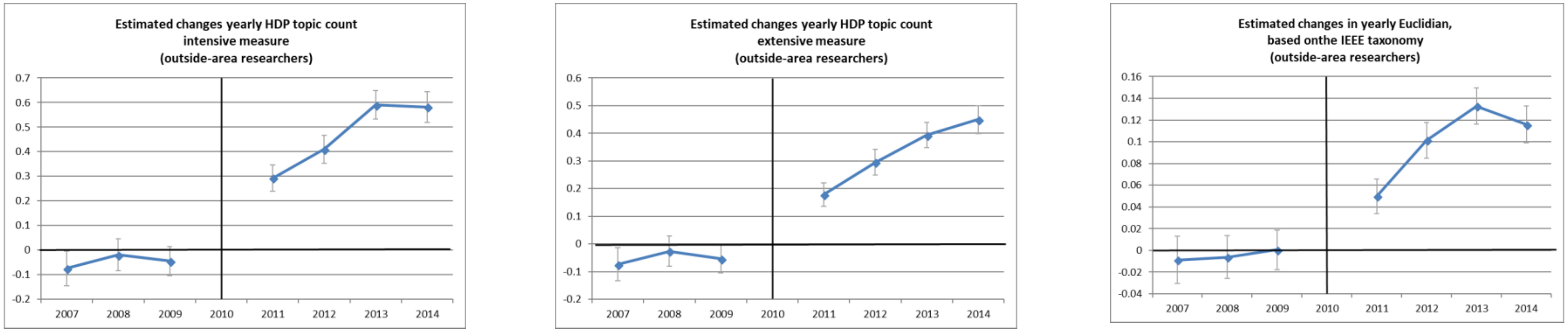
Notes: We base this figure on our 2007-2014 dataset. Each point on the graph represents the coefficient value on the covariate *TreatedResearcher x Year* and thus describes the relative difference in diversification between treated and control authors in that year. The bars surrounding each point represent the 95 percent confidence interval. All values are relative to the base year of 2010.

Figure 2.b: Estimated changes in the type of knowledge output of within-area researchers



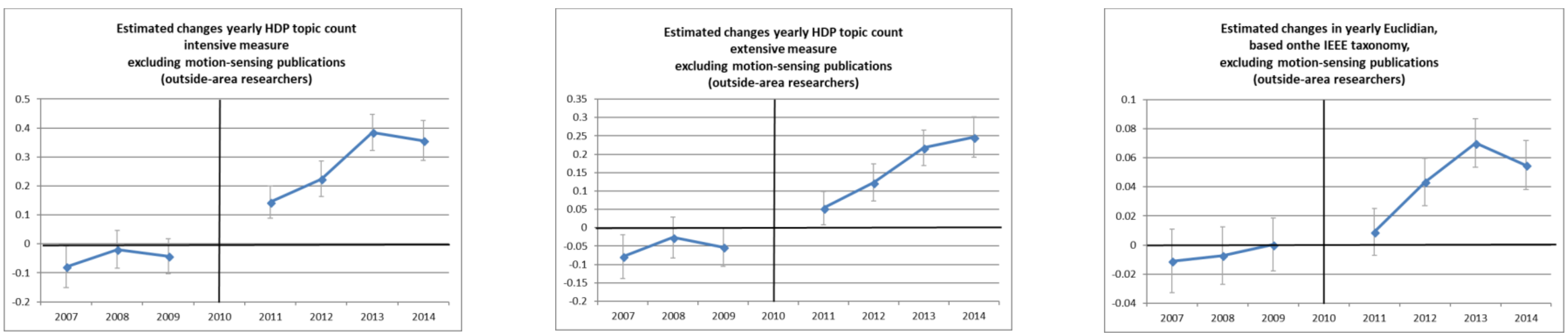
Notes: We base this figure on our 2007-2014 dataset. Each point on the graph represents the coefficient value on the covariate *TreatedResearcher x Year* and thus describes the relative difference in our direction measures between treated and control authors in that year. The bars surrounding each point represent the 95 percent confidence interval. All values are relative to the base year of 2010.

Figure 3.a: Estimated changes in the level of diversification of outside-area researchers



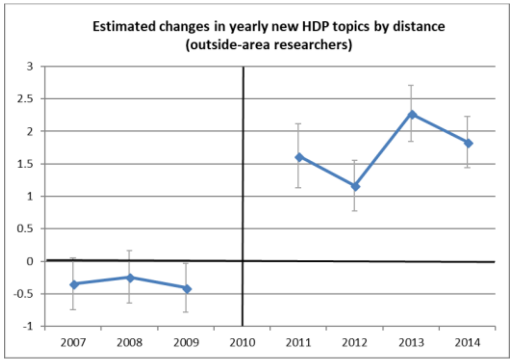
Notes: We base this figure on our 2007-2014 dataset. Each point on the graph represents the coefficient value on the covariate *TreatedResearcher x Year* and thus describes the relative difference in diversification between treated and control authors in that year. The bars surrounding each point represent the 95 percent confidence interval. All values are relative to the base year of 2010.

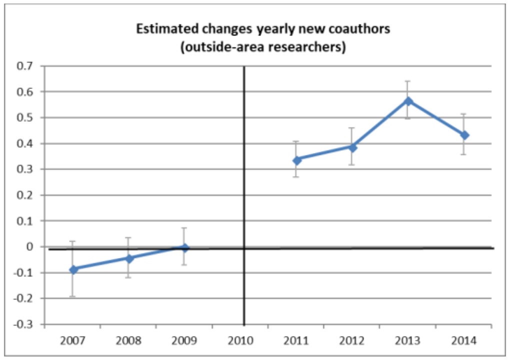
Figure 3.b: Estimated changes in the level of diversification of outside-area researchers, when excluding their motion-sensing publications

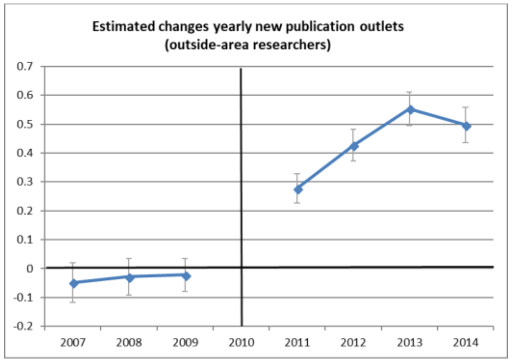


Notes: We base this figure on our 2007-2014 dataset. Each point on the graph represents the coefficient value on the covariate *TreatedResearcher x Year* and thus describes the relative difference in diversification between treated and control authors in that year. The bars surrounding each point represent the 95 percent confidence interval. All values are relative to the base year of 2010.

Figure 4.a: Estimated changes in the type of knowledge output of outside-area researchers

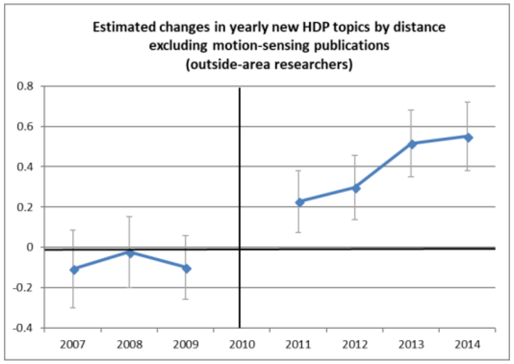


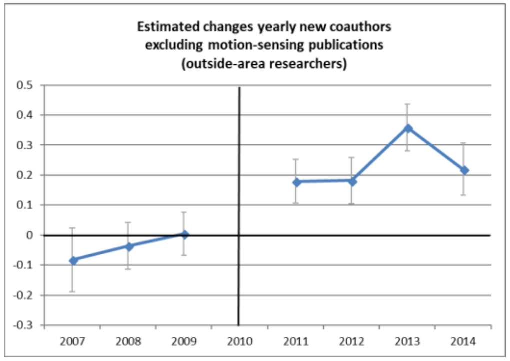


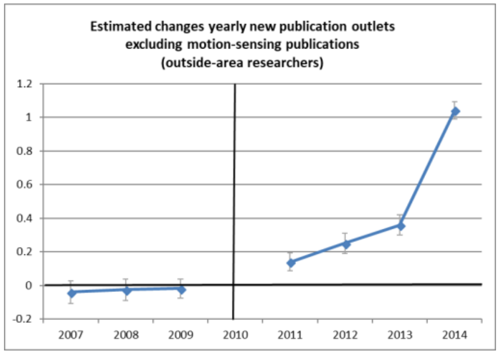


Notes: We base this figure on our 2007-2014 dataset. Each point on the graph represents the coefficient value on the covariate *TreatedResearcher x Year* and thus describes the relative difference in our direction measures between treated and control authors in that year. The bars surrounding each point represent the 95 percent confidence interval. All values are relative to the base year of 2010.

Figure 4.b: Estimated changes in the type of knowledge output of outside-area researchers, when excluding their motion-sensing publications







Notes: We base this figure on our 2007-2014 dataset. Each point on the graph represents the coefficient value on the covariate *TreatedResearcher x Year* and thus describes the relative difference in our direction measures between treated and control authors in that year. The bars surrounding each point represent the 95 percent confidence interval. All values are relative to the base year of 2010.

Table 1.a: Changes in researchers' productivity for our matched samples – Descriptive statistics

Descriptive statistics – Rate (Mean (St. Dev))					
Treated as:		Publication count		Publication count weighted by citations	
		Treated	Controls	Treated	Controls
Within-area researchers	Before	1.182 (1.241)	1.214 (1.295)	2.291 (3.346)	2.256 (3.327)
	After	1.249 (1.669)	1.153 (1.600)	1.767 (3.034)	1.585 (2.941)
Observations		13,168	111,936	13,168	111,936
Outside-area researchers	Before	1.068 (1.199)	1.086 (1.240)	1.900 (3.048)	1.891 (3.058)
	After	1.676 (1.994)	1.022 (1.469)	2.408 (3.740)	1.404 (2.672)
Observations	Before	11,936	79,008	11,936	79,008

Table 1.b: Changes in diversification for our matched samples – Descriptive statistics

Descriptive statistics – Diversification (Mean (St. Dev))							
Treated as:		HDP topic count - intensive		HDP topic count - extensive		Euclidean diversification using IEEE taxonomy	
		Treated	Controls	Treated	Controls	Treated	Controls
Within-area researchers	Before	4.106 (4.634)	4.098 (4.654)	3.091 (2.951)	3.017 (2.890)	0.360 (0.296)	0.358 (0.294)
	After	4.899 (6.858)	4.455 (6.489)	3.189 (3.356)	2.919 (3.216)	0.313 (0.288)	0.283 (0.285)
Observations		13,168	111,936	13,168	111,936	13,168	111,936
Outside-area researchers	Before	3.642 (4.432)	3.662 (4.537)	2.751 (2.872)	2.709 (2.867)	0.325 (0.296)	0.323 (0.297)
	After	6.530 (8.198)	3.934 (5.991)	3.913 (3.583)	2.665 (3.102)	0.372 (0.286)	0.262 (0.281)
Observations		11,936	79,008	11,936	79,008	11,936	79,008

Table 1.c: Changes in the type of knowledge output of research for our matched samples – Descriptive statistics

Descriptive statistics – Direction (Mean (St. Dev))							
Treated as:		Yearly new HDP topics by distance		Yearly new coauthors		Year new publication outlets	
		Treated	Controls	Treated	Controls	Treated	Controls
Within-area researchers	Before	3.466 (5.599)	3.114 (5.163)	2.242 (3.086)	2.258 (3.074)	0.839 (0.964)	0.883 (1.014)
	After	4.415 (7.133)	3.970 (6.525)	2.565 (4.555)	2.359 (4.387)	0.789 (1.111)	0.739 (1.055)
Observations		13,168	111,936	13,168	111,936	13,168	111,936
Outside-area researchers	Before	2.957 (5.156)	2.707 (5.085)	2.061 (3.428)	2.048 (3.367)	0.783 (0.951)	0.806 (0.973)
	After	5.613 (8.208)	3.554 (6.105)	3.512 (5.574)	2.126 (4.195)	1.063 (1.298)	0.674 (0.991)
Observations		11,936	79,008	11,936	79,008	11,936	79,008

Table 2.a: Estimated changes in publication count

Controls determined through Coarsened Exact Matching (CEM)			
	Count of publications		
	Within-area researchers	Outside-area researchers	Outside-area researchers (exclude motion-sensing papers)
Treated x AfterKinect	0.083***	0.497***	0.305***
	(0.017)	(0.017)	(0.019)
Quadratic age	Yes	Yes	Yes
Individual and year FE	Yes	Yes	Yes
LL	-270,358.60	-187,776.56	-186,346.70
Observations	231,666	166,212	166,212

Notes: The data is a panel at the author level based on publication data between 2007 and 2014. All models are Poisson with robust standard errors clustered at the individual level. *significant at 10%, **significant at 5%, ***significant at 1%

Table 2.b: Estimated changes in citation weighted publication count

Controls determined through Coarsened Exact Matching (CEM)			
	Count of publications weighted by citations		
	Within-area researchers	Outside-area researchers	Outside-area researchers (exclude motion-sensing papers)
Treated x AfterKinect	0.083***	0.537***	0.330***
	(0.023)	(0.024)	(0.027)
Quadratic age	Yes	Yes	Yes
Individual and year FE	Yes	Yes	Yes
LL	-424,738.03	-285,514.28	-283,379.44
Observations	231,666	166,212	166,212

Notes: The data is a panel at the author level based on publication data between 2007 and 2014. All models are Poisson with robust standard errors clustered at the individual level. *significant at 10%, **significant at 5%, ***significant at 1%

Table 3: New researcher entries in electrical engineering, computer science, and electronics, per year, as observed through publications logged in the IEEE Xplore database (2007-2014)

Number of authors entering IEEE Xplore academic publication			
	Number of authors entering with at least one motion-sensing publication	Number of authors entering with publications in other areas	Percentage of motion-sensing entry
2007	1,103	122,322	0.90%
2008	1,307	124,670	1.04%
2009	1,188	120,554	0.99%
2010	1,502	152,237	0.99%
2011	1,794	137,233	1.31%
2012	2,449	135,514	1.81%
2013	2,805	116,184	2.41%
2014	2,985	119,265	2.50%

Table 4.a: Estimated changes in the level of diversification of within-area researchers

Treated as within-area researchers. Controls determined through Coarsened Exact Matching (CEM)			
Diversification	HDP topic count (intensive)	HDP topic count (extensive)	Euclidean diversification using IEEE taxonomy
Treated x AfterKinect	0.075***	0.052***	0.025***
	(0.018)	(0.013)	(0.004)
Quadratic age	Yes	Yes	Yes
Individual and year FE	Yes	Yes	Yes
LL / R-sq	-765,976.57	-527,561.97	0.045
Observations	231,263	231,263	231,263

Notes: The data is a panel at the author level based on publication data between 2007 and 2014. All models have robust standard errors clustered at the individual level. Columns 1 and 2 estimate a Poisson model. Column 3 estimates a linear regression model (OLS). *significant at 10%, ** at 5%, *** at 1%

Table 4.b: Specialists in motion-sensing experience a higher level of change in diversification after Kinect

Treated as within-area researchers. Controls determined through Coarsened Exact Matching (CEM)			
Diversification	HDP topic count (intensive)	HDP topic count (extensive)	Euclidean diversification using IEEE taxonomy
Fraction of MS pubs	0.195***	0.146***	0.064***
before x AfterKinect	(0.042)	(0.030)	(0.009)
Quadratic age	Yes	Yes	Yes
Individual and year FE	Yes	Yes	Yes
LL / R-sq	-765,971.96	-527,554.55	0.045
Observations	231,263	231,263	231,263

Notes: The data is a panel at the author level based on publication data between 2007 and 2014. All models have robust standard errors clustered at the individual level. Columns 1 and 2 estimate a Poisson model. Column 3 estimates a linear regression model (OLS). *significant at 10%, ** at 5%, *** at 1%

Table 5.a: Estimated changes in the level of diversification of outside-area researchers

Treated as outside-area researchers. Controls determined through Coarsened Exact Matching (CEM)			
Diversification	HDP topic count (intensive)	HDP topic count (extensive)	Euclidean diversification using IEEE taxonomy
Treated x AfterKinect	0.500***	0.358***	0.103***
	(0.019)	(0.013)	(0.004)
Quadratic age	Yes	Yes	Yes
Individual and year FE	Yes	Yes	Yes
LL / R-sq	-537,282.06	-373,937.94	0.048
Observations	165,868	165,868	166,212

Notes: The data is a panel at the author level based on publication data between 2007 and 2014. All models have robust standard errors clustered at the individual level. Columns 1 and 2 estimate a Poisson model. Column 3 estimates a linear regression model (OLS). *significant at 10%, ** at 5%, *** at 1%

Table 5.b: The change in diversification of outside-area researchers persists outside the set of newly added motion-sensing publications

Treated as outside-area researchers. Controls determined through Coarsened Exact Matching (CEM)			
Diversification	HDP topic count (intensive)	HDP topic count (extensive)	Euclidean diversification using IEEE taxonomy
Treated x AfterKinect	0.310***	0.191***	0.048***
	(0.021)	(0.015)	(0.005)
Quadratic age	Yes	Yes	Yes
Individual and year FE	Yes	Yes	Yes
LL / R-sq	-534,765.16	-372,990.12	0.046
Observations	165,844	165,844	166,212

Notes: The data is a panel at the author level based on publication data between 2007 and 2014. All models have robust standard errors clustered at the individual level. Columns 1 and 2 estimate a Poisson model. Column 3 estimates a linear regression model (OLS). *significant at 10%, ** at 5%, *** at 1%

Table 6.a: Estimated changes in the trajectory of research of within-area researchers

Treated as within-area researchers. Controls determined through Coarsened Exact Matching (CEM)			
Trajectory	Yearly new HDP topics by distance	Yearly new coauthors	Year new Publication outlets
Treated x AfterKinect	0.241**	0.051**	0.074***
	(0.096)	(0.020)	(0.017)
Quadratic age	Yes	Yes	Yes
Individual and year FE	Yes	Yes	Yes
R-sq / LL	0.028	-527,103.78	-208,371.97
Observations	231,266	230,770	231,026

Notes: The data is a panel at the author level based on publication data between 2007 and 2014. All models have robust standard errors clustered at the individual level. Column 1 estimates a linear regression model (OLS). Columns 2 and 3 estimate a Poisson model. *significant at 10%, ** at 5%, *** at 1%

Table 6.b: Specialists in motion-sensing experience a higher level of shift in their trajectory of research

Treated as within-area researchers. Controls determined through Coarsened Exact Matching (CEM)			
Trajectory	Yearly new HDP topics by distance	Yearly new coauthors	Year new Publication outlets
Fraction of MS pubs before x AfterKinect	0.859***	0.170***	0.228***
	(0.196)	(0.051)	(0.039)
Quadratic age	Yes	Yes	Yes
Individual and year FE	Yes	Yes	Yes
R-sq / LL	0.028	-527,090.78	-208,366.32
Observations	231,266	230,770	231,026

Notes: The data is a panel at the author level based on publication data between 2007 and 2014. All models have robust standard errors clustered at the individual level. Column 1 estimates a linear regression model (OLS). Columns 2 and 3 estimate a Poisson model. *significant at 10%, ** at 5%, *** at 1%

Table 7.a: Estimated changes in the trajectory of research of outside-area researchers

Treated as outside-area researchers. Controls determined through Coarsened Exact Matching (CEM)			
Trajectory	Yearly new HDP topics by distance	Yearly new coauthors	Year new Publication outlets
Treated x AfterKinect	1.962***	0.461***	0.448***
	(0.110)	(0.022)	(0.016)
Quadratic age	Yes	Yes	Yes
Individual and year FE	Yes	Yes	Yes
R-sq / LL	0.035	-372,894.72	-145,186.49
Observations	166,212	165,547	165,876

Notes: The data is a panel at the author level based on publication data between 2007 and 2014. All models have robust standard errors clustered at the individual level. Column 1 estimates a linear regression model (OLS). Columns 2 and 3 estimate a Poisson model. *significant at 10%, ** at 5%, *** at 1%

Table 7.b: The shift in the trajectory of outside-area researchers persists outside the set of newly added motion-sensing publications

Treated as outside-area researchers. Controls determined through Coarsened Exact Matching (CEM)			
Trajectory	Yearly new HDP topics by distance	Yearly new coauthors	Year new Publication outlets
Treated x AfterKinect	0.451***	0.261***	0.474***
	(0.052)	(0.024)	(0.016)
Quadratic age	Yes	Yes	Yes
Individual and year FE	Yes	Yes	Yes
R-sq / LL	0.028	-370,110.33	-145,752.72
Observations	166,212	165,515	165,876

Notes: The data is a panel at the author level based on publication data between 2007 and 2014. All models have robust standard errors clustered at the individual level. Column 1 estimates a linear regression model (OLS). Columns 2 and 3 estimate a Poisson model. *significant at 10%, ** at 5%, *** at 1%

APPENDIX A
Robustness to estimating using our main sample without employing CEM

All tables and figures based on regression estimates match the order of the CEM regression estimates included in the main text.

Figure A1. Change in publication counts, raw data, no CEM

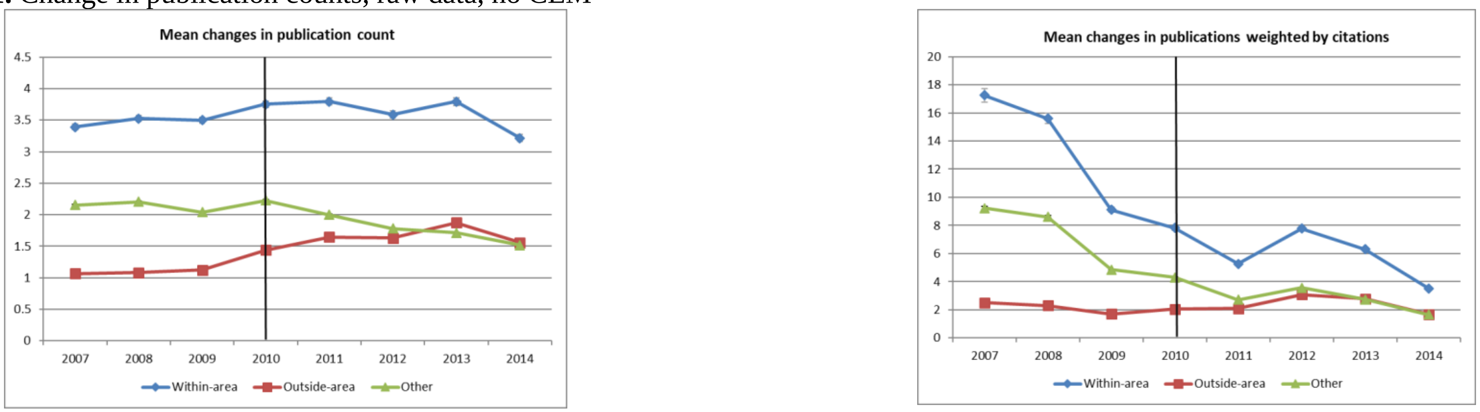


Figure A2. Change in the level of diversification, raw data, no CEM

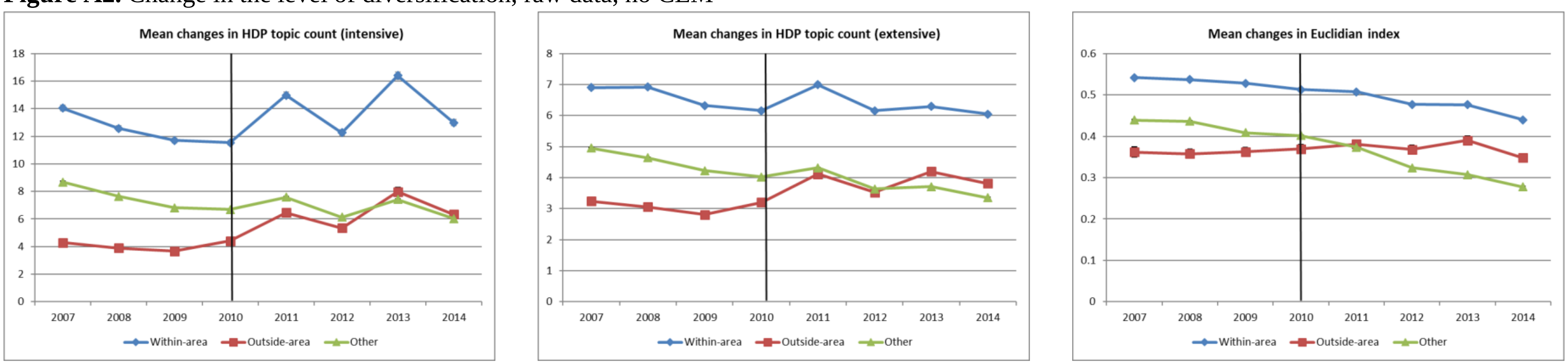


Figure A3. Change in the trajectory of research, raw data, no CEM

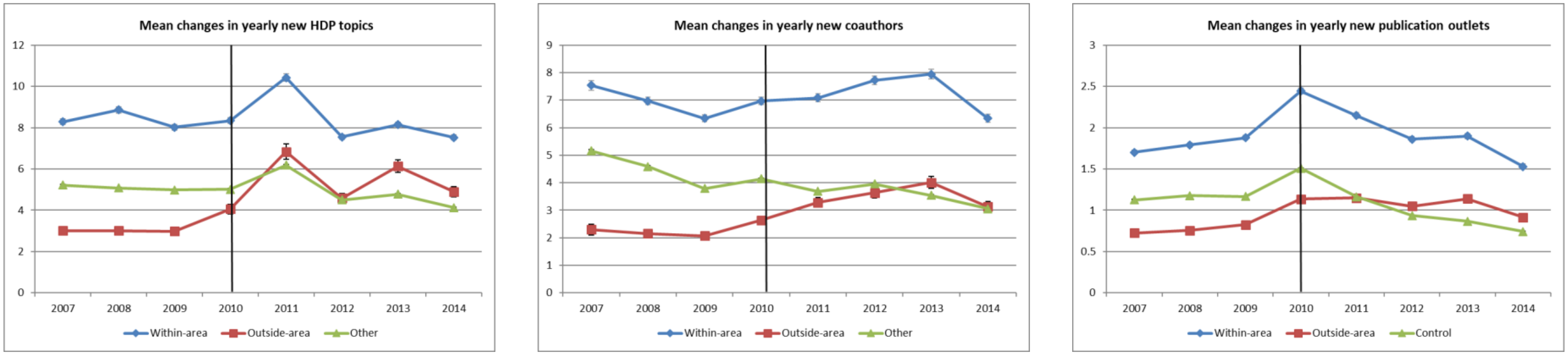
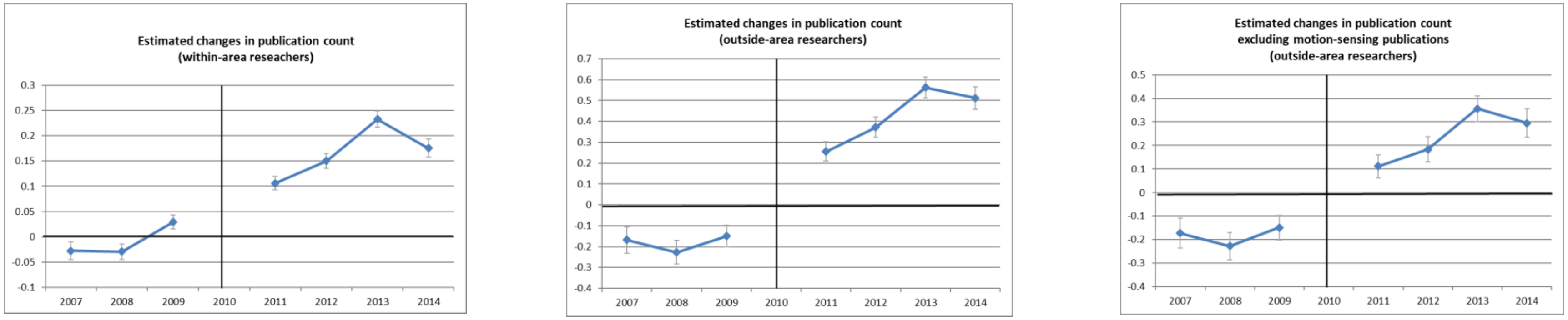
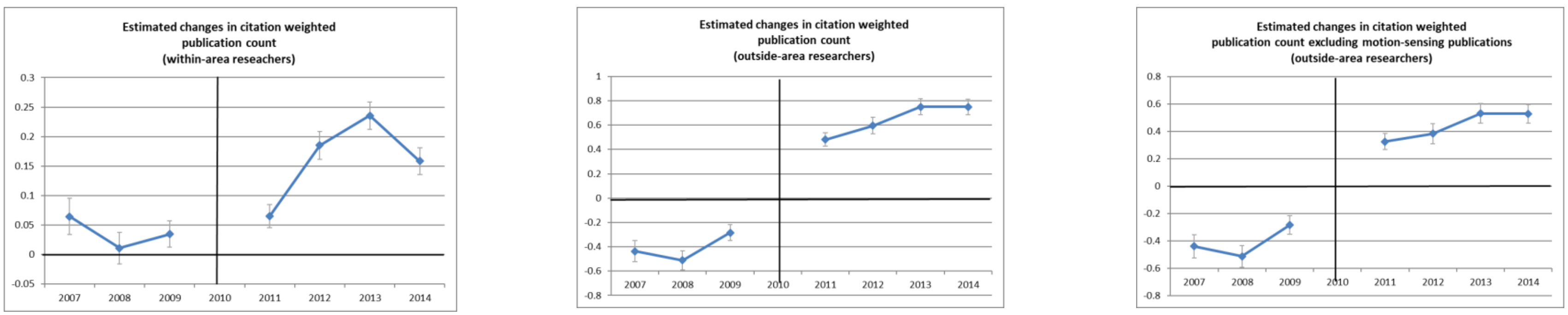


Figure A4.a: Estimated changes in yearly publication counts (original sample without CEM)



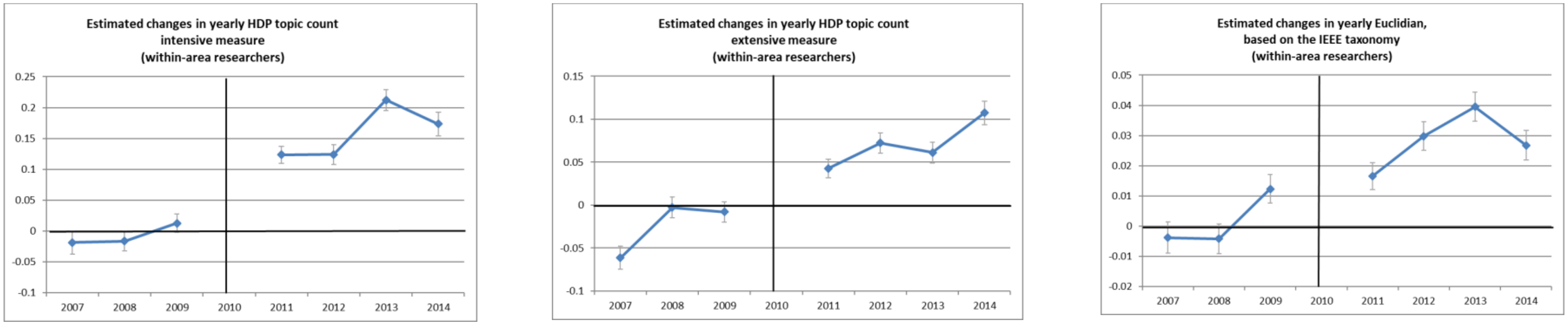
Notes: We base this figure on our 2007-2014 dataset. Each point on the graph represents the coefficient value on the covariate *TreatedResearcher x Year* and thus describes the relative difference in publication counts between treated and control authors in that year. The bars surrounding each point represent the 95 percent confidence interval. All values are relative to the base year of 2010.

Figure A4.b: Estimated changes in yearly publication counts weighted by citations (original sample without CEM)



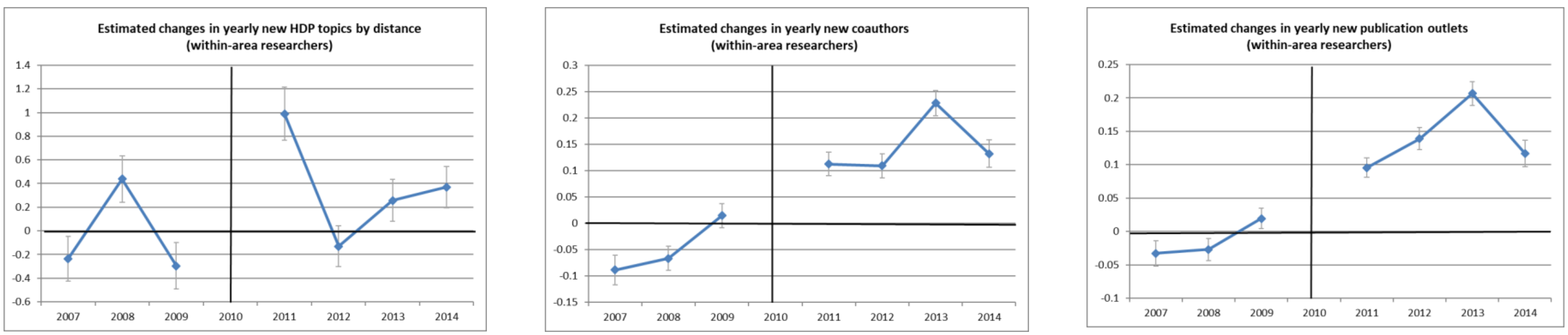
Notes: We base this figure on our 2007-2014 dataset. Each point on the graph represents the coefficient value on the covariate *TreatedResearcher x Year* and thus describes the relative difference in citation weighted publication counts between treated and control authors in that year. The bars surrounding each point represent the 95 percent confidence interval. All values are relative to the base year of 2010.

Figure A5.a: Estimated changes in the level of diversification of within-area researchers (original sample without CEM)



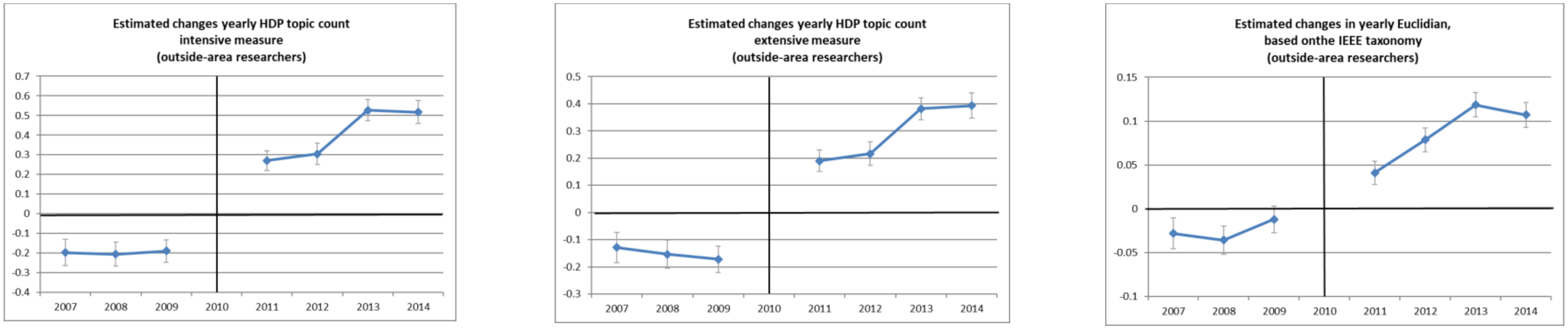
Notes: We base this figure on our 2007-2014 dataset. Each point on the graph represents the coefficient value on the covariate *TreatedResearcher x Year* and thus describes the relative difference in diversification between treated and control authors in that year. The bars surrounding each point represent the 95 percent confidence interval. All values are relative to the base year of 2010.

Figure A5.b: Estimated changes in the direction of within-area researchers (original sample without CEM)



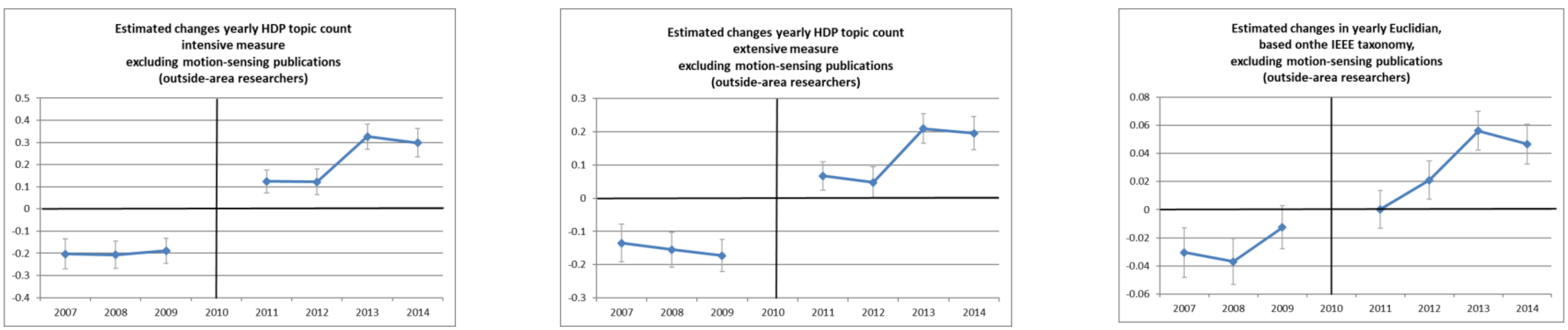
Notes: We base this figure on our 2007-2014 dataset. Each point on the graph represents the coefficient value on the covariate *TreatedResearcher x Year* and thus describes the relative difference in our direction measures between treated and control authors in that year. The bars surrounding each point represent the 95 percent confidence interval. All values are relative to the base year of 2010.

Figure A6.a: Estimated changes in the level of diversification of outside-area researchers (original sample without CEM)



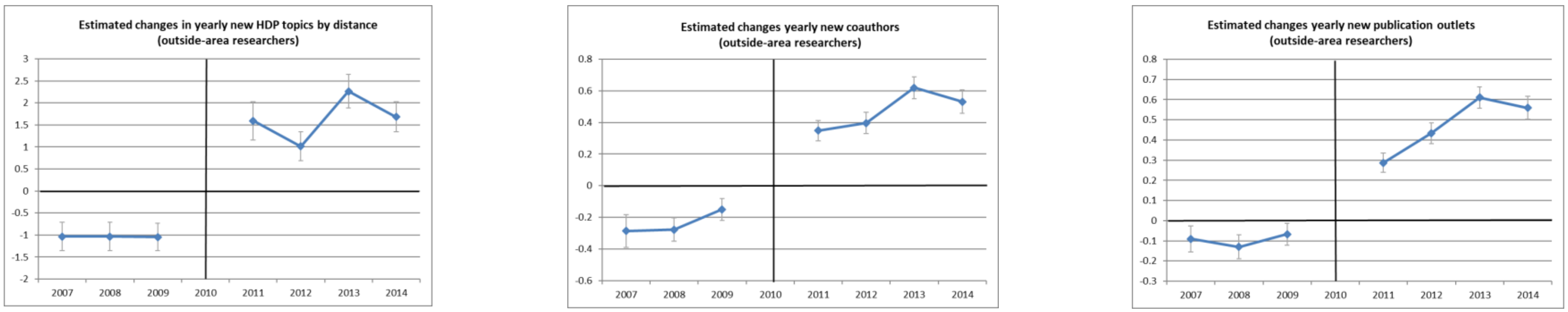
Notes: We base this figure on our 2007-2014 dataset. Each point on the graph represents the coefficient value on the covariate *TreatedResearcher x Year* and thus describes the relative difference in diversification between treated and control authors in that year. The bars surrounding each point represent the 95 percent confidence interval. All values are relative to the base year of 2010.

Figure A6.b: Estimated changes in the level of diversification of outside-area researchers, when excluding their motion-sensing publications (original sample without CEM)



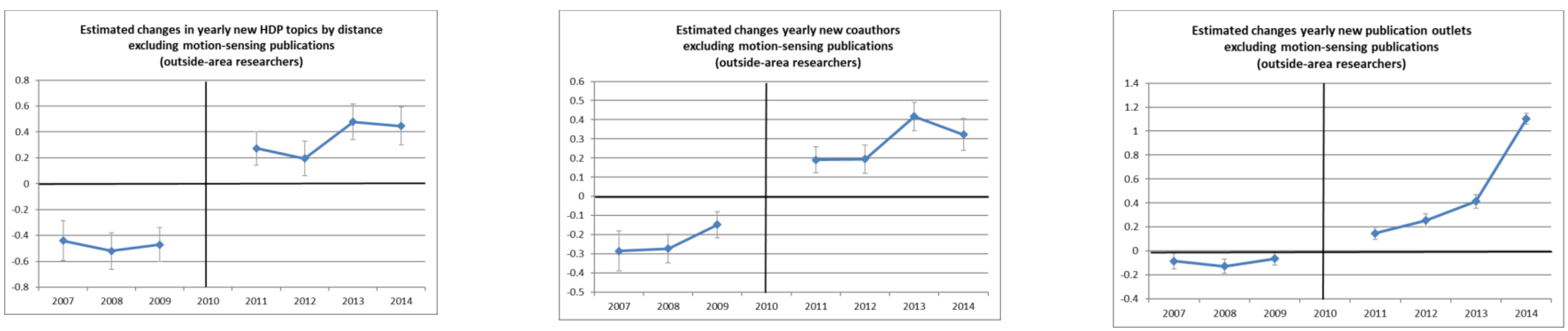
Notes: We base this figure on our 2007-2014 dataset. Each point on the graph represents the coefficient value on the covariate *TreatedResearcher x Year* and thus describes the relative difference in diversification between treated and control authors in that year. The bars surrounding each point represent the 95 percent confidence interval. All values are relative to the base year of 2010.

Figure A7.a: Estimated changes in the direction of outside-area researchers (original sample without CEM)



Notes: We base this figure on our 2007-2014 dataset. Each point on the graph represents the coefficient value on the covariate *TreatedResearcher x Year* and thus describes the relative difference in our direction measures between treated and control authors in that year. The bars surrounding each point represent the 95 percent confidence interval. All values are relative to the base year of 2010.

Figure A7.b: Estimated changes in the direction of outside-area researchers, when excluding their motion-sensing publications (original sample without CEM)



Notes: We base this figure on our 2007-2014 dataset. Each point on the graph represents the coefficient value on the covariate *TreatedResearcher x Year* and thus describes the relative difference in our direction measures between treated and control authors in that year. The bars surrounding each point represent the 95 percent confidence interval. All values are relative to the base year of 2010.

Table A1.a: Estimated changes in publication count

	No CEM		
	Count of publications		
	Within-area researchers	Outside-area researchers	Outside-area researchers (exclude motion-sensing papers)
Treated x AfterKinect	0.170***	0.544***	0.356***
	(0.006)	(0.016)	(0.019)
Quadratic age	Yes	Yes	Yes
Individual and year FE	Yes	Yes	Yes
LL	-1,922,713.80	-1,691,515.20	-1,690,117.20
Observations	1,386,630	1,258,131	1,258,131

Notes: The data is a panel at the author level based on publication data between 2007 and 2014. All models are Poisson with robust standard errors clustered at the individual level. *significant at 10%, **significant at 5%, ***significant at 1%

Table A1.b: Estimated changes in citation weighted publication count

	No CEM		
	Count of publications weighted by citations		
	Within-area researchers	Outside-area researchers	Outside-area researchers (exclude motion-sensing papers)
Treated x AfterKinect	0.134***	0.953***	0.749***
	(0.010)	(0.023)	(0.026)
Quadratic age	Yes	Yes	Yes
Individual and year FE	Yes	Yes	Yes
LL	-4,048,539.40	-3,480,732.90	-3,478,684.10
Observations	1,386,630	1,258,131	1,258,131

Notes: The data is a panel at the author level based on publication data between 2007 and 2014. All models are Poisson with robust standard errors clustered at the individual level. *significant at 10%, **significant at 5%, ***significant at 1%

Table A2.a: Estimated changes in the level of diversification of within-area researchers

Treated as within-area researchers. No CEM			
Diversification	HDP topic count (intensive)	HDP topic count (extensive)	Euclidean diversification using IEEE taxonomy
Treated x AfterKinect	0.166***	0.087***	0.027***
	(0.006)	(0.004)	(0.001)
Quadratic age	Yes	Yes	Yes
Individual and year FE	Yes	Yes	Yes
LL / R-sq	-5,240,308.20	-3,177,377.70	0.003
Observations	1,328,462	1,328,462	1,386,630

Notes: The data is a panel at the author level based on publication data between 2007 and 2014. All models have robust standard errors clustered at the individual level. Columns 1 and 2 estimate a Poisson model. Column 3 estimates a linear regression model (OLS). *significant at 10%, ** at 5%, *** at 1%

Table A2.b: Specialists in motion-sensing experience a higher level of change in diversification after Kinect

Treated as within-area researchers. No CEM			
Diversification	HDP topic count (intensive)	HDP topic count (extensive)	Euclidean diversification using IEEE taxonomy
Fraction of MS pubs	0.122***	0.048**	0.010*
before x AfterKinect	(0.028)	(0.019)	(0.006)
Quadratic age	Yes	Yes	Yes
Individual and year FE	Yes	Yes	Yes
LL / R-sq	-4,977,579.90	-3,040,236.60	0.003
Observations	1,328,462	1,328,462	1,386,630

Notes: The data is a panel at the author level based on publication data between 2007 and 2014. All models have robust standard errors clustered at the individual level. Columns 1 and 2 estimate a Poisson model. Column 3 estimates a linear regression model (OLS). *significant at 10%, ** at 5%, *** at 1%

Table A3.a: Estimated changes in the level of diversification of outside-area researchers

Treated as outside-area researchers. No CEM			
Diversification	HDP topic count (intensive)	HDP topic count (extensive)	Euclidean diversification using IEEE taxonomy
Treated x AfterKinect	0.549***	0.402***	0.104***
	(0.018)	(0.013)	(0.004)
Quadratic age	Yes	Yes	Yes
Individual and year FE	Yes	Yes	Yes
LL / R-sq	-4,649,391.30	-2,867,265.20	0.005
Observations	1,256,557	1,256,557	1,258,131

Notes: The data is a panel at the author level based on publication data between 2007 and 2014. All models have robust standard errors clustered at the individual level. Columns 1 and 2 estimate a Poisson model. Column 3 estimates a linear regression model (OLS). *significant at 10%, ** at 5%, *** at 1%

Table A3.b: The change in diversification of outside-area researchers persists outside the set of newly added motion-sensing publications

Treated as outside-area researchers. No CEM			
Diversification	HDP topic count (intensive)	HDP topic count (extensive)	Euclidean diversification using IEEE taxonomy
Treated x AfterKinect	0.362***	0.239***	0.049***
	(0.020)	(0.015)	(0.004)
Quadratic age	Yes	Yes	Yes
Individual and year FE	Yes	Yes	Yes
LL / R-sq	-4,647,004.50	-2,866,404.40	0.005
Observations	1,256,557	1,256,557	1,258,131

Notes: The data is a panel at the author level based on publication data between 2007 and 2014. All models have robust standard errors clustered at the individual level. Columns 1 and 2 estimate a Poisson model. Column 3 estimates a linear regression model (OLS). *significant at 10%, ** at 5%, *** at 1%

Table A4.a: Estimated changes in the trajectory of research of within-area researchers

Treated as within-area researchers. No CEM			
Trajectory	Yearly new HDP topics by distance	Yearly new coauthors	Year new Publication outlets
Treated x AfterKinect	0.396***	0.182***	0.146***
	(0.046)	(0.007)	(0.006)
Quadratic age	Yes	Yes	Yes
Individual and year FE	Yes	Yes	Yes
R-sq / LL	0.002	-4,064,300.50	-1,371,907.90
Observations	1,386,630	1,384,102	1,380,358

Notes: The data is a panel at the author level based on publication data between 2007 and 2014. All models have robust standard errors clustered at the individual level. Column 1 estimates a linear regression model (OLS). Columns 2 and 3 estimate a Poisson model. *significant at 10%, ** at 5%, *** at 1%

Table A4.b: Specialists in motion-sensing experience a higher level of shift in their trajectory of research

Treated as within-area researchers. No CEM			
Trajectory	Yearly new HDP topics by distance	Yearly new coauthors	Year new Publication outlets
Fraction of MS pubs	0.593***	0.277***	0.189***
before x AfterKinect	(0.145)	(0.035)	(0.027)
Quadratic age	Yes	Yes	Yes
Individual and year FE	Yes	Yes	Yes
R-sq / LL	0.001	-3,848,520.50	-1,295,488.30
Observations	1,386,630	1,384,102	1,380,358

Notes: The data is a panel at the author level based on publication data between 2007 and 2014. All models have robust standard errors clustered at the individual level. Column 1 estimates a linear regression model (OLS). Columns 2 and 3 estimate a Poisson model. *significant at 10%, ** at 5%, *** at 1%

Table A5.a: Estimated changes in the trajectory of research of outside-area researchers

Treated as outside-area researchers. No CEM			
Trajectory	Yearly new HDP topics by distance	Yearly new coauthors	Year new Publication outlets
Treated x AfterKinect	2.376***	0.637***	0.521***
	(0.093)	(0.021)	(0.015)
Quadratic age	Yes	Yes	Yes
Individual and year FE	Yes	Yes	Yes
R-sq / LL	0.001	-3,585,172.90	-1,205,082.80
Observations	1,258,131	1,255,701	1,252,043

Notes: The data is a panel at the author level based on publication data between 2007 and 2014. All models have robust standard errors clustered at the individual level. Column 1 estimates a linear regression model (OLS). Columns 2 and 3 estimate a Poisson model. *significant at 10%, ** at 5%, *** at 1%

Table A5.b: The shift in the trajectory of outside-area researchers persists outside the set of newly added motion-sensing publications

Treated as outside-area researchers. No CEM			
Trajectory	Yearly new HDP topics by distance	Yearly new coauthors	Year new Publication outlets
Treated x AfterKinect	0.688***	0.443***	0.549***
	(0.044)	(0.024)	(0.016)
Quadratic age	Yes	Yes	Yes
Individual and year FE	Yes	Yes	Yes
R-sq / LL	0.001	-3,582,461.10	-1,205,873.40
Observations	1,258,131	1,255,669	1,252,043

Notes: The data is a panel at the author level based on publication data between 2007 and 2014. All models have robust standard errors clustered at the individual level. Column 1 estimates a linear regression model (OLS). Columns 2 and 3 estimate a Poisson model. *significant at 10%, ** at 5%, *** at 1%

APPENDIX B
CEM Balance

Table B1.a: CEM balance, where treated researchers are defined as scientists who published in motion-sensing before Kinect (within-area researchers)

	CEM balance					
	Full Sample			Matched Sample (CEM)		
	Treated	Controls	t-stat	Treated	Controls	t-stat
Citation-weighted publication count 2007	15.868	7.397	87.83	2.323	2.300	0.61
Citation-weighted publication count 2008	14.609	7.802	75.44	2.664	2.615	1.36
Citation-weighted publication count 2009	8.687	4.729	91.31	1.957	1.918	1.73
Citation-weighted publication count 2010	7.287	4.301	87.20	2.220	2.190	1.32
Author count 2007	15.212	9.052	70.02	3.609	3.575	0.73
Author count 2008	15.253	9.890	70.80	4.318	4.329	0.24
Author count 2009	15.290	9.517	89.36	4.585	4.602	0.37
Author count 2010	16.362	10.528	85.91	6.028	5.989	0.67
Diversification index 2007-2010	0.687	0.640	137.17	0.661	0.661	0.42
Observations	100,392	1,286,760		26,336	223,872	

Table B1.b: CEM balance, where treated researchers are defined as scientists who published in motion-sensing after Kinect (outside-area researchers)

	CEM balance					
	Full Sample			Matched Sample (CEM)		
	Treated	Controls	t-stat	Treated	Controls	t-stat
Citation-weighted publication count 2007	11.683	7.397	41.18	1.917	1.888	0.79
Citation-weighted publication count 2008	11.784	7.802	39.81	1.999	2.012	0.39
Citation-weighted publication count 2009	7.252	4.729	52.70	1.630	1.626	0.19
Citation-weighted publication count 2010	6.712	4.301	62.89	2.053	2.038	0.58
Author count 2007	12.314	9.052	32.99	3.419	3.342	0.83
Author count 2008	13.035	9.890	36.89	3.649	3.721	1.36
Author count 2009	13.445	9.517	54.08	4.123	4.122	0.02
Author count 2010	15.150	10.528	60.50	5.601	5.620	0.31
Diversification index 2007-2010	0.677	0.640	94.98	0.650	0.650	0.31
MS Distance	2.651	2.955	76.72	3.068	3.068	0.00
Observations	76,720	1,286,760		23,872	158,016	

APPENDIX C
Difference in magnitude of effects for within-area and outside-area researchers

Tables C1: Estimated changes in publication count and citation weighted publication count

Controls determined through Coarsened Exact Matching (CEM)		
	Productivity of within-area and outside-area researchers	
	Count of publications	Count of publications
Within-area x AfterKinect	0.182***	0.166***
	(0.017)	(0.022)
Outside-area x AfterKinect	0.299***	0.415***
	(0.023)	(0.031)
Quadratic age	Yes	Yes
Individual and year FE	Yes	Yes
LL	-396,598.41	-625,301.33
Observations	342,223	342,223

Notes: The data is a panel at the author level based on publication data between 2007 and 2014. All models are Poisson with robust standard errors clustered at the individual level. *significant at 10%, **significant at 5%, ***significant at 1%

Tables C2: Estimated changes in the level of diversification

Controls determined through Coarsened Exact Matching (CEM)			
	Diversification		
	HDP topic count (intensive)	HDP topic count (extensive)	Euclidean diversification using IEEE taxonomy
Within-area x AfterKinect	0.177***	0.116***	0.041***
	(0.018)	(0.013)	(0.004)
Outside-area x AfterKinect	0.303***	0.230***	0.063***
	(0.025)	(0.018)	(0.005)
Quadratic age	Yes	Yes	Yes
Individual and year FE	Yes	Yes	Yes
LL / R-sq	-1,129,953.00	-775,453.09	0.043
Observations	341,600	341,600	342,223

Notes: The data is a panel at the author level based on publication data between 2007 and 2014. All models have robust standard errors clustered at the individual level. Columns 1 and 2 estimate a Poisson model. Column 3 estimates a linear regression model (OLS). *significant at 10%, **significant at 5%, ***significant at 1%

Tables C3: Estimated changes in the trajectory of research

Controls determined through Coarsened Exact Matching (CEM)			
	Research trajectory		
	Yearly new HDP topics by distance	Yearly new coauthors	Year new Publication outlets
Within-area x AfterKinect	0.518***	0.133***	0.155***
	(0.092)	(0.020)	(0.016)
Outside-area x AfterKinect	1.427***	0.326***	0.299***
	(0.125)	(0.028)	(0.022)
Quadratic age	Yes	Yes	Yes
Individual and year FE	Yes	Yes	Yes
R-sq / LL	0.030	-769,345.22	-304,278.57
Observations	342,223	341,045	341,351

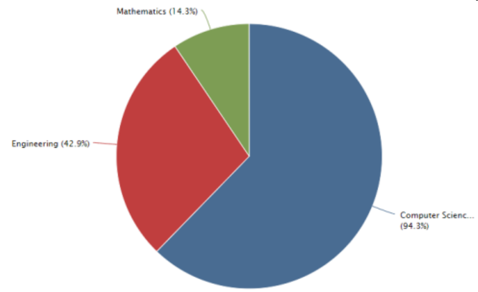
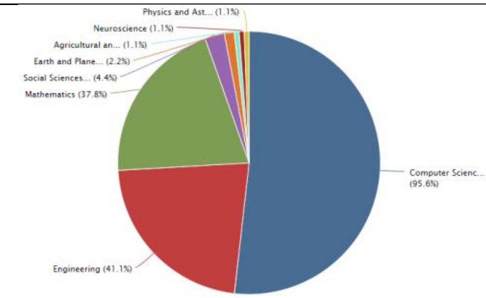
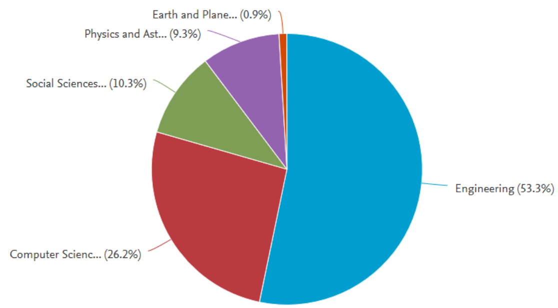
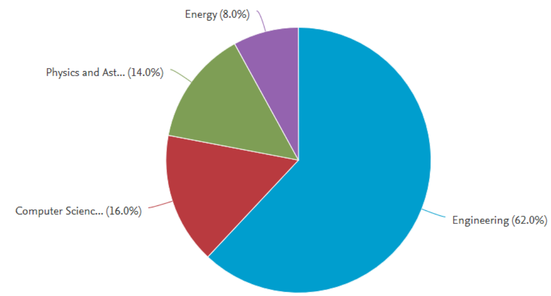
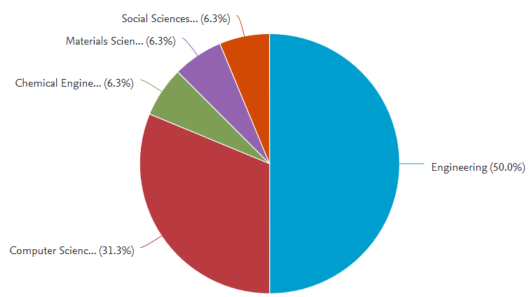
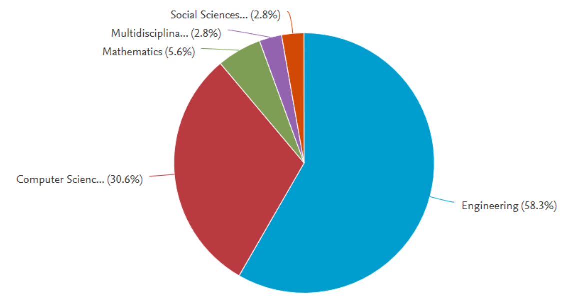
Notes: The data is a panel at the author level based on publication data between 2007 and 2014. All models have robust standard errors clustered at the individual level. Columns 1 and 2 estimate a Poisson model. Column 3 estimates a linear regression model (OLS). *significant at 10%, **significant at 5%, ***significant at 1%

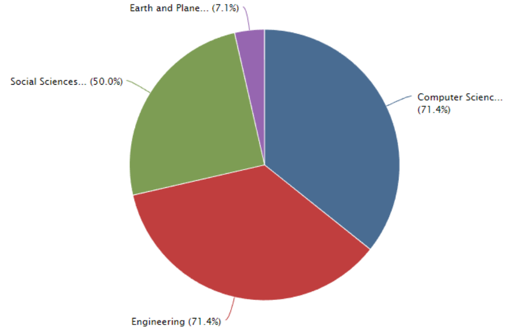
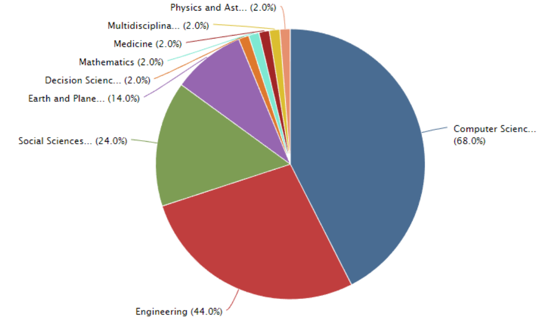
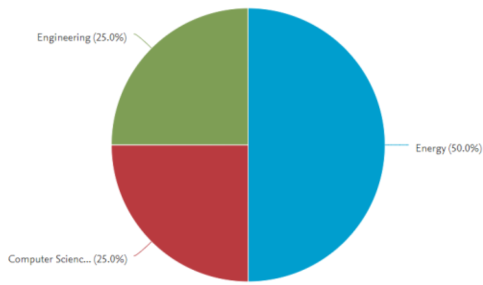
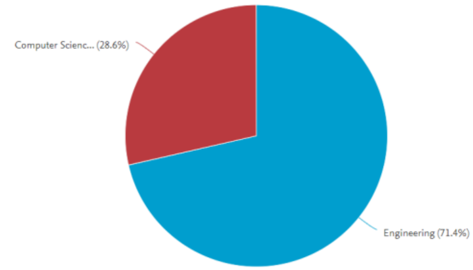
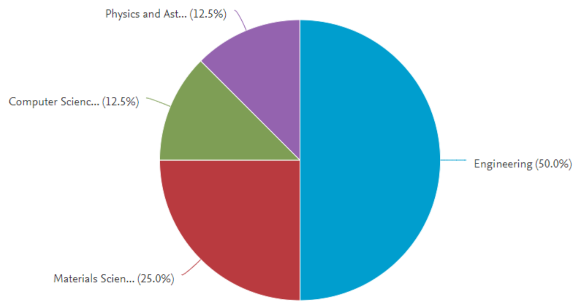
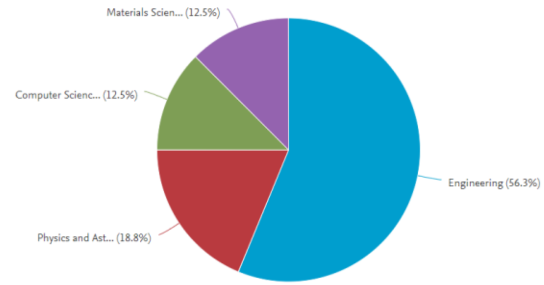
APPENDIX D

Examples of changes in publication portfolio distribution across research topics, as defined by Scopus

(names not shown due to privacy concerns)

In order to provide additional suggestive evidence on the change in research trajectories following the Kinect shock, we leverage data from Scopus. Specifically, we select a random group of researchers from our treated and control samples and look up their publication portfolio in Scopus, a comprehensive bibliographical database tracking academic publication across all areas of science. We take advantage of a particular feature of this database, namely its "analyze" function. Scopus allows users to select a group of papers and to analyze its properties using a proprietary Scopus algorithm. The analysis displays, among other things, pie charts with the distribution of publications across knowledge areas, as defined by the Scopus taxonomy. Since the data in our main analysis derives from a different database (IEEE *Xplore*), which processes and classifies papers with no relation to Scopus, our analysis based on the Scopus analysis algorithm provides an additional piece of evidence to our findings. We include below examples of randomly-selected treated and control researchers and the pie-chart analysis of their pre- and post-Kinect publication portfolios. These charts suggest a change in the composition of project portfolios following the launch of Kinect that is consistent with our econometric analysis and that provide a glimpse into the type of knowledge these researchers create.

<table>
<tr><th colspan="3">Example of within-area researcher changes in diversification and trajectory, relative to weighted CEM controls</th></tr>
<tr><th></th><th>2007-2010</th><th>2011-2014</th></tr>
<tr><td>Within-area researcher</td><td>
</td><td>
</td></tr>
<tr><td>Control for above within-area researcher (weighted CEM)</td><td>
</td><td>
</td></tr>
<tr><td>Control for above within-area researcher (weighted CEM)</td><td>
</td><td>
</td></tr>
</table>

Example of outside-area researcher changes in diversification and trajectory, relative to CEM matched controls		
	2007-2010	**2011-2014**
Outside-area researcher		
Control for above outside-area researcher (weighted CEM)		
Control for above outside-area researcher (weighted CEM)		

APPENDIX E
Institutions that benefited most from the Kinect phenomenon

To investigate which institutions experience the greatest publication boost following the Kinect shock, we leverage the Analyze feature of the Scopus online database. Specifically, we use the Scopus analyze feature to collect motion-sensing publication counts for the period before and after the launch of Kinect, by the affiliation institutions of the authors of these publications. We start by searching for motion-sensing academic publications using the same set of keywords used to determine motion-sensing publications in our main IEEE dataset. Next, we split the search into two groups: the group of motion-sensing papers published in the four years before the launch of Kinect (2007-2010) and that of such papers published in the four years after the launch (2011-2014). For the latter group, we further restrict the set of motion-sensing papers to publications in the top 25% and top 10%, respectively, of the most cited motion-sensing papers during the pre-Kinect period. We do so to make sure we capture institutions that produced quality output in motion-sensing as a result of the Kinect phenomenon. The results do not change substantially when considering the full set of motion-sensing papers, regardless of citation ranking. Last, we extract the count of such motion-sensing publications per affiliation institution of motion-sensing authors in both periods. Scopus counts a publication multiple times, as many as the number of authors on each publication. Unfortunately, the Scopus feature does not facilitate options to also consider additional information when generating these counts, such as order of authors or the number of authors per publication.

In the table below, we display the top 50 institutions that published in motion-sensing after Kinect, as per our citations criteria. In fact, the table lists 51 institutions, with Microsoft as the top institution that benefited most from the launch of Kinect. While this is reassuring, it does not provide generalizable insights to the type of institutions that would benefit most from availability of automating research technology. Thus, we chose to include Microsoft in addition to our list of top 50 institutions that benefited most from Kinect. Columns 4 to 7 list counts of motion-sensing papers in the top 25% and top 10% of most cited motion-sensing papers, before and after the launch of Kinect. We sort the list by the average change in number of publications in the two citations groups, however, other sort orders preserve the picture painted here. Columns 8 and 9 list two sets of rankings for the 50 institutions. In column 8, we follow the Times Higher Education university ranking from 2013-2014. In column 9, we present values from another internationally recognized ranking system, the QS World University Ranking, which has the advantage to offer a specialized ranking for universities in computer science and engineering. We utilize the 2015 ranking; we could not locate an earlier ranking but believe the time difference does not carry substantial implications.

We observe that the list of institutions that benefited most from the Kinect phenomenon includes a mixture of top ranked universities and well-known private institutions, and less prestigious universities. This is important because, on the one hand, the benefit of Kinect as technology automating research tasks might manifest more for research institutions that are or aim to be leaders in research involving motion-sensing topics. The assumption is that these institutions are time constrained, rather than financially constrained, and the automation would help propel their productivity forward. On the other hand, the benefit of Kinect as technology that reduces the costs of performing certain research tasks might manifest more for institutions that are financially constrained. It follows that a research technology that is automating by substituting for human capital and hence significantly reducing the cost of performing certain tasks should benefit both types of institutions, which is what we observe.

	Institution Name	**Country**	**Count publications in top 25% cited (2007-2010)**	**Count publications in top 25% cited (2011-2014)**	**Count publications in top 10% cited (2007-2010)**	**Count publications in top 10% cited (2011-2014)**	**Ranking Times Higher Education[24]**	**Ranking QS World University for Computer Science and Engineering[25]**
1	Microsoft Research	USA	9	117	4	84	n/a	n/a
2	Intel Corporation	USA	2	23	1	21	n/a	n/a
3	University of Missouri-Columbia	USA	6	18	1	14	301-350	Not ranked
4	Tsinghua University	China	2	15	1	9	50	11
5	Universitat Bonn	Germany	2	20	2	12	181	324
6	University of Washington, Seattle	USA	6	38	4	29	25	85
7	Technische Universitat Wien	Austria	2	15	1	6	226-250	93
8	KU Leuven	Belgium	2	10	1	8	61	74
9	Universitat Freiburg im Breisgau	Germany	7	20	1	10	152	293

[24] https://www.timeshighereducation.com/world-university-rankings/2014/world-ranking#!/page/0/length/25/sort_by/scores_overall/sort_order/asc/cols/undefined
[25] https://www.topuniversities.com/university-rankings/world-university-rankings/2015

10	South China University of Technology	China	2	11	1	7	Not ranked	Not ranked
11	National University of Singapore	Singapore	10	25	2	19	26	3
12	CNRS Centre National de la Recherche Scientifique	France	5	25	2	13	n/a	n/a
13	Scuola Superiore Sant'Anna di Studi Universitari e di Perfezionamento	Italy	2	13	1	5	Not ranked	Not ranked
14	Universita di Pisa	Italy	2	11	1	6	301-350	208
15	Zhejiang University	China	11	24	1	9	301-350	70
16	Cornell University	USA	2	12	2	10	19	34
17	Columbia University in the City of New York	USA	2	10	1	6	13	56
18	City University of Hong Kong	China	2	14	1	3	201-225	60
19	UCL	UK	3	16	2	9	21	50
20	University of Sheffield	UK	4	11	1	7	112	112
21	Imperial College London	UK	16	22	2	16	10	7
22	Korea Advanced Institute of Science & Technology	South Korea	2	10	1	4	56	13
23	University of Texas at Arlington	USA	9	12	1	7	Not ranked	Not ranked
24	University of Southern California	USA	9	29	3	15	70	128
25	University of California, Los Angeles	USA	2	11	2	5	12	23
26	Trinity College Dublin	Ireland	5	10	1	6	129	100

27	Nanyang Technological University	Singapore	7	25	3	13	76	6
28	University of Science and Technology of China	China	8	15	1	6	201-225	93
29	Northwestern University	USA	4	10	1	5	22	61
30	Chinese Academy of Sciences	China	11	31	4	18	n/a	n/a
31	University of Michigan, Ann Arbor	USA	4	17	3	9	18	36
32	Technical University of Munich	Germany	13	34	5	21	87	30
33	University of California, Berkeley	USA	5	24	9	17	8	8
34	Harvard University	USA	3	12	3	8	2	10
35	Istituto Italiano di Tecnologia	Italy	2	13	1	0	n/a	n/a
36	Università degli Studi di Padova	Italy	4	15	1	2	301-350	199
37	University of Waterloo	Canada	7	10	1	4	226-250	74
38	Georgia Institute of Technology	USA	7	22	7	11	28	19
39	University of Illinois at Urbana-Champaign	USA	5	15	6	9	29	30
40	Texas A and M University	USA	8	12	3	9	159	30
41	Karlsruhe Institute of Technology	Germany	4	12	5	6	154	62
42	The Walt Disney Company	USA	8	14	3	6	n/a	n/a
43	Seoul National University	South Korea	6	12	4	7	44	15
44	Stanford University	USA	11	31	19	14	4	2
45	The University of North	USA	8	12	3	6	47	224

	Carolina at Chapel Hill							
46	The Royal Institute of Technology KTH	Sweden	4	10	3	3	117	36
47	Massachusetts Institute of Technology	USA	24	40	21	29	5	1
48	ETH Zurich	Switzerland	28	37	14	21	14	5
49	Friedrich-Alexander-Universität Erlangen-Nürnberg	Germany	7	10	5	6	Not ranked	214
50	Carnegie Mellon University	USA	41	52	26	27	24	29
51	Universität Bielefeld	Germany	6	10	1	0	301-350	Not ranked